CrORR

Croatian Operational Research Review

vol. 17 / no. 2 / 2026

ISSN (print): 1848 – 0225

ISSN (online): 1848 – 9931

CRORS

CROATIAN OPERATIONAL RESEARCH SOCIETY

ISSN: 1848–0225 Print
ISSN: 1848–9931 Online
UDC: 519.8 (063)
Abbreviation: Croat. Oper. Res. Rev.

Publisher:
CROATIAN OPERATIONAL RESEARCH SOCIETY

Co-publishers:
University of Zagreb, Faculty of Economics & Business
J. J. Strossmayer University of Osijek, Faculty of Economics in Osijek
University of Split, Faculty of Economics, Business and Tourism
J. J. Strossmayer University of Osijek, School of Applied Mathematics and Informatics
University of Zagreb, Faculty of Organization and Informatics

Croatian Operational Research Review

Volume 17 (2026)
Number 2
pp. 181-354

Zagreb, 2026

Croatian Operational Research Review is viewed/indexed in: Current Index to Statistics, DOAJ, EBSCO host, EconLit, Genamics Journal Seek, INSPEC, Mathematical Reviews, Current Mathematical Publications (MathSciNet), Proquest, SCOPUS, Web of Science Emerging Sources Citation Index (WoS ESCI), Zentralblatt für Mathematik/Mathematics Abstracts (CompactMath).

ISSN: 1848–0225 Print
ISSN: 1848–9931 Online
UDC: 519.8 (063)
Abbreviation: Croat. Oper. Res. Rev.
Edition: 40 copies

Croatian Operational Research Review

Croatian Operational Research Review (CRORR) is an open access scientific journal published bi-annually. Starting from 2014, Number 1, the main publisher of the CRORR journal is the Croatian Operational Research Society (CRORS). Co-publishers of the journal are:

- ◇ University of Zagreb, Faculty of Economics & Business,
- ◇ J. J. Strossmayer University of Osijek, Faculty of Economics in Osijek,
- ◇ University of Split, Faculty of Economics, Business and Tourism,
- ◇ J. J. Strossmayer University of Osijek, School of Applied Mathematics and Informatics and
- ◇ University of Zagreb, Faculty of Organization and Informatics.

Aims and Scope of the Journal

The aim of the Croatian Operational Research Review journal is to provide high quality scientific papers covering the theory and application of operational research and related areas, mainly quantitative methods and machine learning. The scope of the journal is focused, but not limited to the following areas: combinatorial and discrete optimization, integer programming, linear and nonlinear programming, multiobjective and multicriteria programming, statistics and econometrics, macroeconomics, economic theory, games control theory, stochastic models and optimization, banking, finance, insurance, simulations, information and decision support systems, data envelopment analysis, neural networks and fuzzy systems, and practical OR and application.
Occasionally, special issues are published dedicated to a particular area of OR or containing selected papers from the International Conference on Operational Research.
Papers should be written according to journal guidelines available at https://hrcak.srce.hr/ojs/index.php/crorr/about/submissions#authorGuidelines. Papers blindly reviewed and accepted by two independent reviewers are published in this journal.

Publication Ethics and Publication Malpractice Statement

CRORR journal is committed to ensuring ethics in publications and quality of articles. Conformance to standards of ethical behaviour is therefore expected of all parties involved: Authors, Editors, Reviewers, and the Publisher. Before submitting or reviewing a paper read our Publication ethics and Publication Malpractice Statement available at the journal homepage: http://www.hdoi.hr/crorr-journal.
The journal is financially supported by the Ministry of Science, Education and Sports of the Republic of Croatia, and by its co-publishers. All data referring to the journal with full papers texts since the beginning of its publication in 2010 can be found in the Hrcak database – Portal of scientific journals of Croatia (see http://hrcak.srce.hr./crorr).

Open Access Statement

Croatian Operational Research Review is an open access journal which means that all content is freely available without charge to the user or his/her institution. Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the articles in the journal without asking prior permission from the publisher or the author. This is in accordance with the BOAI definition of open access. The journal has an open access to full text of all papers through our Open Journal System (http://hrcak.srce.hr/ojs/index.php/crorr) and the Hrcak database (http://hrcak.srce.hr/crorr).

Croatian Operational Research Review is viewed/indexed by:

- ◇ Current Index to Statistics
- ◇ Directory of Open Access Journals (DOAJ)
- ◇ EBSCO host
- ◇ EconLit
- ◇ Genamics Journal Seek
- ◇ INSPEC
- ◇ Mathematical Reviews, Current Mathematical Publications (MathSciNet)
- ◇ Proquest
- ◇ SCOPUS
- ◇ Web of Science Emerging Sources Citation Index (WoS ESCI)
- ◇ Zentralblatt für Mathematik/Mathematics Abstracts (CompactMath)

Editorial Office and Address:

Editors-in-Chief:	Blanka Škrabić Perić Tea Šestanović
Co-Editors:	Mateja Đumić Danijel Grahovac Goran Lešaja Branka Marasović Ivan Papić Snježana Pivac Silvija Vlah Jerić
Technical Editors:	Ivana Mravak Tea Kalinić Milićević Marija Vuković
Design:	Tea Kalinić Milićević
Cover image:	www.123rf.com

Contact:
Croatian Operational Research Society (CRORS), Trg J. F. Kennedyja 6,
HR-10000 Zagreb, Croatia
URL: http://www.hdoi.hr/crorr-journal
e-mail: crorr.journal@gmail.com

Editorial Board

Republic of North Macedonia, e-mail: vcvetkoska@eccf.ukim.edu.mk
Maja Ćukušić, Faculty of Economics, Business and Tourism, University of Split, Croatia, e-mail: maja.cukusic@efst.hr
Anita Čeh Časni, Faculty of Economics & Business, University of Zagreb, Croatia, e-mail: aceh@efzg.hr
Mirjana Čižmešija, Faculty of Economics & Business, University of Zagreb, Croatia, e-mail: mcizmesija@efzg.hr
Ozren Despić, Aston Business School, College of Business and Social Sciences, UK, e-mail: o.despic@aston.ac.uk
Mateja Đumić, School of Applied Mathematics and Informatics, J. J. Strossmayer University of Osijek, Croatia, e-mail: mdjumic@mathos.hr
Margareta Gardijan Kedžo, Faculty of Economics & Business, University of Zagreb, Croatia, e-mail: mgardijan@efzg.hr
Danijel Grahovac, School of Applied Mathematics and Informatics, J. J. Strossmayer University of Osijek, Croatia, e-mail: dgrahova@mathos.hr
Tihomir Hunjak, Faculty of Organization and Informatics, University of Zagreb, Croatia, e-mail: tihomir.hunjak@foi.hr
Mario Jadrić, Faculty of Economics, Business and Tourism, University of Split, Croatia, e-mail: jadric@efst.hr
Nikola Kadoić, Faculty of Organization and Informatics, University of Zagreb, Croatia, e-mail: nikola.kadoic@foi.unizg.hr
Vedran Kojić, Faculty of Economics & Business, University of Zagreb, Croatia, e-mail: vkojic@efzg.hr
Ulrike Leopold-Wildburger, Institut für Statistik un Operations Research, Karl-Franzens-Universität Graz, Austria, e-mail: ulrike.leopold@uni-graz.at
Goran Lešaja, Department of Mathematical Sciences, Georgia Southern University, USA, e-mail: goran@georgiasouthern.edu
Zrinka Lukač, Faculty of Economics & Business, University of Zagreb, Croatia, e-mail: zlukac@efzg.hr
Robert Manger, Department of Mathematics, University of Zagreb, Croatia, e-mail: manger@math.hr
Branka Marasović, Faculty of Economics, Business and Tourism, University of Split, Croatia, e-mail: branka.marasovic@efst.hr
Milica Maričić, Faculty of Organisational Sciences, Serbia, University of Belgrade, e-mail: milica.maricic@fon.bg.ac.rs
Markéta Matulová, Department of Applied Mathematics and Computer Science, Faculty of Economics and Administration, Czech Republic, e-mail: Marketa.Matulova@econ.muni.cz
Josip Mesarić, Faculty of Economics in Osijek, J. J. Strossmayer University of Osijek, Croatia, e-mail: mesaric@efos.hr
Tea Mijač, Faculty of Economics, Business and Tourism, University of Split, Croatia, e-mail: tea.mijac@efst.hr
Ivan Papić, School of Applied Mathematics and Informatics, J. J. Strossmayer University of Osijek, Croatia, e-mail: ipapic@mathos.hr
Snježana Pivac, Faculty of Economics, Business and Tourism, University of Split, Croatia, e-mail: spivac@efst.hr
Krunoslav Puljić, Faculty of Economics & Business, University of Zagreb, Croatia, e-mail: kpuljic@efzg.hr
Jose Rui Figueria, Instituto Superior Tecnico, Technical University of Lisbon, Portugal, e-mail: figueira@ist.utl.pt
Kristian Sabo, School of Applied Mathematics and Informatics, J. J. Strossmayer University of Osijek, Croatia, e-mail: ksabo@mathos.hr

Rudolf Scitovski, School of Applied Mathematics and Informatics, J. J. Strossmayer University of Osijek, Croatia, e-mail: scitowsk@mathos.hr
Marc Sevaux, Université de Bretagne-Sud, Lorient, France, e-mail:
marc.sevaux@univ-ubs.fr
Kenneth Sörensen, Department of Engineering Management, University of Antwerp, Belgium, e-mail: kenneth.sorensen@uantwerpen.be
Petar Sorić, Faculty of Economics & Business, University of Zagreb, Croatia, e-mail:
psoric@efzg.hr
Greys Sošić, Information and Operations Management Department, The Marshall School of Business, University of Southern California, Los Angeles, CA, USA, e-mail:
sosic@marshall.usc.edu
Alemka Šegota, Faculty of Economics, University of Rijeka, Croatia, e-mail:
alemka.segota@efri.hr
Tea Šestanović, University of Split, Faculty of Economics, Business and Tourism, Croatia, e-mail: tea.sestanovic@efst.hr
Blanka Škrabić Perić, Faculty of Economics, Business and Tourism, University of Split, Croatia, e-mail: blanka.skrabic@efst.hr
Kristina Šorić, Rochester Institute of Technology Croatia, Croatia, e-mail:
ksoric@zsem.hr
Begoña Vitoriano, Complutense University of Madrid, Faculty of Mathematical Sciences, Spain, e-mail: bvitoriano@mat.ucm.es
Silvija Vlah Jerić, Faculty of Economics & Business, University of Zagreb, Croatia, e-mail:
svlah@net.efzg.hr
Richard Wendell, Joseph M. Katz Graduate School of Business, University of Pittsburgh, USA, e-mail: wendell@katz.pitt.edu
Lidija Zadnik Stirn, Biotechnical Faculty, University of Ljubljana, Slovenia, e-mail:
Lidija.Zadnik@bf.uni-lj.si
Sanjo Zlobec, Department of Mathematics and Statistics, McGill University, Canada, e-mail:
zlobec@math.mcgill.ca
Nikolina Žajdela Hrustek, University of Zagreb, Faculty of Organization and Informatics, Croatia, e-mail: nikolina.zajdela@foi.unizg.hr
Berislav Žmuk, Faculty of Economics & Business, University of Zagreb, Croatia, e-mail:
bzmuk@efzg.hr

ISSN: 1848–0225 Print
ISSN: 1848–9931 Online
UDC: 519.8 (063)

Croatian Operational Research Review

Volume 17 (2026), Number 2

CONTENTS

Original scientific papers

Zagreb, 2026

Optimization of the traveling salesman problem through a genetic algorithm guided by self-organizing maps

Fahad Rafique[1,*], **Hengjian Cui**[1], **Abid Hussain**[2], and **Sadaf Amin**[1]

[1] *School of Mathematical Sciences, Capital Normal University Beijing, West Third Ring Road North, Haidian District, 105, 100048, Beijing, China.*
[2] *Department of Statistics, PMAS-Arid Agriculture University Rawalpindi, Shamsabad, Muree Road Rawalpindi, Pakistan.*
E-mail: ⟨fahadgchaudry@gmail.com, hjcui@bnu.edu.cn, abid0100@gmail.com, 4200538002@cnu.edu.cn⟩

Abstract. Optimization is a fundamental process for achieving objectives, exemplified by the Traveling Salesman Problem (TSP), which minimizes resource consumption. This study focuses on optimizing travel routes to reduce fuel costs and maximize tourist visits through efficient time management. We employ a genetic algorithm (GA), a powerful optimization technique, to determine routes. A key challenge in GAs is premature convergence, which can prevent optimal solutions. To address this, we introduce a novel approach integrating Self-Organizing Maps with GAs, specifically through a new selection operator designed to enhance GA performance. An application was developed using real-world coordinates of Pakistani cities. Simulation results demonstrate that our proposed method significantly outperforms existing selection operators in terms of final best distance, average best distance, standard deviation, and computational time. Field validation further confirmed an 18.3% distance reduction and 4.53 million in annual savings in a logistics case study.

Keywords: genetic algorithms; premature convergence; selection operators; self-organizing map; traveling salesman problem.

Received: April 14, 2025; accepted: November 4, 2025; available online: February 9, 2026

DOI: 10.17535/crorr.2026.0014

Original scientific paper.

1. Introduction and Background

In contemporary problem-solving, identifying a solution is insufficient; solutions must be optimized for performance within constraints like limited resources, time, and budget. Optimization techniques are thus developed to find the best possible solution efficiently [20, 21]. As traditional methods often fail with complex real-world problems [1], meta-heuristic algorithms provide approximate solutions for high-dimensional, intractable problems in fields like finance, routing, and scheduling [13].

A key combinatorial challenge is the TSP, which seeks the shortest route visiting each city exactly once and returning to the origin effectively the minimal Hamiltonian cycle in a weighted graph. The TSP is NP-hard, with variations like the Hamiltonian path problem where return is unnecessary [5, 19]. Determining a Hamiltonian path's existence is NP-complete, illustrating the TSP's complexity and relevance to real-world issues like trip planning.

GAs, inspired by natural selection [9], are meta-heuristic techniques valued for global search and robustness [2]. A standard GA involves: (I) initial population creation, (II) fitness evaluation, (III) selection of fitter parents, (IV) crossover to create offspring, and (V) mutation

*Corresponding author.

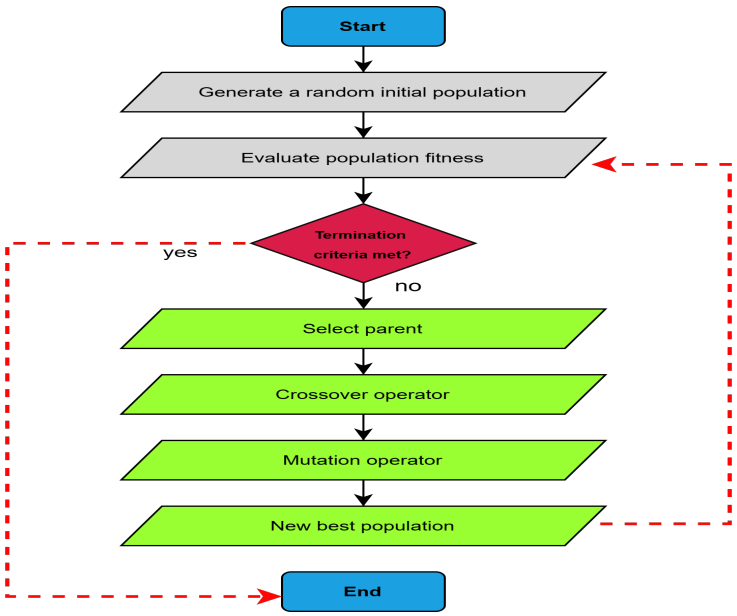


Figure 1: *Typical flowchart of standard GA.*

to prevent premature convergence. This iterative process aims to achieve optimal solutions [18], as shown in Figure 1. However, GAs face premature convergence due to low population diversity, trapping them in local optima [10, 11, 12, 16]. The selection operator is thus critical for maintaining diversity and balancing exploration with exploitation.

This research introduces a novel selection operator to enhance GA convergence by integrating a SOM, forming a hybrid GA-SOM approach. The SOM, an artificial neural network, organizes individuals by normalized fitness ranking, dynamically balancing exploration and exploitation. Validated on TSP and a Pakistan logistics case study where fuel costs are 60% of operational expenses our method reduces journey distance by 18.3% for TCS Express, the largest local provider, servicing 200+ cities with 800 vehicles. Table 1 summarizes related work and our contribution.

2. Selection procedure

The selection process in Genetic Algorithms (GAs) comprises two primary phases. Initially, selection probabilities are assigned to each individual based on their fitness values. This assignment generates a probability vector, given by $P = (p_1, p_2, \cdots, p_N)$, where $p_i \in [0, 1]$ and $\sum_{i=1}^{N} p_i = 1$, with N representing the population size. For a comprehensive understanding of selection probability assignment, readers may refer to [10] and [12]. Subsequently, a sampling procedure is utilized to identify the most suitable parents from the current population for reproduction. This procedure adheres to the Darwinian principle of "survival of the fittest."

Beyond theoretical performance, the efficacy of selection operators is proven in applied Operational Research (OR) contexts like logistics and vehicle routing. Recent work emphasizes hybrid GA approaches to handle real-world constraints. For instance, [14] surveyed hybrid metaheuristics for complex humanitarian vehicle routing, while [7] combined RL with a GA for electric vehicle routing and charging. This trend of hybridizing GAs for specialized logistics challenges motivates our work. We contribute to this domain by integrating SOMs to enhance GA scalability and convergence for large-scale routing problems, as validated in our case study.

This research significantly contributes to the evolution of GA selection methodologies by

Study	Year	Method	Key Features	Limitations
Applegate et al. [3]	2006	Exact (Branch-and-Bound)	Comprehensive TSP study; exact solutions.	High computational cost for large-n.
Whitley et al. [23]	2012	GA selection survey	Review of selection mechanisms.	No novel operator proposed.
Dorigo et al. [6]	2016	ACO	Meta-heuristic TSP solver.	Premature convergence.
Zelinka [24]	2016	GA-SOM	SOM clustering + GA optimization.	Static selection operators.
Papa et al. [17]	2021	Adaptive GA	Dynamic rank-based selection.	Standalone GA only.
Abualigah et al. [1]	2022	Hybrid GA	Survey of hybrid frameworks.	Limited focus on selection operators.
Maroof et al. [14]	2023	Stochastic TSP	Real-world application (developing economies).	Traditional methods.
Ebrahimi et al. [7]	2024	RL + ES (Hybrid)	EV routing	Charging; battery life optimization. & Computational complexity.
Tang et al. [22]	2024	GA + heuristics	Jumping gene operators for TSP.	No hybrid frameworks.
Proposed	2025	**GA-SOM + RRS**	**Novel RRS; Hybrid GA-SOM; OX/CX comparison; TSPLIB + Case-Study validation.**	—

Table 1: *Summary of existing literature and contributions*

proposing a novel selection operator. This new operator is expected to enhance the characteristic traits of the population while optimizing the balance between exploitation and exploration.

2.1. Assignment of probability

This study utilizes a selection of established selection operators, the fundamental characteristics of which are detailed below.

Fitness Proportional Selection (FPS), initially proposed by [9], assigns selection probabilities to individuals in a population in direct proportion to their fitness values. The selection probability, p_i, for the i^{th} individual is calculated as

$$p_i = \frac{1}{N-1}\left(1 - \frac{f_i}{\sum_{j=1}^{N} f_j}\right); \quad i \in \{1, 2, \cdots, N\}, \tag{1}$$

where f_i represents the fitness of the i^{th} individual, and N denotes the population size. While FPS offers conceptual simplicity, it is recognized for its susceptibility to scaling issues, as discussed by [8].

To mitigate the issue of premature convergence often associated with FPS, [4] introduced Linear Rank Selection (LRS). LRS aims to apply a more graduated selection pressure, thereby facilitating the selection of individuals with lower fitness. In LRS, the selection probability, p_i, for the i^{th} ranked individual is determined by

$$p_i = \frac{1}{N}\left(\theta^- + (\theta^+ - \theta^-)\frac{i-1}{N-1}\right); \qquad i \in \{1, 2, ..., N\}, \tag{2}$$

where i denotes the individual's rank, and θ^- and θ^+ represent the selection probabilities assigned to the worst and best-ranked individuals, respectively. These parameters are subject to the constraint $\theta^- + \theta^+ = 2$. However, despite its rank-based approach, LRS may still struggle to effectively capture subtle fitness differences, potentially reducing the discriminatory power of the selection process and compromising the representation of significant fitness variations.

To address the limitations of LRS, [15] proposed an alternative rank-based selection operator: Exponential Rank Selection (ERS). Unlike LRS, ERS assigns selection probabilities that increase exponentially with fitness rank, from the least to the most fit individual. The selection probability for an individual ranked i^{th} is defined as

$$p_i = \frac{\gamma^{N-i}(1-\gamma)}{1-\gamma^N}, \quad i \in \{1, 2, \cdots, N\}, \tag{3}$$

where $0 < \gamma < 1$, and γ represents a constant ratio reflecting the relative weights assigned to individuals based on their fitness ranks. [15] recommends values of γ approaching unity to maximize selection gain.

[12] employed a probabilistic threshold mechanism for selecting tournament winners, termed Probabilistic 2-Tournament Selection (PTS). In this scheme, the winner advances with a probability q, where $0.5 < q < 1$, while the loser is given a subsequent opportunity to compete with a probability of $1 - q$. The selection probability for the i^{th} ranked individual within the PTS method is defined by the following equation

$$p_i = \frac{2(i-1)}{N(N-1)}q + \frac{2(N-i)}{N(N-1)}(1-q), \quad i \in \{1, 2, \cdots, N\}. \tag{4}$$

[16] introduced a fitness-based selection operator. In their procedure, m_i be the size of i-th individual, where $i = 1, 2, \cdots, N$. For the i-th individual, f_i be the corresponding fitness value. Then,

$$M_{f,t} = \text{Median}(f_{(1)}, f_{(2)}, f_{(3)}, \cdots, f_{(k)}),$$

where $f_{(1)}, f_{(2)}, f_{(3)}, \cdots, f_{(k)}$ are the fitness values arranged in ascending order of magnitude, i.e., from worst to best individuals. The expected number of individuals at the next generation will be

$$M_{i,t+1} = m_{i,t} \cdot n \cdot P_{i,t},$$

where the total number of individuals sums to n. The probability of selection for the i-th individual at the t-th generation is

$$P_{i,t} = \frac{f_i + M_{f,t}}{\sum_{j=1}^{k}(m_{j,t}f_j + M_{f,t})}, \quad i = 1, 2, 3, \cdots, N, \tag{5}$$

where f_i is the fitness value of the i-th individual and $M_{f,t}$ is the median of the current population at the t-th generation.

3. Proposed selection operator

3.1. The rationale

The existing literature describes various selection strategies for GAs. Two common methods, LRS and FPS, each present distinct trade-offs. LRS prioritizes population diversity, which can lead to a slower convergence rate due to reduced selection pressure. Conversely, FPS exerts strong selection pressure, potentially compromising diversity and increasing the risk of premature convergence. To address these limitations, this study adopts the relative rank methodology proposed by [10]. This approach preserves the magnitude differences between successive fitness values. We integrate these relative ranks into a linear ranking operator, creating a hybrid methodology that combines the benefits of both FPS and LRS. This new method aims to achieve a more effective balance between exploration and exploitation within the evolutionary process, maintaining adequate selection pressure while preserving population diversity. Furthermore, we incorporate SOMs to enhance solution initialization and improve scalability for future

optimization efforts. This integration contributes to a more robust and efficient GA system. The following sections will detail the fundamental components and operational procedures of our proposed GA system.

3.2. Relative-rank-based selection operator

The Relative-Rank-Based Selection (RRS) operator is designed to assign selection probabilities to individuals in a population while providing explicit control over the trade-off between selection pressure and population diversity. This is achieved by first establishing a nuanced "relative rank" for each individual, which accounts for its position within the ordered fitness landscape, and then mapping these ranks to probabilities using adjustable parameters.
Here's how RRS operates:

- Let $\Omega = (f_1, f_2, \cdots, f_N)$ represent the ascendingly ordered set of fitness values for all individuals in the population.
- The relative rank of the worst individual, $\mathcal{R}_1^{(r)} = 1$, is defined as 1.
- For individuals $i = 2, 3, \cdots, N$, the subsequent expression

$$z_i = i - 1 + \frac{(N-1)(f_{(i)} - f_{(i-1)})}{f_{(N)} - f_{(1)}},$$

  is used to calculate an intermediate value that incorporates both the individual's integer rank and its fitness proximity to neighbors.
- The relative rank $\mathcal{R}_i^{(r)}$ for $i = 2, 3, \cdots, N$ is then defined as the $(i-1)$th smallest value of $\{z_2, z_3, \ldots, z_N\}$. This meticulous assignment ensures a strict ordering of relative ranks:

$$\mathcal{R}_1^{(r)} < \mathcal{R}_2^{(r)} < \cdots < \mathcal{R}_N^{(r)}. \tag{6}$$

- Finally, selection probabilities p_i are assigned using the expression:

$$p_i = \frac{1}{N}\left(\alpha^- + (\alpha^+ - \alpha^-)\frac{\mathcal{R}_i^{(r)} - 1}{N-1}\right); \qquad i \in \{1, 2, ..., N\}. \tag{7}$$

  Here, α^- and α^+ denote the selection probabilities assigned to the worst and best-ranked individuals, respectively. These parameters are constrained by $\alpha^- + \alpha^+ = 2$, ensuring a balanced probability distribution.

3.3. Validating the RRS probability distribution

1. The assigned probabilities must form a complete probability distribution, meaning their sum across all individuals must equal 1.

For any set of probabilities to be valid, their sum must always be 1. In the context of RRS, this means that even though we're assigning probabilities based on relative ranks and using adjustable parameters (α^-, α^+), the scaling within the formula must naturally ensure that all chances of selection add up to 100%. The algorithmic ingenuity resides in the interplay between its linear time complexity and the invariant structural property of the relative rank sum, $(\mathcal{R}_i^{(r)} = \frac{N(N+1)}{2})$ work together to guarantee this sum, regardless of the population size N or the specific values of α^- and α^+. Essentially, the design ensures that increasing the probability of the best-ranked individuals by adjusting α^+ is precisely balanced by a corresponding decrease

in the probabilities of the worse-ranked individuals, all while maintaining the total sum of 1. Its formal derivation is provided in Result 1.

Result 1. The expression (7) is a complete probability distribution, i.e. $\sum_{i=1}^{N} p_i = 1$.

Proof: We begin by summing the probability expression (7) over all individuals i=1,. . . ,N:

$$\sum_{i=1}^{N} p_i = \sum_{i=1}^{N} \frac{1}{N}\left(\alpha^- + (\alpha^+ - \alpha^-)\frac{\mathcal{R}^{(r)}i - 1}{N-1}\right)$$

We can factor out the common term $\frac{1}{N}$ and separate the summation:

$$= \frac{1}{N}\left(\sum i = 1^N \alpha^- + \sum_{i=1}^{N}(\alpha^+ - \alpha^-)\frac{\mathcal{R}^{(r)}i - 1}{N-1}\right)$$

Since α^- is a constant, $\sum i = 1^N \alpha^- = N\alpha^-$. We can also factor out the constant term $\frac{\alpha^+ - \alpha^-}{N-1}$:

$$= \frac{1}{N}\left(N\alpha^- + \frac{\alpha^+ - \alpha^-}{N-1}\sum_{i=1}^{N}(\mathcal{R}_i^{(r)} - 1)\right)$$

$$= \frac{1}{N}\left(N\alpha^- + \frac{\alpha^+ - \alpha^-}{N-1}\left(\sum_{i=1}^{N}\mathcal{R}_i^{(r)} - N\right)\right)$$

At this juncture, we leverage a previously established result by [10], which proves that the sum of the relative ranks $\mathcal{R}_i^{(r)}$ for $i = 1, 2, \cdots, N$ is equivalent to the sum of the first N natural numbers: $\sum_{i=1}^{N} \mathcal{R}_i^{(r)} = \frac{N(N+1)}{2}$. Substituting this identity into our expression:

$$= \frac{1}{N}\left(N\alpha^- + \frac{\alpha^+ - \alpha^-}{N-1}\left(\frac{N(N+1)}{2} - N\right)\right)$$

$$= \frac{1}{N}\left(N\alpha^- + \frac{\alpha^+ - \alpha^-}{N-1}\left(\frac{N(N-1)}{2}\right)\right)$$

We can cancel out the $(N-1)$ term in the numerator and denominator:

$$= \frac{1}{N}\left(N\alpha^- + \frac{\alpha^+ - \alpha^-}{1}\left(\frac{N}{2}\right)\right)$$

$$= \frac{1}{N}\left(N\alpha^- + \frac{N(\alpha^+ - \alpha^-)}{2}\right)$$

Finally, applying the constraint $\alpha^- + \alpha^+ = 2$, the desire result can be obtained.

2. The selection must adhere to Darwin's "survival-of-the-fittest" principle, where fitter individuals are assigned a higher or equal probability of selection than less fit individuals ($p_j \geq p_i$ for $j > i$).

It should align with Darwin's survival-of-the-fittest principle, specifically represented as $p_{i+1} > p_i$ for $j > i$ where both i and j are integers falling within the range of $1, 2, \cdots, N$. Its formal derivation is provided in Result 2.

Result 2. According to Darwin's theory, the expression (7) should follows the survival-of-fittest-criterion, i.e. $p_{i+1} > p_i$.

Proof: As $p_{i+1} > p_i$

$$\frac{1}{N}\left(\alpha^- + (\alpha^+ - \alpha^-)\frac{\mathcal{R}_{i+1}^{(r)} - 1}{N-1}\right) > \frac{1}{N}\left(\alpha^- + (\alpha^+ - \alpha^-)\frac{\mathcal{R}_i^{(r)} - 1}{N-1}\right)$$

$$\left((\alpha^+ - \alpha^-)\frac{\mathcal{R}_{i+1}^{(r)} - 1}{N-1}\right) > \left((\alpha^+ - \alpha^-)\frac{\mathcal{R}_{i}^{(r)} - 1}{N-1}\right)$$

$$\left((\alpha^+ - \alpha^-)\frac{\mathcal{R}_{i+1}^{(r)}}{N-1}\right) > \left((\alpha^+ - \alpha^-)\frac{\mathcal{R}_{i}^{(r)}}{N-1}\right)$$

$$\mathcal{R}_{i+1}^{(r)} > \mathcal{R}_{i}^{(r)}.$$

This inequality is always true, it can be seen in equation (7). Thus, individuals with better fitness are assigned higher selection probabilities.

4. Adopted methodology

This study employs two main computational approaches to solve the TSP.

A genetic algorithm (GA) maintains a population of candidate routes, each encoded as an ordered city list. The population evolves through selection, crossover, and mutation, with fitness measured by total route distance. Selection favors shorter routes to serve as parents, crossover merges parent segments to produce valid offspring, and mutation introduces random changes to maintain diversity and avoid premature convergence. This process iterates until a termination condition is met, aiming to approach the optimal minimal-distance route (see Algorithm 1 in the online appendix).

The second method integrates the GA with a Self-Organizing Map (SOM), an artificial neural network inspired by the animal visual cortex. The SOM uses competitive learning to reduce dimensionality while preserving input topology. Nodes' weight vectors $\mathbf{W}_v$ are updated iteratively based on the best matching unit u for input $D(t)$ by

$$\mathbf{W}_v(s+1) = \mathbf{W}_v(s) + \theta(u, v, s) \cdot \alpha(s) \cdot (D(t) - \mathbf{W}_v(s)),$$

where s is the current iteration, θ is the neighborhood function, and α is the learning rate. Traditional static SOM structures may underperform on complex data, so the GA optimizes the SOM architecture, resulting in the hybrid GA-SOM method that adapts to dataset complexity. This approach improves mapping quality, verified by error metrics across multiple datasets (see Algorithm 2 in the online appendix).

5. Experimental setup

All experiments were conducted on a computing system running Windows 11. The system was equipped with an AMD Ryzen 7 4700U processor featuring Radeon Graphics and 32 GB of RAM. The application was developed in R (version 4.4.2), leveraging the following libraries: `parallel`, `ggplot2`, `Rfast`, `TSP`, and `kohonen`. Effectively addressing a problem necessitates not only a meticulous selection of appropriate algorithms but also a judicious choice of parameters. The significant impact of parameter selection on the results is demonstrated in (see Table A2 in the online appendix).

6. Results and discussion

A summary of key results for the most effective method combinations is presented in Table 2. Comprehensive results for all operators, crossover methods, and problem sizes are provided in the Online Appendix (Tables 4-8).

This section evaluates the performance of the proposed Relative-Rank-Based Selection (RRS) operator, both within a standalone Genetic Algorithm (GA) and a hybrid GA-Self

Organizing Map (GA-SOM) framework. Performance is assessed using benchmark TSPLIB instances and a real-world case study with Pakistani city coordinates, measured by solution quality (FinalBest, AvgBest), stability (Standard Deviation, SD), and computational efficiency (Average Run Time, ART).

6.1. Overall performance and comparative analysis

The proposed RRS operator consistently outperformed existing selection operators (FPS, LRS, ERS, PTS) across all problem sizes and configurations. A summary of key results for the most effective method combinations is presented in Table 2.

Problem Context	Method Combination	FinalBest (km)	AvgBest (km)	SD (km)	Time (s)
TSPLIB ($n = 500$)	RRS (GA-SOM + CX)	1910.56	1950.12	103.44	4.21
TSPLIB ($n = 300$)	RRS (GA + OX)	2042.73	2089.45	115.80	1.85
Pakistan ($n = 200$)	RRS (GA-SOM + CX)	612.72	628.15	13.08	0.46
Pakistan ($n = 200$)	RRS (GA + CX)	629.06	645.91	15.50	0.17

Table 2: *Summary of Key Results for Top-Performing RRS Combinations*

Standalone GA Performance: When integrated into the standard GA, RRS demonstrated superior solution quality and faster convergence than other operators. For example, on the Pakistan cities dataset, RRS with Cycle Crossover (CX) found a 629 km route, significantly shorter than routes found by FPS (720 km) or LRS (∼690 km). Furthermore, RRS reduced the average runtime by approximately 30% compared to other operators, highlighting its computational efficiency.

Hybrid GA-SOM Performance: The integration of RRS within the GA-SOM framework yielded the most robust and effective results, particularly for larger problems. The GA-SOM's ability to pre-cluster cities reduced the search space complexity by an estimated 60%, leading to more stable convergence and higher-quality solutions. As shown in Table 2, **RRS (GA-SOM + CX)** achieved the best overall performance, delivering the lowest FinalBest and AvgBest distances with the highest stability (lowest SD). This combination effectively balanced exploration and exploitation, preventing premature convergence.

6.2. Real-World logistics case study and economic impact

The practical value of the proposed method was validated through a composite case study based on the operations of Trazum Courier Service (TCS) in Pakistan.

This optimization translates to substantial economic and environmental savings:

- **Annual Fuel Savings:** Calculated as $\frac{750-612.72}{8}$ liters × \$1.10/liter × 800 vehicles × 300 days = **\$4.53 Million**.
- **CO_2 Reduction:** Estimated at **28.5 tons** per day.

Note: Calculations of Saved fuel = $\frac{750-612.72}{8}$ liters; Emissions = 2.68kg CO_2/liter.

The logistics application shows even greater practical impact. For 200 cities, RRS (GA-SOM + CX) achieves 18.3% shorter routes than current industry operations shown in Table 3, translating to \$4.53M annual savings for a typical medium-sized operator.

Method	Distance (km)	Fuel Cost (USD/day)	Savings
TCS Current	750	103.13	-
FPS (GA)	720.10	99.01	4.0%
RRS (GA-SOM + CX)	**612.72**	**84.25**	**18.3%**

Table 3: *Economic Impact of optimized routes (200 cities)*

6.3. Discussion of optimal combinations and visual analysis

The comparative performance of different RRS configurations reveals distinct advantages for specific scenarios, supported by comprehensive visual evidence in Table 4 and Figures 3-4.

RRS (GA-SOM + CX) is the recommended combination for large-scale, critical routing problems where solution quality and reliability are paramount. Its low SD indicates consistent performance, while Figure 3 visually demonstrates the more efficient route pathing achieved by this hybrid approach.

RRS (GA + CX) is the optimal choice when computational speed is the highest priority, especially for smaller problems, offering a good balance of speed and solution quality. The computational efficiency of RRS-based approaches is clearly illustrated in Figure 2, which ranks selection methods by average run times.

RRS (GA + OX) performed well for medium-sized problems, demonstrating the versatility of the RRS operator. The comprehensive comparison in Table 4 further validates the consistent superiority of RRS operators across different crossover methods and problem sizes.

These visual representations collectively highlight the computational efficiency of the RRS operator compared to existing methods and demonstrate the practical applicability of the optimized routes generated by both standalone RRS and hybrid GA-SOM approaches.

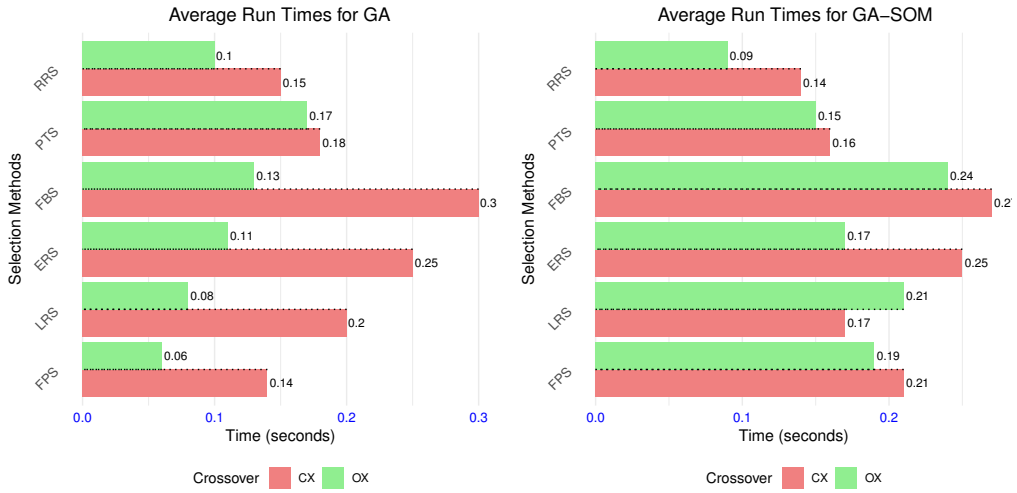


Figure 2: *Ranking of Selection methods by average run times.*

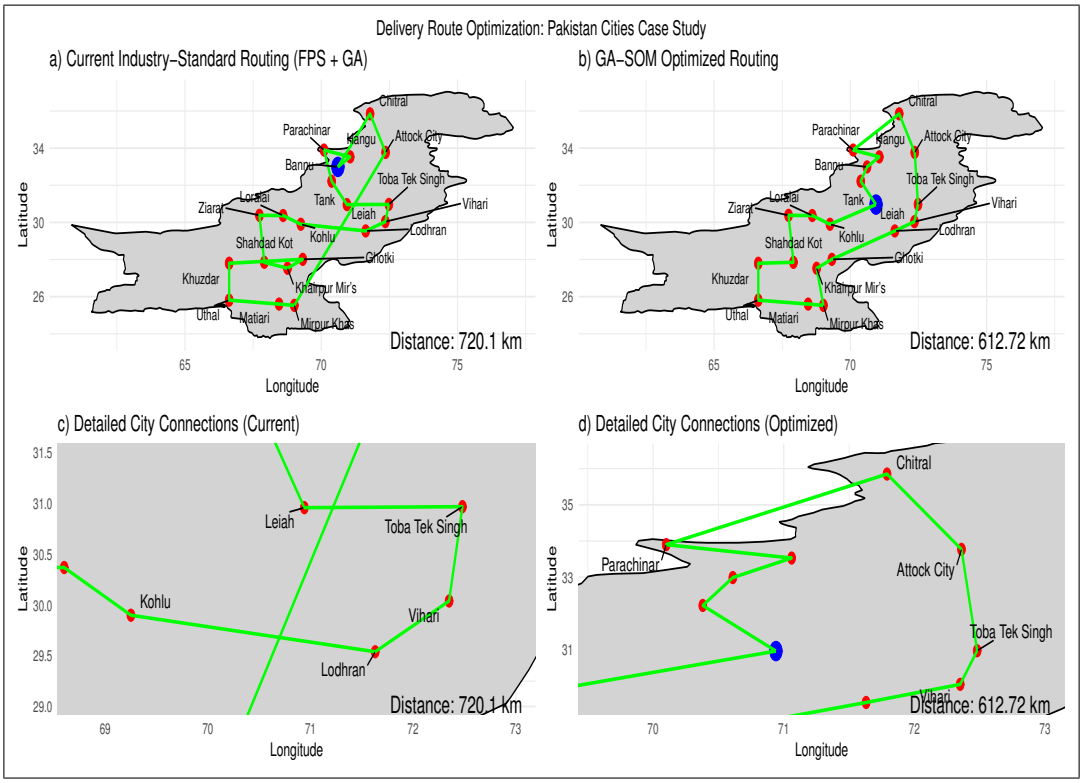


Figure 3: *Delivery route optimization: Current industry-standard routing (750km), GA-SOM optimized routes (613km), and detailed city connections.*

Method	n	Operator	GA				GA-SOM			
			FinalBest	AvgBest	SD	AvgRunTime	FinalBest	AvgBest	SD	AvgRunTime
CX	20	FPS	708.32	1099.62	25.17	0.14	650.04	1035.76	20.17	0.06
		LRS	696.63	1003.81	33.17	0.20	699.19	1013.88	41.59	0.08
		ERS	700.12	1010.71	34.42	0.25	700.72	1020.41	42.10	0.11
		FBS	705.11	1015.23	35.21	0.30	710.07	1030.18	43.19	0.13
		PTS	660.21	680.10	24.37	0.18	615.32	705.12	22.00	0.17
		RRS	619.87	726.33	27.26	0.15	609.18	711.70	24.91	0.10
OX	20	FPS	691.02	870.52	24.12	0.21	690.08	817.13	18.17	0.19
		LRS	692.63	830.93	29.81	0.17	674.00	814.02	24.67	0.21
		ERS	685.12	790.35	30.21	0.25	702.32	877.41	25.14	0.17
		FBS	719.41	895.13	31.05	0.27	705.00	870.52	26.71	0.24
		PTS	640.32	720.06	23.30	0.16	670.62	690.41	20.61	0.15
		RRS	629.06	718.73	24.86	0.14	612.89	706.03	26.92	0.09

Table 4: *Performance comparison of existing selection operators and RRS method using CX and OX crossover and EM mutation operators for 20 Pakistan' cities.*

7. Summary

This study meticulously compares the performance of a novel RRS operator against current selection methods, employing diverse crossover techniques and problem sizes. The RRS operator consistently surpassed existing operators across all evaluated key metrics, including solution quality, stability, and computational efficiency. When integrated with the GA-SOM framework RRS+GA-SOM, its performance exceeded that of RRS+GA, yielding more effective and prac-

tical solutions across all tested scenarios. Among the assessed combinations, RRS (GA-SOM + CX) demonstrated superior efficacy. This combination consistently produced the lowest SD values, achieved optimal FinalBest and AvgBest results, and maintained commendably competitive average run times. This particular configuration proves especially advantageous for extensive optimization challenges where stability and solution quality are paramount. Furthermore, our RRS (GA + OX) exhibited exceptional performance in specific medium-sized problems, thereby reinforcing the adaptability and robustness of the RRS operator across various optimization scenarios.

Experimental results obtained from the Pakistan city coordinates dataset further underscore the advantages of the RRS operator in terms of solution quality, stability, and computational economy. The hybrid GA-SOM framework significantly amplifies these benefits, particularly for extensive optimization techniques. Among all combinations, RRS + CX (GA-SOM) proved to be the most efficacious, consistently yielding the lowest FinalBest and AvgBest values, demonstrating the most steady convergence (minimal SD), and achieving notably competitive run times. These findings collectively illustrate the practical utility of the RRS operator and the hybrid GA-SOM framework for real-world logistics and routing challenges. Our composite logistics case study indicates annual savings of $4.53 million for a medium-sized operator, thereby substantiating its direct utility in developing economies where fuel expenses are a primary driver of operational costs.

This study is limited to symmetric TSP; future work should assess performance on asymmetric and constrained vehicle routing problems. While RRS shows superior performance, its parameters require tuning, suggesting a need for adaptive control mechanisms. Scalability to ultra-large networks (>10,000 nodes) also warrants investigation through parallelization.

Policymakers should incentive logistics firms to adopt AI driven optimization through tax benefits or other grants. Establishing a national digital logistics platform could democratize access to these tools. Furthermore, integrating fuel savings and emission reductions into transportation policies would align economic and environmental objectives.

Data accessibility and Online Appendix

In the Online Appendix, containing detailed results for all algorithmic pseudocode, parameter setup and all large comprehensive tables is available at `https://doi.org/10.5281/zenodo.17509716`.

The primary dataset used to validate the conclusions presented in this manuscript was obtained from the well-established TSPLIB repository, a publicly available resource widely recognized for its collection of traveling salesman problem instances. This repository can be accessed at `https://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95`.

The findings presented in this manuscript are supported by data obtained from the publicly available Kaggle platform through following link: `https://www.kaggle.com/datasets/tayyarhussain/pakistani-cities-latitude-andlongitude-dataset?resource=download`.

State Bank of Pakistan financial Stability Review retrieved from `https://www.sbp.org.pk/fsr`.

Pakistan logistics performance assessment retrieved from `https://www.worldbank.org`.

Tranzum Courier Service (TCS) company overview can accessed at `https://www.tcsexpress.com`.

Pakistan Automotive Manufacturers Association commercial vehicle fuel efficiency report assessed at `https://www.pama.org.pk/publications/Table7:Avg.7.5-8.5km/literformediumtrucks`.

Oil & Gas regulatory authority fuel Price can accessed at `https://www.ogra.org.pk`.

References

[1] Abualigah, L., Elaziz, M. A., Khasawneh, A. M., Alshinwan, M., Ibrahim, R. A., Al-qaness, M. A. A., Mirjalili, S., Sumari, P., and Gandomi, A. H. (2022). Meta-heuristic optimization algorithms for solving real-world mechanical engineering design problems: a comprehensive survey, applications, comparative analysis, and results. Neural Computing and Applications, 34, 4081–4110. doi: 10.1007/s00521-021-06747-4

[2] Alanzi, E., Menai, and M. E. B. (2025). Solving the traveling salesman problem with machine learning: a review of recent advances and challenges. Artificial Intelligence Review, 58(9), 267. doi:10.1007/s10462-025-11267-x

[3] Applegate, D. L., Bixby, R. E., Chvátal, V., and Cook, W. J. (2006). *The Traveling Salesman Problem: A Computational Study*. Princeton University Press.

[4] Baker, J. E. (1985). Adaptive selection methods for genetic algorithms. In: *Proceedings of an International Conference on Genetic Algorithms and their Applications*, 101-111. Hillsdale, New Jersey.

[5] Bokhari, S. W. A., Ali, N., and Hussain, A. (2025). Optimizing maritime routes: A multi-method analysis from Shanghai to Vladivostok. Croatian Operational Research Review, 16(1), 73–83. doi:10.17535/crorr.2025.0007

[6] Dorigo, M., Birattari, M., and Stutzle, T. (2016). *Ant Colony Optimization*. IEEE Computational Intelligence Magazine, 1(4), 28 - 39. doi:10.1109/MCI.2006.329691

[7] Ebrahimi, D., Kashefi, S., de Oliveira, T. E. A., and Alzhouri, F. (2024). Efficient routing and charging strategy for electric vehicles considering battery life: A reinforcement learning approach. In: *2024 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 429-435. doi:10.1109/CCECE59415.2024.10667251

[8] Grefenstette, J.J. (1986). Optimization of control parameters for genetic algorithms. IEEE Transactions on Systems, Man and Cybernetics, 16(1), 122-128. doi:10.1109/TSMC.1986.289288

[9] Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press.

[10] Hussain, A., and Muhammad, Y. S. (2020). Trade-off between exploration and exploitation with genetic algorithm using a novel selection operator. Complex & Intelligent Systems, 6(1), 1-14. doi:10.1007/s40747-019-0102-7

[11] Hussain, A. and Cheema, S. A. (2020). A new selection operator for genetic algorithms that balances between premature convergence and population diversity. Croatian Operational Research Review, 11(2020), 107–119. doi:10.17535/crorr.2020.0009

[12] Julstrom, B. A. (1999). It's all the same to me: Revisiting rank-based probabilities and tournaments. In *Proceedings of the 1999 Congress on Evolutionary Computation-CEC99*, 2, 1501-1505. IEEE. doi:10.1109/CEC.1999.782661

[13] Khurshid, B., Maqsood, S., Khurshid, Y., Naeem, K., and Khalid, Q. S. (2024). A hybridization of evolution strategies with iterated greedy algorithm for no-wait flow shop scheduling problems. Scientific Reports, 14(1), 2376. doi:10.1038/s41598-023-47729-x

[14] Maroof, A., Khalid, Q. S., Mahmood, M., Naeem, K., and Saeed, S. (2023). Vehicle Routing Optimization for Humanitarian Supply Chain: A Systematic Review of Approaches and Solutions. IEEE Access, 11, 127157–127175. doi:10.1109/ACCESS.2023.3331062

[15] Michalewicz, Z., Janikow, C. Z., and Krawczyk, J. B. (1992). A modified genetic algorithm for optimal control problems. Computers & Mathematics with Applications, 23(12), 83–94. doi:10.1016/0898-1221(92)90094-X

[16] Naqvi, F. B. and Shad, M. Y. (2022). Seeking a balance between population diversity and premature convergence for real-coded genetic algorithms with crossover operator. Evolutionary Intelligence, 1-16. doi:10.1007/s12065-021-00636-4

[17] Papa, G. (2021). Applications of dynamic parameter control in evolutionary computation. In: *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 1064–1088. doi:10.1145/3449726.3461435

[18] Rafique, F., Cui, H., and Amin, S. (2025). An adaptive multi-strategy selection method for improving genetic algorithm performance. Advances and Applications in Statistics, 92(8), 1105–1142. doi:10.17654/0972361725050

[19] Ruan, Y., Cai, W., and Wang, J. (2024). Combining reinforcement learning algorithm and genetic

algorithm to solve the traveling salesman problem. The Journal of Engineering, 2024(6), e12393. doi:10.1049/tje2.12393

[20] Seyyedabbasi, A., and Kiani, F. (2023). Sand Cat swarm optimization: A nature-inspired algorithm to solve global optimization problems. Engineering with Computers, 39(4), 2627–2651. doi:10.1007/s00366-022-01604-x

[21] Shabani, A., Asgarian, B., Salido, M., and Gharebaghi, S. A. (2020). Search and rescue optimization algorithm: A new optimization method for solving constrained engineering optimization problems. Expert Systems with Applications, 161, 113698. doi:10.1016/j.eswa.2020.113698

[22] Tang, L., Wang, X., He, Z., Wang, S., and Lin, J. (2024). Review of Traveling Salesman Problem Solution Methods. In Pan, L., Wang, Y., Lin, J. (Eds.), Bio-Inspired Computing: Theories and Applications. BIC-TA 2023. Communications in Computer and Information Science (Vol. 2062, 3-16). Springer, Singapore. doi:10.1007/978-981-97-2275-4_1

[23] Whitley, D. and Sutton, A. M. (2012). Genetic algorithms A survey of models and methods. In G. Rozenberg, T. Back, J. N. Kok (Eds.), Handbook of Natural Computing (Vol. 4, 637–671). Springer, Berlin, Heidelberg. doi:10.1007/978-3-540-92910-9_21

[24] Zelinka, I. (2016). SOMA—self-organizing migrating algorithm. Self-Organizing Migrating Algorithm: Methodology and Implementation, 3-49. Springer International Publishing. doi:10.1109/ITIS50118.2020.9321044

Croatian Operational Research Review
CRORR **17:2**(2026), 195–207

Data inversion in MCDM problems: nonlinear 1/a and linear ReS inversion

Irik Z. Mukhametzyanov[1,*] **and Ildar U. Zulkarnay**[1]

[1] *Institute of Socio-Economic Research – subdivision of the Ufa Federal Research Centre of the Russian Academy of Sciences, October Avenue, 71, 450054 Ufa, Russia*
E-mail: ⟨izmukhametzyanov@gmail.com, zulkar@mail.ru ⟩

Abstract. A comprehensive analysis of the procedures for consistent normalization/inversion of benefit and cost attributes for multi-criteria decision making (MCDM) and multivariate classification problems is performed. This study demonstrates that the commonly used $1/a$ transformation for cost attribute inversion introduces structural inconsistencies in normalized data. Nonlinear data inversion does not have a reasonable interpretation of values and leads to a violation of mutual distances in the original data. The measurement scales of various attributes are not consistent and there is a shift in the domains of normalized values. Elimination of these problems is achieved by using the linear inversion Reverse Sorting algorithm (ReS). The ReS algorithm offers more consistent, linear, and interpretable results for handling cost attributes in MCDM. The ReS algorithm is a linear transformation and preserves the original information about the object: dispositions of attribute values, preserves the relative positions of the domains of different attributes and can be applied to both the original and normalized data sets. The ReS algorithm eliminates all the shortcomings of nonlinear inversion and is recommended for inversion of values when coordinating the optimization goals of a multi-criteria problem, as well as in the weighing methods based on information contained in the decision matrix.

Keywords: MCDM, normalization of multivariate data, non-linear inversion, Reverse Sorting algorithm (ReS), scale coordination

Received: May 19, 2025; accepted: October 16, 2025; available online: February 9, 2026

DOI: 10.17535/crorr.2026.0015

Original scientific paper.

1. Introduction

The standard problem of multi-criteria choice on a discrete set of alternatives (MCDM) is formed as follows [4, 12]: for an available set of alternatives $A_i, (i = 1, \ldots, m)$ the attributes a_{ij} (decision matrix) are determined in the context of the selected evaluation criteria $C_j, (j = 1, \ldots, n)$. Then a decision model is formed. The most widespread models are based on Value Measurement methods, which include the method of assessing the significance of criteria (W), the method of normalizing the decision matrix ($Norm$) and the method of aggregation (F) of various attributes of each alternative into an indicator of its effectiveness Q_i:

$$Q_i = F(x_{ij}, w_j), \quad \forall i = 1, \ldots, m. \tag{1}$$

Where, F is one of the methods or functions (including multi-step) of data aggregation. F transforms the normalized values of the attributes $x_{ij} = Norm(a_{ij})$, taking into account the significance of the criteria w_j, to the numerical value Q_i — the rating of the alternative (or the

*Corresponding author.

performance indicator, or the assessment score). For example, one of the simplest and most widely used aggregation methods is a weighted sum of the normalized values of the attributes of each alternative ($Q_i = \sum_{j=1}^{n} w_j x_{ij}$).

The choice of the trio of methods (F, W, $Norm$) is not formalized and is carried out on the basis of some general principles [8].

Additive (or multiplicative) aggregation functions F assume consistency in the direction of improvement of each attribute. There are three possible types of goals for different criteria: larger the best (LTB), smaller the best (STB), and target the best (TTB). Aggregation involves matching the goals of the criteria if the goal type of different criteria is different. For Value Measurement Methods, matching the goals is achieved by data inversion. We define data inversion as a rearrangement of the sorted list in reverse order. With a common LTB goal, the cost criteria are inverted, and with an STB goal, the benefit criteria are inverted. For criteria with a specified target value, with a common LTB goal, the inversion is performed for all values greater than the target, and with an STB goal, the inversion is performed for all values less than the target [8]. Inversion methods are also multivariate. Following the review by Jahan (2015) [5], and other studies, the understanding of this fact is exaggerated and the term inversion is not used, but separate coordinated normalization of benefit criteria (LTB type) and cost criteria (STB type) is applied. In this case, two separate formulas are used. For example, for Max-method of normalization, the formula is:

$$x_{ij} = \frac{a_{ij}}{a_j^{max}}, \forall j, \quad \text{then} \quad x_{ij} := 1 - x_{ij}, \text{ for } \forall j \in C_j^- \text{— cost criteria.} \tag{2}$$

That is first normalization of all attributes is performed, and then a linear inversion of the form (1–x) is performed for cost attributes (or benefits for the general goal of STB).

In formula (2) and further, the assignment operator ":=" is used, which is interpreted as follows: the variable on the left is assigned the result of the operation performed on the right. This will allow one variable to be operated upon when performing the normalization/inversion operation.

Normalization of multidimensional data is not only bringing data to a dimensionless form, but also coordinating normalization scales. This is a necessary element of transforming multidimensional data. The requirement is that none of the data sets receive priority over the others at the normalization stage. In fact, this is impossible [8], since each of the features uses its own normalization scale. Considering that inversion is an integral part of multidimensional data normalization, the same requirements are imposed on the inversion procedure as on normalization.

In this paper, we will focus on the analysis of the procedure for the consistent normalization of the benefit and cost criteria $x_{ij} = Norm(a_{ij})$, or more precisely, on the inversion of the objective depending on the overall objective of multi-criteria optimization. Normalization/inversion is relevant for MCDM problems for Value Measurement models and for Goal or Reference Level models. In the second case, inversion is not required. Therefore, this article discusses various data inversion methods as applied to solving MCDM problems using Value Measurement Methods. This inversion is often used as a component of such models as, for example, Weighted Sum Model (WSM), Weighted Product Model (WPM), Weighted Aggregated Sum Product Assessment (WASPAS) [15], COmplex PRoportional ASsessment (COPRAS) [14], Additive Ratio Assessment (ARAS) [16], COmbinative Distance-based ASsessment (CODAS) [3], Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) [12], etc. Exactly the same approach is used in objective methods for assessing the weight of criteria based on the information contained in the decision matrix: Entropy weight method (EWM) [13], MEthod based on the Removal Effects of Criteria (MEREC) [6], Simultaneous Evaluation of Criteria and Alternatives (SECA) [2], method of Criterion Impact LOSs (CILOS) [1], etc.

It is clear that the monotonic a strictly decreasing function $1/a$ will preserve the order of

the values in the list and invert the data, so that larger values become smaller and vice versa. In this study, based on the analysis and simple examples, we will show that such an inversion is not correct. A nonlinear transformation changes the data, or rather changes the distances between the values. This means that the normalized data is fundamentally distorted. Thousands of studies in MCDM using nonlinear inversion of cost attributes based on a transformation of the type $1/a$ should be recognized as inaccurate. As an alternative, the paper proposes a linear method, Reverse Sorting (ReS), an algorithm that eliminates all the problems of nonlinear transformation. The ReS algorithm is a linear transformation and preserves the original information about the object: dispositions of attribute values, preserves the relative positions of the domains of different attributes and can be applied to both the original and normalized data sets.

Although the basic publication of the linear inversion of ReS in the journal IJITDM [7] was in 2020, very few studies use this method. Meanwhile, this is the simplest and most universal inversion for both univariate and multivariate data. In this paper, the authors aim to show that the only correct solution is to use the ReS inversion, which satisfies all the necessary requirements for normalizing multidimensional data.

2. Methods: data transformation in multi-criteria optimization problems

In problems with a multidimensional data set, as a rule, preliminary transformation of the values of all features is required so that they fall into intervals of comparable size. This procedure of bringing values measured in different scales to a conditionally common scale is defined as normalization. Multidimensional normalization transforms data independently for each attribute.

2.1. Linear normalization

The general formula for linear normalization is:

$$x_{ij} = \frac{a_{ij} - a_j^*}{k_j}, \quad i = 1, ..., m, j = 1, ...n. \tag{3}$$

The shift coefficients a_j^* and stretching-compression coefficients k_j are different for each criterion. This is due to the fact that the attributes of objects and the ranges of their values can differ greatly from each other. Therefore, each of the attributes uses its own normalization scale and the normalized values depend on the measurement scale and on the range natural values of the attributes ($r_j = range_i(a_{ij}) = a_j^{max} - a_j^{min}$).

2.2. Invariance of attribute value dispositions under linear transformations

We define the disposition of the values a_{ij} of the jth attribute for alternatives A_p and A_q as:

$$\Delta a_{pq}^{(j)} = \frac{|a_{pj} - a_{qj}|}{a_j^{max} - a_j^{min}}. \tag{4}$$

Without taking into account the sign, the disposition is the relative distance between two attribute values of alternatives A_p and A_q. It is easy to show [8, 9] that the disposition of the values of the jth attribute for the alternatives A_p and A_q is an invariant with respect to linear transformations $x = L(a)$:

$$\Delta a_{pq}^{(j)} = \frac{|a_{pj} - a_{qj}|}{a_j^{max} - a_j^{min}} = \frac{|x_{pj} - x_{qj}|}{x_j^{max} - x_j^{min}} = \Delta x_{pq}^{(j)}, \quad \forall p, q = 1, ..., m, \quad \forall j. \tag{5}$$

The converse is also true: if the dispositions of two data sets are identical, then they are linearly transformed into each other.

The interpretation of this property is that relative distances between data are preserved under linear transformations. For a sorted list, property (5) has the following form:

$$\Delta a_i^{(j)} = \frac{a_{i+1,j} - a_{ij}}{a_{mj} - a_{1j}} = \frac{x_{i+1,j} - x_{ij}}{x_{mj} - x_{1j}} = \Delta x_i^{(j)}, \;\; i = 1, ..., m-1, \;\; \forall j. \tag{6}$$

Obviously, two such lists are ordered identically and differ in scale and position.

2.3. Invariance of dispositions of rating under linear transformations of data

We apply a linear transformation of the form:

$$u_{ij} = x_{ij}/k + h_j, \tag{7}$$

to the normalized values. Here, the compression coefficient k is a fixed constant for $\forall j$, and the bias h_j can be different for different attributes. That is, all attributes will be subjected to the same compression or expansion, but the shift of the domains of different attributes may be different. In the study [9], the invariance of rating dispositions was proven for the case of linear (WSM) or homogeneous (TOPSIS) aggregation functions F, according to formulas (1) under otherwise identical conditions.

Let us explain this important property.

Let us apply the MCDM rank model and determine the ratings of alternatives $Q_i^{(1)} = F(x_{ij}, w_j)$ in accordance with (1) for the normalized decision matrix x_{ij}, and determine the ratings of alternatives $Q_i^{(2)} = F(u_{ij}, w_j)$ for the transformed decision matrix u_{ij}. As the aggregation function F, we use a linear function (WSM and similar) or a homogeneous function (TOPSIS and similar). In both cases, we use the same aggregation functions and the same weighting coefficients. Invariance of the rating dispositions means that

$$\Delta Q_i^{(1)} = \Delta Q_i^{(2)}, \;\; \forall i = 1, \ldots, m-1. \tag{8}$$

From the equality of dispositions it follows that the ranks R_i of the alternatives are the same in both cases

$$R_i^{(1)} = R_i^{(2)}, \;\; \forall i = 1, \ldots, m. \tag{9}$$

Thus, a linear transformation of normalized data of the form $u_{ij} = x_{ij}/k + h_j$ does not change the dispositions of the rating list and does not change the ranks of the alternatives if a linear or homogeneous aggregation function is used.

2.4. Basic principles of multivariate data normalization

The basic principles of multivariate data normalization are set out in the recently published monograph [8]. Preservation of the information content of the data after transformation includes:

(*i*) Preservation of the information content of the data in the context of each attribute: For all linear normalization methods and linear transformations, the dispositions of natural values for each attribute are preserved according to property (5). The use of nonlinear transformations violates the structure of the original data and requires meaningful justification.

(*ii*) Alignment of normalized value scales: Each attribute has its own normalization scale, which is not aligned with the normalization scale of other attributes. This is due to the different range of values and different distribution of features in the domain. For multi-criteria problems,

linear methods produce anisotropic scaling, when at least one of the scaling factors differs from the others. There is a deformation of the multidimensional feature cloud. This can lead to the priority of the contribution of individual attributes to the efficiency indicator of alternatives. The degree of difference depends on the decision matrix D (on the problem) and the choice of the data normalization method. Since it is impossible to reach a consensus on the alignment of normalized value scales at the normalization stage, it is necessary to minimize possible consequences when choosing normalization/inversion methods.

3. Data inversion procedures

3.1. Linear inversion of normalized values of the form (1-x)

One popular approach combines normalization and inversion. First, normalize all attributes, and then perform a linear inversion of the form $(1 - x_{ij})$ for the cost attributes. For example, for the Max method, the normalization/inversion is:

$$x_{ij} = \frac{a_{ij}}{a_j^{max}}, \ \ \forall i, j, \tag{10}$$

$$x_{ij} := 1 - x_{ij}, \text{ for } j \in C_j^- \text{ — cost criteria.} \tag{11}$$

In this case, if the normalized values x_{ij} are shifted to 1, as for the Max method, then $(1 - x_{ij})$ will be shifted to 0 and the contribution of cost criteria to the integral indicator will obviously be lower than the contribution of benefit criteria. However, since (11) is a linear transformation of the form (7) (k=-1, $h_j = 1$), then in accordance with property (8)-(9), this will not affect the ranking and does not change the dispositions of alternatives in the case of using the WSM or TOPSIS aggregation methods, including analogues.

Note that the above is also true for the inversion $-x$:

$$x_{ij} := -x_{ij}, \text{ for } j \in C_j^- \text{ — cost criteria.} \tag{12}$$

In this case, if the normalization transforms the values into [0, 1], then when using the inversion (12) $x_{ij} \in [-1, 0]$.

Inversions (11), (12) are obviously problematic when using nonlinear methods of aggregating normalized values of different attributes.

3.2. Nonlinear inversion of the form $1/a$

Linear normalization procedures for benefit (C_j^+) and cost (C_j^-) criteria for the overall LTB goal are as follows:

$$x_{ij} = Norm(a_{ij}) = \frac{a_{ij} - a_j^*}{k_j}, \ \text{ for } j \in C_j^+, \quad x_{ij} = \frac{1/a_{ij}}{t_j}, \quad \text{for } j \in C_j^-. \tag{13}$$

1) Max/ iMax (inversion Max with non-linear inversion iMax)

$$x_{ij} = Max(a_{ij}) = \frac{a_{ij}}{k_j}, \quad a_j^* = 0, \ k_j = a_j^{max}, \text{ for } j \in C_j^+, \quad t_j = \frac{1}{a_j^{min}}, \text{ for } j \in C_j^-. \tag{14}$$

2) Sum/ iSum (inversion Sum with non-linear inversion iSum)

$$x_{ij} = Sum(a_{ij}) = \frac{a_{ij}}{k_j}, \ \ a_j^* = 0, \ k_j = \sum_{i=1}^{m} a_{ij}, \text{ for } j \in C_j^+, \quad t_j = \sum_{i=1}^{m} \frac{1}{a_{ij}}, \text{ for } j \in C_j^-. \tag{15}$$

3) Vec/ iVec (inversion Vec with non-linear inversion iVec)

$$x_{ij} = Vec(a_{ij}) = \frac{a_{ij}}{k_j}, \quad a_j^* = 0, k_j = \sqrt{\frac{1}{m}\sum_{i=1}^{m} a_{ij}^2}, \text{ for } j \in C_j^+, \quad t_j = \sqrt{\frac{1}{m}\sum_{i=1}^{m} \frac{1}{a_{ij}}^2}, \text{ for } j \in C_j^-. \quad (16)$$

Note that the normalization coefficients t_j are constants for fixed j (and similarly k_j), and therefore all three inverse normalizations (13)-(16) are inversions of the same type using the $1/a$ transformation. Note also that when choosing the overall goal of the STB problem, it is necessary to swap the formulas for the benefit and cost criteria. In MCDM problems, as a rule, the general goal is LTB and transformations (13)-(16) are used.

Disadvantages of the $1/a$ transformation:

1) with the $1/a$ transformation, interpretation of the normalized values is difficult. To interpret the normalized values, the connecting term "share of" is used, for example, the share of the largest value. In the case of normalization (13)-(16), the value $1/a_{ij}$ in the numerator of these formulas does not make sense,

2) the first requirement for normalization is to preserve the proportions between the values (preserve dispositions by Eq.(6)). In accordance with Section 2.3, dispositions are invariant with respect to linear data transformations. Inverse normalizations iMax, iSum, iVec are nonlinear, and, therefore, the data and the object are distorted. The acceptability of formulas (13)-(16) is due only to the fact that these distortions are not so great, and in some cases do not affect the final result,

3) the second requirement for normalization is the agreement of the domains of normalized values of different attributes. Agreement of the normalizations for the benefit criteria and the cost criteria is achieved if the domains of the normalized values x_{ij} are comparable. With the $1/a$ transformation, this requirement is met only for the iMax normalization, but for the iSum and iVec methods this is not the case. The domains of values for the benefit and cost criteria are shifted, which leads to a change in the rating.

3.3. Reverse Sorting inversion (ReS)

The Reverse Sorting algorithm is universal, preserves value dispositions, preserves domain positions, and can be applied to both the original (17) and normalized (18) data set for jth criteria [7, 8]:

$$a_{ij} := -a_{ij} + a_j^{max} + a_j^{min}, \text{ for } j \in C_j^-, \quad (17)$$

$$x_{ij} := -x_{ij} + x_j^{max} + x_j^{min}, \text{ for } j \in C_j^-. \quad (18)$$

According to (17) and (18), the following are possible for the normalization option. First, using (17), an inversion is performed for the cost criteria for the original decision matrix, and then one of the linear normalization methods is applied. Another option is to first perform one of the linear normalization methods for the entire decision matrix, and then, using (18), an inversion of the normalized cost criteria values is performed.

In the study [7] it is shown that $\text{ReS}(\text{Max}(a)) = \text{Max}(\text{ReS}(a))$, but $\text{ReS}(\text{Sum}(a)) \neq \text{Sum}(\text{ReS}(a))$ and $\text{ReS}(\text{Vec}(a)) \neq \text{Vec}(\text{ReS}(a))$. The dispositions of the values are preserved, but there is a shift in the domains for the case of $\text{Sum}(\text{ReS}(a))$ and $\text{Vec}(\text{ReS}(a))$ from the corresponding domains when normalizing $\text{Sum}(a)$ and Vec(a). Therefore, it is recommended to use the sequence: normalization-inversion, i.e. formula (18).

The ReS transformation equally inverts both the cost criteria and the benefit criteria. For example, you can invert the values of the cost attributes and then solve the LTB problem, or invert the values of the benefit attributes and then solve the STB problem.

As is known, for optimization problems, the LTB goal is changed to the STB goal by simply changing the sign of the objective function. The known inversions using this simple technique are presented above in Section 3.1:

$$x_{ij} := -x_{ij}, \tag{19}$$

$$x_{ij} := 1 - x_{ij}. \tag{20}$$

In the first case, the cost criteria domain becomes negative in contrast to the positive values for the benefit criteria. In the second case, there is a strong shift in the data — antiphase. It is obvious that ReS is a modification of the above methods. In formulas (17) and (18), an exact shift in the domain of values is performed, which allows preserving the domain of values after the inversion. This is precisely what eliminates the problem of changing the contribution of the attribute to the integral rating under the shift.

According to the results of Section 2.3, the WSM (or TOPSIS) model with the ReS inversion is equivalent to the WSM model with the inversion $x_{ij} := -x_{ij}$ (for $j \in C_j^-$ — cost criteria) and is also equivalent to the WSM with the inversion $x_{ij} := 1 - x_{ij}$ (for $j \in C_j^-$ — cost criteria), where x_{ij} is obtained using any linear normalization method. In all three cases, ΔQ_i are the same, which produces the same ranks of alternatives. In particular, this property excludes the approach of separately accounting for the contribution of benefit and cost criteria in the WSM model (the Ratio System approach) in the form:

$$Q_i = \sum_{j=1}^{g} w_j x_{ij} - \sum_{j=g+1}^{n} w_j x_{ij}, \tag{21}$$

Where g is the number of benefit criteria.

Note also that the ReS transformation is equivalent to the generally accepted inversion for the Max-Min normalization method:

$$x_j^{max} + x_j^{min} - x_{ij} = 1 + 0 - \frac{a_j^{max} - a_{ij}}{a_j^{max} - a_j^{min}} = \frac{a_{ij} - a_j^{min}}{a_j^{max} - a_j^{min}}. \tag{22}$$

A significant advantage of the ReS transformation is the universality of combination with any normalization methods (including normalization of TTB-type criteria and nonlinear normalization) and combination with any methods of aggregation of various attributes.

3.4. Illustrations comparing the nonlinear 1/a inversion and the linear ReS inversion

The final illustration in Fig. 1, 2 and 3 compares the nonlinear inversion $1/a$ and the linear ReS for the Max, Sum and Vec normalizations for the following vector of one of the attributes for 8 alternatives: [1056 2680 1230 1480 1350 2065 1650 1750]. The ReS transformation is performed for the pre-normalized data using formula (17). Figure 1 illustrates the linearity property for ReS and the nonlinearity of the inversion based on the $1/a$ transformation for the Max, Sum and Vec normalization.

The areas of the normalized values are equal only for the Max/ iMax normalization, but there is a slight shift for the Sum/ iSum and Vec/ iVec methods. For the case of iSum, iVec, a shift in the domain of values towards a decrease is observed, which means a decrease in the contribution of this attribute to the overall integral indicator. Figure 2 shows the dispositions between the normalized values for ReS and the nonlinearity of the $1/a$ transformation.

Figure 2 shows the dispositions Δx_i by Eq.(6) between the normalized values for ReS and the nonlinear inversion of iSum. The ReS transformation preserves the dispositions Δx_i of the original data and has a center of symmetry when displayed. For the inversion of iMax, iSum,

iVec, the dispositions do not correspond to the dispositions of the original data; the center of symmetry is absent. For some points, the dispositions differ by up to $\times 2$ times. This is a significant violation of the structure of the original data. Note that Δx_i for iSum and iVec will be the same as for iMax, since the formulas differ only in the denominator, which is constant for fixed j and therefore does not affect Δx_i (see also the values of Δx_i in Figure 2). Note also that ReS inversion together with linear normalization (Sum, Vec, Max-Min,...) preserves the dispositions of the attribute values unchanged according to property (5) or (6). They are the same as for natural values. This is the universality of ReS inversion.

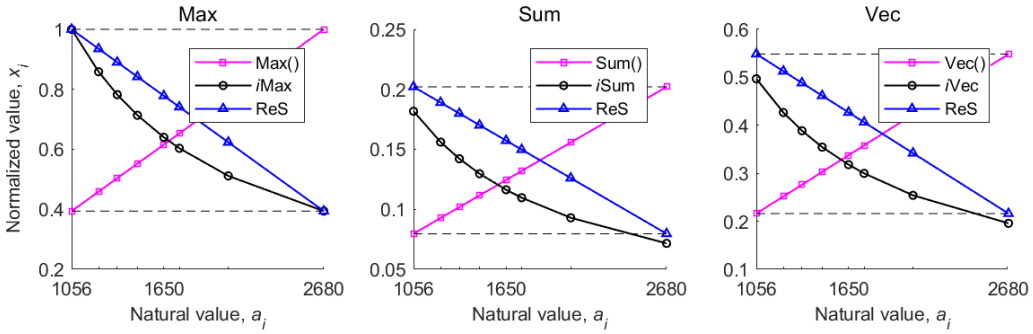


Figure 1: *Normalized and inverted values using the 1/a and ReS technique.*

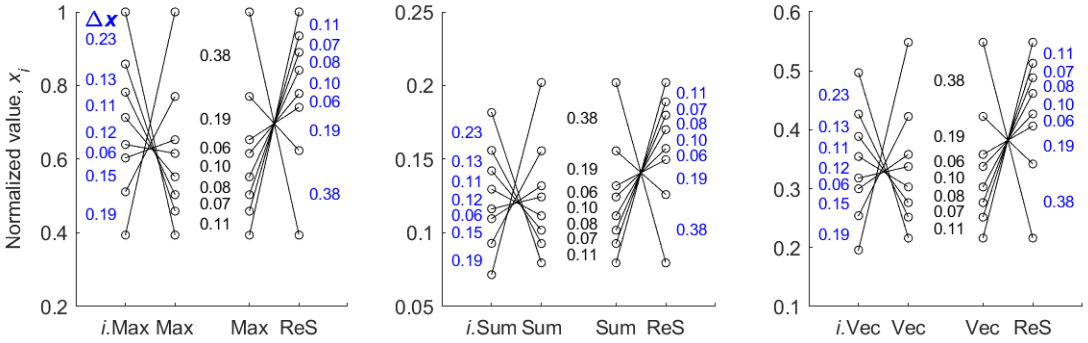


Figure 2: *Distortion of dispositions with a nonlinear inversion of 1/a compared to a linear inversion of ReS.*

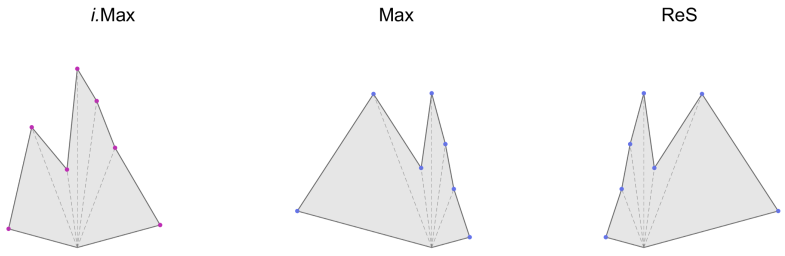


Figure 3: *Polyhedron of attribute dispositions for different normalization/inversion methods.*

Figure 3 shows the polyhedron of attribute dispositions for the original data normalization method Max and the inversion of iMax and ReS. The data polyhedron for Max normalization (which is the same for natural attribute values) coincides with the polyhedron for data

obtained with ReS inversion up to mirror symmetry, but differs from the data polyhedron for iMax normalization. Figure 3 demonstrates object scaling during ReS transformation, i.e. the proportions in the data are preserved, unlike the nonlinear transformation based on $1/a$.

3.5. A simple example demonstrating that the widely used $1/a$ transformation for inverting values of attributes introduces structural inconsistencies into normalized data

In addition to the example presented in Section 3.4, we will give a simpler and more illustrative example of incorrectness of the inversion of the $1/a$ type. Let the settlements A, B, and C have 1, 2, and 3 thousand people, respectively. Several more features are defined for these three settlements, which we will omit from consideration. Let the problem be to select one of the settlements for the construction of a logistics warehouse. Let the decision maker give priority to the settlement where the smallest number of people live. Accordingly, "opulation" is the STB criterion. We normalize and invert the data using the Max, iMax, and ReS methods. The results are presented in Table 1.

Settlement	Population (a), 1000 people	x=Max(a)	ReS(a)		iMax		ReS	
			a	$\Delta a, \%$	x	$\Delta x, \%$	x	$\Delta x, \%$
A	1	1/3	3	-	1	-	1	-
B	2	2/3	2	50	1/2	75	2/3	50
C	3	1	1	50	1/3	25	1/3	50

Table 1: *Results of normalization/inversion for the problem of choosing a settlement.*

Now let's compare the ratios Δ of the population in settlement A, B and C obtained by different methods using the formula of disposition (5). The iMax inversion significantly changed the population shares in localities A, B and C after normalization by 75 and 25%, respectively, compared to the original values of 50 and 50%. This simple example demonstrates that the nonlinear inversion of cost attributes of the 1/a type introduces structural contradictions into the normalized data. The current data do not correspond to the attributes of the original object and, therefore, the results of further data processing will be incorrect.

4. Numeric examples

Example 1. The first example shows how significantly the ranking of alternatives A_1 and A_4 for the CODAS/iMax method changed when using the ReS inversion (Table 2 and 3). The original data (Table 2) are from the article promoting the CODAS method [3].

Criteria	C_1^+	C_2^+	C_3^-	C_4^+	C_5^-
A_1	60.00	0.40	2540	500	990
A_2	6.35	0.15	1016	3000	1041
A_3	6.80	0.10	1727	1500	1676
A_4	10.00	0.20	1000	2000	965
A_5	2.50	0.10	560	500	915
A_6	4.50	0.08	1016	350	508
A_7	3.00	0.10	1778	1000	920
$Weight, w_j$	0.036	0.326	0.192	0.326	0.120

Table 2: *Input data: Decision matrix.*

The results of solving the multi-criteria choice problem for various MCDM models are presented in Table 3.

	CODAS/Max/iMax			CODAS/Max/ReS			WSM/Max/iMax			WSM/Max/ReS		
R	Q_i	A_i	ΔQ_i	Q_i	A_i	ΔQ_i	Q_i	A_i	ΔQ_i	Q_i	A_i	ΔQ_i
1	2.19	2	-	2.31	2	-	0.63	2	-	0.68	2	-
2	1.21	1	22.8	1.26	4	26.0	0.56	4	18.1	0.61	4	18.2
3	1.11	4	2.5	0.86	1	9.7	0.53	1	8.7	0.53	1	22.0
4	-0.31	3	32.9	-0.33	3	29.5	0.43	3	25.7	0.47	3	14.9
5	-0.47	5	3.9	-0.78	5	11.1	0.40	5	9.4	0.40	5	20.2
6	-1.63	7	27.0	-1.58	7	19.6	0.32	7	19.9	0.36	7	9.6
7	-2.10	6	10.9	-1.74	6	4.0	0.25	6	18.3	0.30	6	15.1

Table 3: *Rank list (R) of alternatives A_i and performance indicators Q, $\Delta Q(\%)$ for models CODAS/Max/iMax, CODAS/Max/ReS, WSM/Max/iMax, WSM/Max/ReS.*

Even a small difference in the inversion of just one attribute changed the ranks, not to mention the ratings. WSM in this example serves as a test method. The value of ΔQ for the sorted list of ratings shows how much the alternatives of rank p and $p+1$ differ. With a small difference, such alternatives can be defined as indistinguishable or as alternatives of the same rank [8, 10].

Example 2. The second example shows how significantly the weight of the criteria for the objective MEREC/Max/iMax method has changed in contrast to MEREC/Max/ReS. The original data (Table 4) are from the article promoting the MEREC method [6].

Criteria	C_1^+	C_2^+	C_3^+	C_4^-	C_5^-	C_6^-	C_7^-
A_1	23	264	2.37	0.05	167	8900	8.71
A_2	20	220	2.20	0.04	171	9100	8.23
A_3	17	231	1.98	0.15	192	10800	9.91
A_4	12	210	1.73	0.20	195	12300	10.21
A_5	15	243	2.00	0.14	187	12600	9.34
A_6	14	222	1.89	0.13	180	13200	9.22
A_7	21	262	2.43	0.06	160	10300	8.93
A_8	20	256	2.60	0.07	163	11400	8.44
A_9	19	266	2.10	0.06	157	11200	9.04
A_{10}	8	218	1.94	0.11	190	13400	10.11

Table 4: *Input data: Decision matrix.*

The weights of the criteria are presented in Table 5.

method	w_1	w_2	w_3	w_4	w_5	w_6	w_7
$w^{(1)}: MEREC, Max/iMax$	0.324	0.055	0.086	0.368	0.044	0.077	0.045
$w^{(2)}: MEREC, Max/ReS$	0.266	0.058	0.082	0.409	0.049	0.085	0.050
Relative change,%	-18	5	-5	11	11	10	11

Table 5: *Criteria weights based on the MEREC method, using iMax and ReS inversion.*

For the MEREC method, when using the linear inversion of ReS, the weight of the first and third criteria decreased, and the weight of the remaining criteria increased. The weight changes are significant. Table 6 presents the results of ranking alternatives for the WSM model using different weighting coefficients (according to Table 5) and within the iMax and ReS inversion methods.

	$w^{(1)}, iMax$			$w^{(1)}, ReS$			$w^{(2)}, iMax$			$w^{(2)}, ReS$		
R	Q_i	A_i	ΔQ_i	Q_i	A_i	ΔQ_i	Q_i	A_i	ΔQ_i	Q_i	A_i	ΔQ_i
1	0.930	2	-	0.969	1	-	0.937	2	-	0.967	1	-
2	0.913	1	3.6	0.931	2	7.8	0.905	1	6.7	0.938	2	5.8
3	0.828	7	18.6	0.917	7	2.8	0.818	7	18.3	0.917	7	4.2
4	0.785	9	9.5	0.884	8	6.6	0.780	9	8.1	0.883	8	6.6
5	0.778	8	1.3	0.873	9	2.1	0.766	8	2.8	0.878	9	1.0
6	0.589	3	41.5	0.659	3	43.3	0.571	3	41.2	0.649	3	45.6
7	0.565	5	5.2	0.646	5	2.7	0.553	5	3.9	0.643	5	1.4
8	0.550	6	3.2	0.641	6	1.0	0.541	6	2.4	0.642	6	0.1
9	0.481	10	15.2	0.586	10	11.0	0.488	10	11.2	0.606	10	7.2
10	0.472	4	2.0	0.473	4	22.8	0.463	4	5.3	0.465	4	28.0

Table 6: *Rank list (R) of alternatives A_i and performance indicators Q, $\Delta Q(\%)$ for models WSM/Max/iMax, WSM/Max/ReS.*

MEREC weighting method use iMax and ReS inversion. In this example, the ratings and ranks of alternatives change in the WSM/Max/iMax and WSM/Max/ReS models with fixed weights. For the model with different weights: $w^{(1)}$ and $w^{(2)}$ the integral rating of alternatives changes insignificantly, and the ranks of alternatives do not change.

Example 3. The third example shows how significantly the weight of criteria for the objective SECA, Max/iMax method has changed in contrast to SECA, Max/ReS method. The original data (Table 4) are from the article promoting the SECA method [2]. The weights of the criteria are presented in Table 7.

method	w_1	w_2	w_3	w_4	w_5	w_6	w_7
$w^{(3)} : SECA, Max/iMax$	0.153	0.151	0.126	0.178	0.118	0.159	0.116
$w^{(4)} : SECA, Max/ReS$	0.156	0.147	0.124	0.173	0.113	0.169	0.118
Relative change,%	2	-3	-2	-3	-4	6	2

Table 7: *Criteria weights based on the SECA method, using iMax and ReS inversion.*

For the SECA method, when using the linear inversion of ReS, the weights of the first, sixth and seventh criteria increased, and the weights of the remaining criteria decreased. The change is small, about 5%, which apparently will not affect the ranking. Table 8 presents the results of ranking within the weight change.

	$w^{(3)}, iMax$			$w^{(3)}, ReS$			$w^{(4)}, iMax$			$w^{(4)}, ReS$		
R	Q_i	A_i	ΔQ_i	Q_i	A_i	ΔQ_i	Q_i	A_i	ΔQ_i	Q_i	A_i	ΔQ_i
1	0.939	1	-	0.967	1	-	0.940	1	-	0.968	1	-
2	0.921	2	5.5	0.932	7	10.3	0.922	2	5.6	0.931	7	10.7
3	0.884	7	11.8	0.924	2	2.4	0.884	7	12.0	0.925	2	2.0
4	0.856	8	8.8	0.912	8	3.4	0.856	8	8.9	0.911	8	4.1
5	0.847	9	2.9	0.895	9	5.0	0.846	9	3.1	0.893	9	5.1
6	0.711	3	42.9	0.750	3	42.4	0.713	3	41.9	0.752	3	41.5
7	0.698	5	3.9	0.741	5	2.5	0.699	5	4.5	0.741	5	3.1
8	0.678	6	6.3	0.725	6	4.8	0.679	6	6.4	0.724	6	5.0
9	0.632	10	14.5	0.684	10	11.9	0.632	10	14.9	0.682	10	12.3
10	0.621	4	3.5	0.625	4	17.2	0.623	4	2.8	0.627	4	16.2

Table 8: *Rank list (R) of alternatives A_i and performance indicators Q, $\Delta Q(\%)$ for models WSM/Max/iMax, WSM/Max/ReS.*

SECA weighting method use iMax and ReS inversion. In this example, the ratings and ranks of the alternatives change in the WSM/Max/iMax and WSM/Max/ReS models with fixed weights. For the model with different weights: $w^{(3)}$ and $w^{(4)}$ the integral rating of the alternatives changes insignificantly, and the ranks of the alternatives do not change.

5. Conclusions

All the above arguments and examples indicate that the use of nonlinear inversion of the $1/a$ type transforms the data incorrectly. Nonlinear data inversion does not have a reasonable interpretation of values and leads to a violation of mutual distances in the original data. The measurement scales of various attributes are not consistent and there is a shift in the domains of normalized values. The ReS algorithm eliminates all the shortcomings and is recommended for use in inversion of values when coordinating the optimization goals of a multi-criterial problem.

ReS inversion is superior in all respects to existing inversion methods. Cost criteria are inverted consistently with normalized benefit criteria. It should be emphasized that ReS inversion is a logical development of the -x, and (1-x) transformations.

There is no need to exclude from the collection of methods MCDM methods that use the $1/a$ inversion. The authors systematically promote the concept of a multi-model [10], including the triple (F, W, $Norm$) of Normalization/Inversion-Weighting-Aggregation methods. It is sufficient to use ReS inversion instead of the nonlinear $1/a$ inversion. There are no technical problems to replace the normalization/inversion method in any of the MCDM methods. Examples are Max/ReS, Sum/ReS, Vec/ReS and others. Moreover, the ReS transformation is compatible and consistent with any of the nonlinear normalization methods. In this case, changes in the rating of alternatives in many cases, as shown in the examples of our article, will be insignificant and may not lead to changes in the rank of alternatives.

Data inversion is also relevant in objective methods of assessing the weight of criteria based on information contained in the decision matrix. Therefore, in such methods as EWM, MEREC, SECA, CILOS, instead of nonlinear inversion $1/a$, it is recommended to use ReS inversion.

In MCDM methods, it is difficult to check the adequacy of the result, since the optimality criterion is relative. However, our criticism of nonlinear inversion is simple and clear — it is a violation of mutual distances in the data, which distorts the description of the objects of choice.

A more radical result is presented in a recently published [11]. It is shown that as soon as MCDM models use the same linear normalization method and linear inversion, it turns out that a whole range of MCDM models (WSM, RS, MABAC, TOPSIS(L_1), MAIRCA and RAWEC) are equivalent — they always have the same ranking of alternatives and the same dispositions of the performance indicator.

One of the unresolved issues remains the question of which data set should the ReS inversion be applied to, the original data, according to (17), or the normalized data, according to (18). Although in both cases the dispositions in the data are preserved, the values are shifted. The authors also believe that further research is required in the direction of synthesizing a solution based on the ratings obtained in different MCDM models combining different combinations of the Normalization/Inversion-Weighting- and Aggregation methods. The tendency to synthesize solutions based on a large number of MCDM models is based on the assumption that each method includes different relationships between alternatives, which increases the overall information content and improves the reliability of the result. This premise suggests defining a list of acceptable methods for solving a specific problem. These methods should be independent and reflect different relationships between alternatives.

Acknowledgements

This work was carried out within the framework of the state assignment of the Ministry of Science and Higher Education of the Russian Federation (no. 075-00571-25-00 for the period 2025-2027).

References

[1] Ayan, B., Abacıoğlu, S. and Basilio, M.P. (2023). A Comprehensive Review of the Novel Weighting Methods for Multi-Criteria Decision-Making. Information, 14(5), 285. doi: 10.3390/info14050285

[2] Ghorabaee, M.K., Amiri, M., Zavadskas, E.K., Turskis, Z. and Antuchevičienė, J. (2018). Simultaneous Evaluation of Criteria and Alternatives (SECA) for Multi-Criteria Decision-Making. Informatica, 29(2), 265-280. doi: 10.15388/INFORMATICA.2018.167

[3] Ghorabaee, M.K., Zavadskas, E.K., Turskis, Z. and Antuchevičienė, J. (2016). A new Combinative Distance-Based Assessment (CODAS) method for Multi-Criteria Decision-Making. Economic Computation and Economic Cybernetics Studies and Research, 50, 25-44. url: etalpykla.vilniustech.lt/handle/123456789/116529 [Accessed 13/11/2025]

[4] Hwang, C. L. and Yoon, K. (1981). *Multiple Attributes Decision Making: Methods and Applications. A State-of-the-Art Survey.* Berlin, Heidelberg: Springer-Verlag.

[5] Jahan, J. and Edwards, K.L. (2015). A state-of-the-art survey on the influence of normalization techniques in ranking: Improving the materials selection process in engineering design. Materials Design, 65, 335–342. doi: 10.1016/j.matdes.2014.09.022

[6] Keshavarz-Ghorabaee, M., Amiri, M., Zavadskas, E.K., Turskis, Z. and Antucheviciene, J. (2021). Determination of Objective Weights using a new MEthod Based on the Removal Effects of Criteria (MEREC). Symmetry, 13(4), 525. doi: 10.3390/sym13040525

[7] Mukhametzyanov, I. Z. (2020). ReS-algorithm for converting normalized values of cost criteria into benefit criteria in MCDM tasks. International Journal of Information Technology and Decision Making, 19(5), 1389–1423. doi: 10.1142/S0219622020500327

[8] Mukhametzyanov, I. Z. (2023). *Normalization of Multidimensional Data for Multi-Criteria Decision Making Problems: Inversion, Displacement, Asymmetry.* Springer.

[9] Mukhametzyanov, I. Z. (2024). Elimination of the domain's displacement of the normalized values in MCDM tasks: the IZ-method. International Journal of Information Technology and Decision Making, 23(1), 289-326. doi: 10.1142/S0219622023500037

[10] Mukhametzyanov, I. Z. and Pamucar, D. (2024). "Thin" Structure of Relations in MCDM Models. Equivalence of the MABAC, TOPSIS(L1) and RS Methods to the Weighted Sum Method. Decision Making: Applications in Management and Engineering, 7(2), 418–422. doi: 10.31181/dmame7220241088

[11] Mukhametzyanov, I. Z. and Pamučar, D. (2025). Equivalence of MCDM Methods and Synthesis of Solution Based on Ratings Obtained in Different Models. Decision Making: Applications in Management and Engineering. 8(2), 1–20. doi: 10.31181/dmame8220251473

[12] Tzeng, G-H. and Huang, J-J. (2011). *Multiple Attribute Decision Making: Methods and Application.* Chapman and Hall/CRC.

[13] Wu, J., Sun, J., Liang, L. and Zha, Y. (2011). Determination of weights for ultimate cross efficiency using Shannon entropy. Expert Systems with Applications, 38(5), 5162–5165. doi: 10.1016/j.eswa.2010.10.046

[14] Zavadskas, E. K., Kaklauskas, A., Peldschus, F. and Turskis, Z. (2007). Multi-attribute assessment of road design solutions by using the COPRAS method. The Baltic Journal of Road and Bridge Engineering, 2(4), 195-203. url: bjrbe-journals.rtu.lv/bjrbe/article/view/1822-427X.2007.4.195%E2%80%93203 [Accessed 13/11/2025]

[15] Zavadskas, E.K., Antuchevičienė, J., Šaparauskas, J. and Turskis Z. (2013). MCDM methods WASPAS and MULTIMOORA: verification of robustness of methods when assessing alternative solutions. Economic computation and economic cybernetics studies and research, 47(2), 5–20. etalpykla.vilniustech.lt/handle/123456789/143184 [Accessed 13/11/2025]

[16] Zavadskas, E.K. and Turskis, Z. (2010). A new Additive Ratio Assessment (ARAS) method in Multicriteria Decision-Making. Technological and Economic Development of Economy, 16(2), 159-172. doi: 10.3846/tede.2010.10

Croatian Operational Research Review
CRORR **17:2**(2026), 209–220

Optimizing an imperfect production system with varying production cost under selling price and advertising-driven demand

Shilpy Tayal[1], **Shiv Raj Singh**[2], **Mohit Rastogi**[3,*], **and Saurabh Bahuguna**[1]

[1] *Department of Mathematics, Graphic Era Hill University, Bell Road Society Area, Clement Town, Dehradun, Uttarakhand, 248002, India*
E-mail: ⟨agarwal_ shilpy83@yahoo.com,bahugunasauri1996@gmail.com⟩

[2] *Department of Mathematics, Ch. Charan Singh University, Main Road, New Prabhat Nagar, Meerut, Uttar Pradesh, 250001, India*
E-mail: ⟨shivrajpundir@gmail.com⟩

[3] *Department of Applied Sciences and Humanities, ABES Engineering College, Ghaziabad, Uttar Pradesh, 201009, India*
E-mail: ⟨mohitrastogi84@gmail.com⟩

Abstract. This study develops a production inventory model that accounts for imperfect production processes, acknowledging that a certain percentage of items produced are defective, reason being labor inefficiencies and machinery malfunctions. These defective items are identified and are segregated during the quality check and inspection phase and sold at a discounted price. The model incorporates economies of scale, whereby production costs decreases as productivity increases, and represents production costs as a diminishing function of productivity. Furthermore, the demand rate is modeled as dependent on advertising frequency and selling price, impacting significantly the product demand. The primary objective of this research is to determine the optimal advertising frequency, selling price, and the production cycle that maximize overall system profit. A numerical illustration is provided to validate the model, and a sensitivity analysis is conducted to assess the stability of the system. The findings reveal that increased demand reduces production time and raises pricing potential that leads to boosting of profits. Higher production rates improve efficiency and lower unit costs, while a higher defect rate and rising holding costs adversely impact profitability, which illustrates the importance of quality control and efficient inventory management. These insights highlight the model's practicality and innovation, offering valuable guidance for optimizing operations, managing costs, and achieving sustainable growth.

Keywords: advertisement dependent demand, imperfect production, production model, selling price, variable production cost

Received: April 26, 2025; accepted: December 9, 2025; available online: February 9, 2026

DOI: 10.17535/crorr.2026.0016

Original scientific paper.

1. Introduction

Many inventory practitioners investigated numerous aspects of inventory modeling in the recent decades by assuming an unvarying demand rate. Yet, in reality, an item's demand, over the time remains in a dynamic condition. A product's selling price is an extremely crucial aspect in today's cut-throat market competition. In addition to price, advertising is another marketing factor that influences demand. Advertising enables businesses to reach a broader audience, including new or untapped markets. As a result, it can lead to increased demand from previously untargeted consumer segments. Thus, advertising is a vital tool for businesses to attract

*Corresponding author.

customers, increase sales, and stabilize success in competitive markets. Consequently, an item's demand may be thought of as being impacted by both its advertisement's frequency as well as the selling price. By incorporating price-reliant demand, researchers like [9], [11] and [20] presented their studies for deteriorating items. An inventory framework was discussed by [8] for products that do not deteriorate immediately, with a price-reliant demand. With sales returns, scrap as well as reworking, [24] created models in an imperfect production system. However, authors considered the demand rate as price-reliant in their study. [12] presented an inventory framework with deterioration as well as amelioration under trade-credit policy. Shortages were not permissible in this analysis, while the demand rate was thought to be price-reliant.

In all the above models the impact of advertising on demand was not taken into account in their studies. But authors like [19] and [17] projected their studies by incorporating demand as dependent upon the advertisement's frequency as well as the selling price. A retailer's joint pricing study was explored by [16] by utilizing a trade-credit strategy that incorporates price as well as advertisement's frequency dependent demand. Authors permitted shortages in this analysis while both the deterioration rate and the cost of holding of items were constant. Recently, a supply chain paradigm was presented by [6] incorporating demand as dependent upon advertisement's frequency as well as the selling price. Authors allowed shortages to occur in this presented study.

Among the most crucial components of a business tactic is its production system. An efficient production system ensures that resources are used optimally, minimizing waste of materials, time, and labor. This leads to cost savings and improved profitability for businesses. To minimize product overstocks and shortages, the firm should properly manage its production process in order to benefit the business. Researchers like [27] and [23] introduced production inventory models involving decaying products. An EPQ paradigm for items that were ameliorating deteriorating in nature by incorporating ramp type demand was formulated by [10]. For decaying goods which were non-instantaneous in nature, [25] gave an EPQ framework under no shortages. This study evaluated holding cost as contingent on time whereas the rate of production was treated as demand rate dependent. To identify the optimum rate of production, [1] proposed a production inventory framework for an agricultural commodity. Costs such as deterioration, transportation and holding cost were considered as unvarying in the proposed study. After that, [13] presented a production model by incorporating advertisement and price-reliant demand for defectives. As production rate was treated as higher in comparison of the demand rate, thus shortages were not permitted in this developed study.

In all the above models impact of varying production cost has not been considered in their study while it is essential in determining a firm's profitability, competitiveness, pricing strategy and overall financial health. It is commonly seen that when demand raises, production rates rise, which decreases the cost of production per unit. Hence, the unit production cost per unit depends upon production or demand. In line with this, only few authors considered variable production cost in their studies. [14] proposed a deterministic model by incorporating variable production cost as well as stock-reliant demand for damageable products. [22] explored an imperfect production system stating the unit production cost was treated as a function of the finite production rate. Two EPQ models with or without shortages were studied by [15] for deteriorating items. For developing models, authors treated that the production rate along with the unit production cost were inversely correlated whereas rate of production rate was proportional to the demand rate. [5] demonstrated a production replica by treating the production cost per unit that significantly increases with regard to production rate of the system, where as in the presented model we have considered that as the system's production rate increases, the cost of production per unit decreases.

A number of research studies that are available believe the products generated by a production system are perfect. While in reality it is not true. Due to the various challenges encountered in a lengthy production process, a certain ratio of produced items is defective. Therefore, the

things that are defective may be addressed as a consequence of the production's imperfection. Authors, including [3] and [4] presented their inventory models incorporating imperfect production. After that, [21] presented imperfect production structure that incorporates an EPL framework under no shortages. In this study, the production's rate was used as a decision-making measure. [26] introduced their study for deteriorating and imperfect quality under no shortages. However, the demand rate, deterioration rate and the production rate was unvarying in nature. In support of imperfect quality products, [18] demonstrated an inventory paradigm by including partial trade credit. But demand for the products was treated to be reliant upon advertisement, only in the mentioned model. After that, [7] gave a framework than incorporates an imperfect production model under screening and shortage constraints, considering a non-linear time-reliant demand function in this study. In order to maximise total profit, [2] recommended an imperfect manufacturing system. This research was developed under shortages by incorporating demand as reliant upon selling price only.

2. Research gap

A key research gap identified in the existing literature on inventory and production models is the limited exploration of variable production costs in imperfect production systems, particularly where production costs are influenced by economies of scale. While some studies have incorporated price and advertisement-dependent demand and even addressed imperfect production and defect rates, many models assume fixed production costs, which do not account for the real-world scenario, where the production costs per unit decrease as production rates increase. Furthermore, although some studies have addressed imperfect production, few have considered the combined impact of both advertising and production costs on overall profitability. Existing models also often neglect fluctuating production costs as a function of both demand and production rate, which is crucial for accurate decision-making in dynamic and competitive markets. Addressing this gap by developing a model that integrates variable production costs, selling price, advertisement-dependent demand, and imperfect production processes could lead to a more realistic approach to optimizing inventory and production systems. Thus,

- A production inventory model with imperfect production has been presented.
- Several business tactics, like advertisement frequency, selling price, and variable production cost are combined to create a system to support decisions.
- As the production process is not 100% reliable; during the inspection procedure, defective products are isolated and offered for sale at a reduced cost.
- The production rate is thought to be reliant on demand and reflects real-world scenarios.
- Production cost per unit reduces as the production rate of the system rises
- Optimal values of advertisement frequency, selling price, and production period are calculated, which optimize the system's overall average profit.

3. Assumptions and notations

The underlying presumptions that guided the creation of the suggested model are outlined in this part.

1. This is a single item inventory system with finite rate of replenishment.
2. The period between placing an order and its delivery is assumed to be negligible.
3. The rate of demand is reliant upon advertisement's frequency as well as selling price which is given by $(\alpha - \beta P^{\gamma})A^{\theta}$ where $\alpha > 0,\ 0 < \beta < 1,\ \lambda > 0,\ 0 < \theta < 1$.
4. Production rate is regarded as contingent upon demand rate which is given by $P = \eta(\alpha - \beta P^{\gamma})A^{\theta}$.

5. An inspection is done after the production of items. The inspection cost is given by $c_i = c_2 + c_3 \int_0^{t_1} P dt$, where c_2 represents the initial cost to set up the inspection process, regardless of the number of items produced; this cost remains constant and c_3 represents the cost per unit of inspection effort.
6. During assessment, a specific percentage of the total number of produced is discovered to be defective and these flawed products are available for sale at a discounted value.
7. This model assumes that the production cost per unit decreases as the system's production rate rises. It is given by the term $(\lambda - \mu P)$. Here $(\lambda - \mu P)$ must remain positive for meaningful production. If P exceeds a certain level, the cost per unit could become zero or negative, which is unrealistic. λ reflects the base cost of manufacturing one unit without considering production scale. μ is the cost sensitivity to production rate. As the production rate (P) increases, the per-unit cost decreases. This models situations where higher production reduces fixed costs per unit or achieves efficiencies (e.g., bulk purchasing, better utilization of machinery).
The key applications for this cost function are mentioned as follows.
Cost Optimization: Businesses can use this function to determine the optimal production rate (P) where per-unit costs are minimized, balancing economies and diseconomies of scale.
Pricing Strategies: Understanding how costs change with production allows firms to set prices competitively while ensuring profitability.

The notations are as follows:

$\alpha, \beta, \gamma, \theta$: demand parameters,

η : production parameter, $\eta > 1$,

A : number of advertisement frequencies conducted per month, $A \in \mathbb{N}$,

p : selling price in \$ / unit,

T : replenishment cycle,

t_1 : production period,

λ : initial production cost in \$ / unit,

μ : production cost parameter, $\mu > 0$,

ξ : set-up cost,

δ : percentage of imperfect products, $(0 < \delta < 1)$,

c_1 : cost for per advertisement,

P : production rate,

χ : percentage discount offered to imperfect products,

h : holding cost per unit per unit time,

c_2 : fixed cost of inspection,

c_3 : variable part of inspection cost.

4. Mathematical formulation

The production cycle (Figure 1) starts with zero-inventory and production begins at $t = 0$. During time $[0, t_1]$ inventory builds up, adjusting demand in the market. At $t = t_1$, production stops. After that, through the period $[t_1, T]$, this stock level reduces owing to market demand

and approaches the zero level at $t = T$.

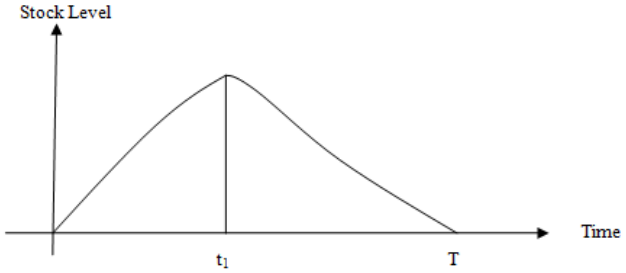


Figure 1: *Inventory behavior with respect to time.*

The differential equations showing the inventory system's behavior at any moment t are given by

$$\frac{dI(t)}{dt} = \eta(\alpha - \beta p^{\gamma})A^{\theta} - (\alpha - \beta p^{\gamma})A^{\theta}, \quad 0 \le t \le t_1, \tag{1}$$

$$\frac{dI(t)}{dt} = -(\alpha - \beta p^{\gamma})A^{\theta}, \qquad t_1 \le t \le T, \tag{2}$$

with boundary condition $I(0) = 0, I(T) = 0$.
These equations' solutions are as follows

$$I(t) = (\eta - 1)(\alpha - \beta p^{\gamma})A^{\theta}t, \quad 0 \le t \le t_1, \tag{3}$$

$$I(t) = (\alpha - \beta p^{\gamma})A^{\theta}(T - t), \quad t_1 \le t \le T, \tag{4}$$

with the help of equation (3) and (4), the relation in and T as

$$T = \eta t_1. \tag{5}$$

Total production $= \int_0^t P dt$.

$$T.P. = \eta A^{\theta}(\alpha - \beta p^{\gamma})t_1 \tag{6}$$

The expense incurred by storing goods for a period of time is known as holding cost:

$$\begin{aligned} \text{Holding cost} &= h\left(\int_0^{t_1} I(t)\,dt + \int_{t_1}^{T} I(t)\,dt\right), \\ \text{H.C.} &= \frac{hD}{2}\left[(\eta - 1)t_1^2 + (T - t_1)^2\right]. \end{aligned} \tag{7}$$

If A is the advertisement's frequency then the cost associated with the advertisement will be

$$\text{Advertisement cost} = Ac_1. \tag{8}$$

If δ is the percentage of imperfect items, then the revenue $(R1)$ generated from the perfect items will be

$$R_1 = p(1 - \delta)\eta t_1(\alpha - \beta p^{\gamma})A^{\theta}. \tag{9}$$

All the defective items are available for sale at discounted price χp where $0 < \chi < 1$ so the sales revenue (R_2) generated from the defective items will be

$$R_2 = \chi p \delta \eta t_1(\alpha - \beta p^{\gamma})A^{\theta}. \tag{10}$$

So, total sales revenue $= R_1 + R_2$,

$$\text{Sales revenue} = p(1-\delta)\eta t_1(\alpha - \beta p^{\gamma})A^{\theta} + \chi p\delta\eta t_1(\alpha - \beta p^{\gamma})A^{\theta}. \tag{11}$$

The expenses related to assembling all necessary equipment for manufacturing are provided by

$$\text{Set up cost} = \xi. \tag{12}$$

Since the production procedure is not 100% reliable, through production a few defective items are also produced. So to separate the imperfect items all the produced items has to undergo an inspection procedure. The cost associated for the inspection of items is given by

$$\begin{aligned} c_i &= c_2 + c_3 \int_0^{t_1} P dt, \\ c_i &= c_2 + c_3\eta A^{\theta}(\alpha - \beta P^{\gamma})t_1. \end{aligned} \tag{13}$$

If λ is the initial production cost per unit in that case the total production cost will be

$$\begin{aligned} P.C. &= \int_0^{t_1} P(\lambda - \mu P)dt, \\ P.C. &= \eta(\alpha - \beta p^{\gamma})A^{\theta}(\lambda - \mu\eta(\alpha - \beta p^{\gamma})A^{\theta})t_1. \end{aligned} \tag{14}$$

The system's overall average profit may now be calculated as

$$\begin{aligned} \text{T.A.P.}(t_1, A, P) = \frac{1}{\eta t_1}\big[&\text{Sales revenue} - \text{Set up cost} - \text{Inspection cost} \\ &- \text{Holding cost} - \text{Advertisement cost} - \text{Production cost}\big], \end{aligned} \tag{15}$$

$$\begin{aligned} \text{T.A.P.}(t_1, A, P) = \frac{1}{\eta t_1}\Big(&P(1-\delta)\eta t_1(\alpha - \beta p^{\gamma})A^{\theta}) + \chi p\delta\eta t_1(\alpha - \beta p^{\gamma})A^{\theta} - \xi - c_2 \\ &-c_3\eta A^{\theta}(\alpha - \beta p^{\gamma})t_1 - \frac{h(\alpha - \beta p^{\gamma})A^{\theta}}{2}[(\eta - 1)t_1^2 + (\eta t_1 - t_1^2)] - Ac_1 \\ &-\eta(\alpha - \beta p^{\gamma})A^{\theta}(\lambda - \mu\eta(\alpha - \beta p^{\gamma})A^{\theta})t_1\Big). \end{aligned} \tag{16}$$

5. Solution procedure and algorithm

Now the problem can be formulated as:

Max T.A.P.(t_1, A, p)
with constraints $0 \le t_1 \le T$ and $p > \lambda$
Under the following conditions
$\alpha > 0,\ 0 < \beta < 1,\ \gamma > 0,\ 0 < \theta < 1,\ \lambda > \mu P,\ 1 \le A \le 10$
We use the following algorithm to determine the model's optimal solution.

5.1. Algorithm

For all the constant parameters employed in the model
Step 1:Initialize $A = 1$
Step 2: Determine optimal values of decision variables t_1 and p, subject to given constraints

Step 3: Calculate the optimal solution T.A.P. (t_1, A, p)
(a) For a fixed value of A, put

$$\frac{\partial \text{T.A.P}(t_1, A, p)}{\partial t_1} = 0, \qquad \frac{\partial \text{T.A.P}(t_1, A, p)}{\partial p} = 0. \tag{17}$$

(b) Solve these equations simultaneously for critical points.
The determinant of the Hessian exists, but it's symbolic form is prohibitively complex for presentation. Thus, we determine negative definiteness by examining the leading principal minors according to Sylvester's criterion and find the optimal value.
Step 4: Set $A_1 = A + 1$
Recalculate T.A.P(t_1, A_1, p_1) using new A_1 & compare T.A.P(t_1, A_1, p_1) with T.A.P(t_1, A, p)
Step 5: If T.A.P$(t_1, A_1, p_1) >$ T.A.P(t_1, A, p), then: Set $A = A_1$
Repeat Step 3 and Step 4
Else: Proceed to Step 6
Step 6: Set $(t_1, A_1, p_1) = (t_1^*, A^*, p^*)$ as the best solution.
Step 7: Calculate T.A.P$^*(t_1^*, A^*, p^*)$
Return the optimal solution (t_1^*, A^*, p^*)
End Algorithm.

5.2. Hessian matrix

For optimal value of p and t_1. We differentiate the total profit function with respect to p and t_1 and equating to zero:

$$\frac{\partial \text{T.A.P}(t_1, A, p)}{\partial t_1} = -\frac{1}{2}A^\theta h(\alpha - p^\gamma \beta)(-1+\eta) + \frac{\xi + Ac_1 + c_2}{\eta t_1^2}, \tag{18}$$

$$\begin{aligned}\frac{\partial \text{T.A.P}(t_1, A, p)}{\partial p} =& \frac{1}{2p}A^\theta \Big(2p\alpha\{1+(-1+\chi)\delta\} - 2p^{1+\gamma}\beta(1+\gamma)\{1+(-1+\chi)\delta\} \\ &+ 4A^\theta p^2 \gamma \beta^2 \gamma \eta \mu + 2p^\gamma \beta \gamma(\lambda - 2A^\theta \alpha \eta \mu) \\ &+ 2p^\gamma \beta \gamma c_3 + h p^\gamma \beta \gamma(-1+\eta)t_1\Big).\end{aligned} \tag{19}$$

The optimal solutions p^* and ${t_1}^*$ can be obtained by solving the resultant system after setting the equations (18) and (19) to zero:

$$H = \begin{pmatrix} \frac{\partial^2 T.A.P}{\partial t_1^2} & \frac{\partial^2 T.A.P}{\partial t_1 \partial p} \\ \frac{\partial^2 T.A.P}{\partial p \partial t_1} & \frac{\partial^2 T.A.P}{\partial p^2} \end{pmatrix}, \tag{20}$$

$$\frac{\partial \text{T.A.P}(t_1, A, p)}{\partial t_1^2} = -\frac{2(\xi + Ac_1 + c_2)}{\eta t_1^3} < 0, for \;\; \xi > 0, A > 0, c_1 > 0, c_2 > 0, \eta > 0. \tag{21}$$

It shows $D_1 < 0$ (negative definite):

$$\frac{\partial^2 \text{T.A.P.}(t_1, A, P)}{\partial p^2} = \frac{1}{2}A^\theta p^{-2+\gamma}\beta\gamma B(p, t_1), \tag{22}$$

where

$$\begin{aligned}B(p, t_1) =& -2p(1+\gamma)(1+(-1+\chi)\delta) + 4A^\theta p^\gamma \beta(-1+2\gamma)\eta\mu \\ &+ 2(-1+\gamma)(\lambda - 2A^\theta \alpha \eta \mu) + 2(-1+\gamma)c_3 + h(-1+\gamma)(-1+\eta)t_1\end{aligned} \tag{23}$$

$$\frac{\partial^2 \text{T.A.P.}(t_1, A, P)}{\partial p \partial t_1} = \frac{\partial^2 \text{T.A.P.}(t_1, A, P)}{\partial t_1 \partial p} = \frac{1}{2} A^{\theta} h p^{-2+\gamma} \beta \gamma J(p, t_1). \tag{24}$$

The determinant of the Hessian is given by

$$|H| = \frac{1}{16\eta^2 t_1^3} A^{\theta} p^{-2+\gamma} \beta \gamma J(p, t_1), \tag{25}$$

where

$$\begin{aligned} J(p,t_1) = \Big(& 32p\xi\eta + 32p\gamma\xi\eta - 32p\delta\xi\eta + 32p\chi\delta\xi\eta - 32p\gamma\delta\xi\eta + 32p\chi\gamma\delta\xi\eta + 32\xi\eta\lambda \\ & - 32\gamma\xi\eta\lambda - 64A^{\theta}\alpha\xi\eta^2\mu + 64A^{\theta}p^{\gamma}\beta\xi\eta^2\mu + 64A^{\theta}\alpha\gamma\xi\eta^2\mu - 128A^{\theta}p^{\gamma}\beta\gamma\xi\eta^2\mu \\ & + 32\xi\eta c_3 - 32\gamma\xi\eta c_3 + 8h\xi t_1 - 8h\gamma\xi t_1 - 24h\eta\xi t_1 + 24h\gamma\xi\eta t_1 + 16h\xi\eta^2 t_1 \\ & - 16h\gamma\xi\eta^2 t_1 - A^{\theta}h^2p^{\gamma}\beta\gamma t_1^3 + 6A^{\theta}h^2p^{\gamma}\beta\gamma\eta t_1^3 - 13A^{\theta}h^2p^{\gamma}\beta\gamma\eta^2 t_1^3 \\ & + 12A^{\theta}h^2p^{\gamma}\beta\gamma\eta^3 t_1^3 - 4A^{\theta}h^2p^{\gamma}\beta\gamma\eta^4 t_1^3 + 8Ac_1(4\eta(p(1+\gamma)(1+(-1+\chi)\delta) \\ & - 2A^{\theta}p^{\gamma}\beta(-1+2\gamma)\eta\mu - (-1+\gamma)(\lambda - 2A^{\theta}\alpha\eta\mu)) - 4(-1+\gamma)\eta c_3 \\ & - h(-1+\gamma)(1-3\eta+2\eta^2)t_1) \quad + 8c_2(4\eta(p(1+\gamma)(1+(-1+\chi)\delta) \\ & - 2A^{\theta}p^{\gamma}\beta(-1+2\gamma)\eta\mu \quad - (-1+\gamma)(\lambda - 2A^{\theta}\alpha\eta\mu)) - 4(-1+\gamma)\eta c_3 \\ & - h(-1+\gamma)(1-3\eta+2\eta^2)t_1)\Big). \end{aligned} \tag{26}$$

The second derivative with respect to t_1 has been shown analytically to be negative under the stated parameter assumptions. The symbolic determinant of the Hessian is prohibitively large for closed-form presentation. Consequently, we verified the sign of the determinant numerically: we evaluated D_2=det(H), across the feasible ranges of p and t_1 and for the parameter values considered in this study. In all tested cases the determinant was strictly positive while $D_1 < 0$. Hence, by Sylvester's criterion the Hessian is negative definite on the feasible domain and the stationary point is a (global) maximizer in the admissible region.

6. Numerical example

These following are the inputs used to calculate the numerical results. Mathematical software Mathematica 11.3 has been used to solve the numerical example.
α=500 units, β=0.02, η=1.5, θ=0.001, γ=2.5, μ=0.1, χ=0.5, δ=0.2, ξ=1000\$ /setup, c_1=5\$ /advertisement, λ=5, c_2=500\$ /Inspection, c_3=1\$ /unit, h=0.2\$ /unit.

We treat A as an integer (ads per month) with a practical upper bound $A_{\max}$. Because $A_{\max}$ is small in realistic scenarios, we use an exhaustive linear search over $\{1, \ldots, A_{\max}\}$.This approach is exact for the discrete problem, computationally trivial, and avoids errors from continuous approximations.

A	t_1	p	T.A.P.
1	7.81631	36.568	9010.57
2	**7.93887**	**36.5699**	**9012.79**
3	8.06058	36.5718	9012.40
4	8.18089	36.5738	9010.98
5	8.29967	36.5758	9009.02
6	8.41690	36.5777	9006.73
7	8.53261	36.5797	9004.23
8	8.64685	36.5816	9001.6
9	8.75964	36.5835	8998.86
10	8.87104	36.5854	8996.06

Table 1: *Solution search procedure for optimal 'A'.*

The results show that when the value of A increases, the cycle time t_1 also increases steadily. This means that higher values of A require more time for operations. Price p does not show any major change and remains nearly constant across all values of A. The Total Annual Profit (T.A.P.) first rises and reaches its highest level when $A = 2$. After this point, the profit starts decreasing slowly as A continues to increase. Therefore, the value $A = 2$ gives the best profit outcome, and further increase in A reduces the profitability even though the cycle time becomes longer. Thus, the maximum value obtained for T.A.P. is the optimal value of the system with optimal frequency of advertisement and corresponding to it the value of production period t_1 and selling price p are also the optimal values for the given system. Here the optimal values are: $A^* = \mathbf{2}, t_1^* = \mathbf{7.93887}$ months, $p^* = \mathbf{36.5699}\$, \text{T.A.P}^* = \mathbf{9012.79}\$, T = 11.90$ months.

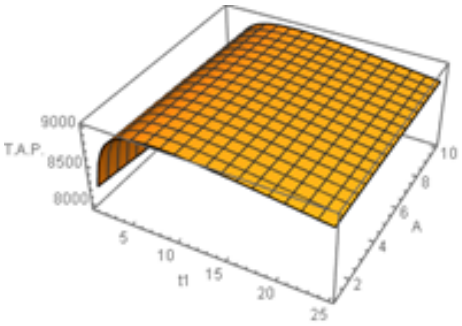


Figure 2: *Concavity of T.A.P. function for fix value of 'p'.*

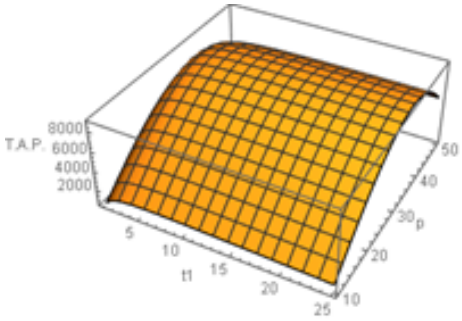


Figure 3: *Concavity of T.A.P. function for fix value of 'A'.*

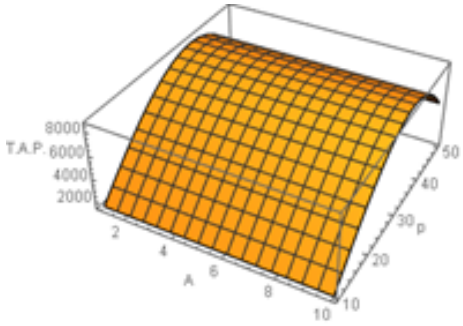


Figure 4: *Concavity of T.A.P. function for fix value of 't_1'.*

7. Sensitivity analysis

To look into how distinct parameters affect the system's total average profit, a sensitivity analysis is carried out. The results are given in Table 2. The following points are noticed:

1. With the increment in initial demand parameter 'α', t_1 decreases whereas p and T.A.P. increases. The value of A remains constant up to a certain range of variation and then shows a slight increase. This happens because higher demand leads to higher sales, which results in increased profit. The rationale is that when demand of a product raises it boost the sale of the product which result an increase in profit.

2. We noticed that as initial production cost parameter 'λ' increases, t_1 & p increases, whereas T.A.P. decreases. The rationale is that as initial production cost/unit increase, the profit per unit decreases due to reduced profit margins. As a result, the overall average profit, which is a measure of the average profit per unit produced or sold, reduces. As a result, the manufacturer needs to monitor the initial production cost parameter, since it has a quick impact on earnings.

3. As the production parameter (η) increases, t_1 and p decreases, and the Total Average Profit (T.A.P.) also decreases, which aligns with practical expectations. The reason is that when the production rate increases, the cost per unit decreases, leading to lower pricing. This ultimately reduces the overall profit earned.

4. Variation in parameter 'δ' (fraction of imperfect items) is presented in Table 2. As parameter 'δ' increases, t_1 and p increases, while T.A.P. decreases. The rationale is that a higher fraction of imperfect items reduces product quality, prompting customers to switch to competitors, offering better-quality products. This leads to reduced sales, ultimately decreasing profit. Therefore, manufacturers should pay close attention to this parameter, as it has a significant impact on earnings.

Parameter	**% Variation**	**Value**	A	t_1	p	**T.A.P.**
α	-20	400	2	8.91552	33.6938	6397.18
	-10	450	2	8.38547	35.1804	7669.40
	0	500	2	7.93887	36.5699	9012.79
	10	550	2	7.55594	37.8773	10423.2
	20	600	3	7.33358	39.1158	11897.4
λ	-20	4	2	7.89495	36.2262	9353.18
	-10	4.5	2	7.91666	36.3974	9182.51
	0	5	2	7.93887	36.5699	9012.79
	10	5.5	2	7.96158	36.7435	8844.03
	20	6	2	7.98480	36.9183	8676.24
η	-20	1.20	3	14.2561	36.6012	9057.96
	-10	1.35	3	10.1585	36.5905	9027.44
	0	1.50	2	7.93887	36.5699	9012.79
	10	1.65	2	6.63630	36.5467	9006.58
	20	1.80	2	5.72479	36.5207	9005.63
δ	-20	0.16	2	7.93363	36.5294	9260.39
	-10	0.18	2	7.93622	36.5494	9136.58
	0	0.20	2	7.93887	36.5699	9012.79
	10	0.22	2	7.94159	36.5908	8889.02
	20	0.24	2	7.94438	36.6122	8765.25

Table 2: *Sensitivity analysis for some crucial parameters.*

8. Conclusion

A realistic inventory problem with varying production costs is addressed by taking the selling price and advertisement dependent demand into consideration. Here, production cost is deemed to be a decreasing function of production rate, as it is realistic that if anyone produces more items, then the production cost per unit decreases, which results in an overall profit of the system. The results presented in sensitivity analysis confirmed this fact also. Throughout the production phase, a few flawed goods are manufactured, that are separated during the inspection procedure. These defective goods are sold at discounted price to generate profit. This research investigates the effects of key operational factors on efficiency and profitability in the production and inventory systems. The proposed model introduces a novel approach by incorporating real-world factors such as imperfect production systems, demand-dependent rates, and strategies for optimizing advertisement frequency and pricing. It addresses defective item management and demonstrates how higher production rates contribute to cost efficiency. These findings offer practical insights into optimizing production operations and achieving sustainable business growth.

References

[1] Bhattacharjee, N., Nath, B. K., Sen, N., Malakar, S., and Jaggi, C. K. (2022). A Production Inventory Model to Study the Supply Chain of Agri-Product for a Time Reliant Population. International Journal of Applied and Computational Mathematics, 8(3), 97. doi: 10.1007/s40819-022-01286-5

[2] Choudri, V., Senthilkumar, C., Sivashankari, C. K., and Helenprabha, K. (2025). Sustainable production for forward–backward shortages, reworking, scrap, with price-dependent demand: decisions on optimal pricing, lot size, and shortages. Journal of Control and Decision, 1-16. doi: 10.1080/23307706.2025.2478254

[3] Chung, C. J., and Wee, H. M. (2008). An integrated production-inventory deteriorating model for pricing policy considering imperfect production, inspection planning and warranty-period- and stock-level-dependant demand. International Journal of Systems Science, 39(8), 823-837. doi: 10.1080/00207720801902598

[4] Chung, K. J., Her, C. C., and Lin, S. D. (2009). A two-warehouse inventory model with imperfect quality production processes. Computers & Industrial Engineering, 56(1), 193-197. doi: 10.1016/j.cie.2008.05.005

[5] Das, S., Ali, H., Shaikh, A. A., and Bhunia, A. K. (2023). Impact of emission and service constraint for an imperfect production system under two level Hamiltonian. Optimal Control Applications and Methods, 44(5), 2796-2820. doi: 10.1002/oca.3003

[6] Debnath, A., and Sarkar, B. (2024). Supply chain model having stochastic lead time demand with variable production rate and demand dependent on price and advertisement. RAIRO-Operations Research, 58(4), 2645-2667. doi: 10.1051/ro/2023191

[7] Dolai, M., and Mondal, S. K. (2024). An imperfect production model with shortage and screening constraint under time varying demand. International Journal of System Assurance Engineering and Management, 15(3), 1183-1202. doi: 10.1007/s13198-023-02202-w

[8] Duary, A., Khan, M. A. A., Pani, S., Shaikh, A. A., Hezam, I. M., Alraheedi, A. F., and Gwak, J. (2022). Inventory model with nonlinear price-dependent demand for non-instantaneous decaying items via advance payment and installment facility. AIMS Mathematics, 7(11), 19794-19821. doi: 10.3934/math.20221085

[9] Dye, C. Y., Ouyang, L. Y., and Hsieh, T. P. (2007). Inventory and pricing strategies for deteriorating items with shortages: A discounted cash flow approach. Computers & Industrial Engineering, 52(1), 29-40. doi: 10.1016/j.cie.2006.10.009

[10] Goyal, S. K., Singh, S. R., and Dem, H. (2013). Production policy for ameliorating/deteriorating items with ramp type demand. International Journal of Procurement Management, 6(4), 444-465. doi: 10.1504/IJPM.2013.054753

[11] Jaggi, C. K., Pareek, S., Khanna, A., and Sharma, R. (2014). Credit financing in a two-warehouse environment for deteriorating items with price-sensitive demand and fully backlogged shortages.

Applied Mathematical Modelling, 38(21-22), 5315-5333. doi: 10.1016/j.apm.2014.04.025

[12] Jayashri, P., and Umamaheswari, S. (2025). Entrepreneurial strategies in livestock inventory management with trade credit. Contemporary Mathematics, 6(1), 516-535. doi: 10.37256/cm.6120255714

[13] Kausar, A., Hasan, A., Maheshwari, S., Gautam, P., and Jaggi, C. K. (2024). Sustainable production model with advertisement and market price dependent demand under salvage option for defectives. Opsearch, 61(1), 315-333. doi: 10.1007/s12597-023-00688-3

[14] Maiti, M. K., and Maiti, M. (2005). Inventory of damageable items with variable replenishment and unit production cost via simulated annealing method. Computers & Industrial Engineering, 49(3), 432-448. doi: 10.1016/j.cie.2005.07.004

[15] Manna, S. K., and Chiang, C. (2010). Economic production quantity models for deteriorating items with ramp type demand. International Journal of Operational Research, 7(4), 429-444. doi: 10.1504/IJOR.2010.03242

[16] Mashud, A. H. M., and Sarkar, B. (2021). Retailer's joint pricing model through an effective preservation strategy under a trade-credit policy. RAIRO-Operations Research, 55(3), 1799-1823. doi: 10.1051/ro/2021018

[17] Mondal, R., Shaikh, A. A., and Bhunia, A. K. (2019). Crisp and interval inventory models for ameliorating item with Weibull distributed amelioration and deterioration via different variants of quantum behaved particle swarm optimization-based techniques. Mathematical and computer modelling of dynamical systems, 25(6) 602-626. doi: 10.1080/13873954.2019.1692226

[18] Moradi, S., Gholamian, M. R., and Sepehri, A. (2023). An inventory model for imperfect quality items considering learning effects and partial trade credit policy. Opsearch, 60(1), 276-325. doi: 10.1007/s12597-022-00602-3

[19] Palanivel, M., and Uthayakumar, R. (2015). Finite horizon EOQ model for non-instantaneous deteriorating items with price and advertisement dependent demand and partial backlogging under inflation. International Journal of Systems Science, 46(10), 1762-1773. doi: 10.1080/00207721.2013.835001

[20] Rastogi, M., and Singh, S. R. (2019). An inventory system for varying deteriorating pharmaceutical items with price-sensitive demand and variable holding cost under partial backlogging in healthcare industries. Sādhanā, 44(4), 95. doi: 10.1007/s12046-019-1075-3

[21] Sana, S. S. (2010). An economic production lot size model in an imperfect production system. European Journal of Operational Research, 201(1), 158-170. doi: 10.1016/j.ejor.2009.02.027

[22] Sana, S.S., Goyal, S.K., and Chaudhuri, K.S. (2007). An imperfect production process in a volume flexible inventory model. International Journal of Production Economics, 105(2), 548–559. doi: 10.1016/j.ijpe.2006.05.005

[23] Sana, S., Goyal, S.K., and Chaudhuri, K.S. (2004). A production-inventory model for a deteriorating item with trended demand and shortages. European Journal of Operational Research, 157(2), 357–371. doi: 10.1016/S0377-2217(03)00222-4

[24] Sivashankari, C. K., and Valarmathi, R. (2023). Optimal pricing and production lot-size policies in imperfect production system with price-sensitive demand, reworking, scrap, and sales return. Operational Research, 23(3), 43. doi: 10.1007/s12351-023-00779-5

[25] Tayal, S., Singh, S.R., Sharma, R., and Singh, A.P. (2015). An EPQ model for non-instantaneous deteriorating item with time dependent holding cost and exponential demand rate. International Journal of Operational Research, 23(2), 145–161. doi: 10.1504/IJOR.2015.069177

[26] Tiwari, S., Daryanto, Y., and Wee, H. M. (2018). Sustainable inventory management with deteriorating and imperfect quality items considering carbon emission. Journal of Cleaner Production, 192, 281-292. doi: 10.1016/j.jclepro.2018.04.261

[27] Yang, P., and Wee, H.M. (2003). An integrated multi-lot-size production inventory model for deteriorating item. Computers & Operations Research, 30(5), 671–682. doi: 10.1016/S0305-0548(02)00032-1

Benchmarking Croatian airports: A data-driven approach through DEA window analysis

Damira Keček[1], **Katerina Fotova Čiković**[1,*] **and Violeta Cvetkoska**[2]

[1] *University North, Trg dr. Žarka Dolinara 1, 48000 Koprivnica, Croatia*
E-mail: ⟨{dkecek, kcikovic}@unin.hr⟩

[2] *Faculty of Economics – Skopje, University of St. Cyrill and Methodius, Boulevard Goce Delchev 9, Skopje 1000, North Macedonia*
E-mail: ⟨vcvetkoska@eccf.ukim.edu.mk⟩

Abstract. This study benchmarks the efficiency of seven certified international airports in Croatia using Data Envelopment Analysis (DEA) with a window framework under variable returns to scale, explicitly addressing the challenge of a small number of decision-making units through dynamic window specification. Efficiency scores were calculated across overlapping three-year periods from 2019-2023, allowing performance dynamics to be captured during both crisis and recovery phases. The results reveal substantial heterogeneity across airports. Osijek Airport emerged as the most consistently efficient, outperforming larger hubs such as Zagreb and Split, while Zadar and Rijeka demonstrated notable efficiency improvements over time. In contrast, Dubrovnik Airport recorded persistently low efficiency despite traffic recovery, suggesting structural cost rigidity and seasonal vulnerability. These findings highlight the operational resilience of smaller regional airports and signal performance constraints among major hubs. By integrating traffic trends with efficiency outcomes, the study shows that rising passenger volumes alone do not ensure improved economic performance. The results provide a timely, evidence-based framework for airport managers and policymakers to enhance strategic planning, improve resource allocation, and strengthen the financial sustainability and competitiveness of Croatia's air transport sector.

Keywords: airport efficiency, benchmarking, Croatian airports, data envelopment analysis, window analysis

Received: November 3, 2025; accepted: December 8, 2025; available online: February 9, 2026

DOI: 10.17535/crorr.2026.0017

Original scientific paper.

1. Introduction

Airports are not just transit hubs but also engines of economic growth, connectivity, and tourism development. They facilitate the transport of passengers and goods, stimulating trade, tourism, and investment. According to the report "The Economic and Social Impact of European Airports and Air Connectivity", European airports significantly impact regional economic development and social connectivity [25]. The report's key findings indicate that European airports positively impact the economy, contributing to the employment of approximately 6% of all jobs and generating 5% of European GDP. However, the report also emphasized striking a balance between the economic factors of air transport and its impact on the climate, all to reduce negative environmental impacts. Furthermore, research on the impact of airports on the environment is increasingly important in the context of the global goal of reducing CO emissions.

*Corresponding author.

[10] emphasized the need for sustainable operations, including the introduction of alternative energy sources to reduce the environmental footprint of airports.

Airport efficiency directly impacts the entire air transport system, including flight punctuality, resource utilization, and operating costs. According to [13], airports implementing advanced technologies, such as real-time traffic management systems and digital baggage handling, experience reduced delays and increased passenger satisfaction. Additionally, infrastructure optimization, such as terminal expansion or runway improvements, can significantly increase capacity and reduce congestion. [9] noted that operational efficiency increases airport capacity and enables better integration with other modes of transport, thereby improving the overall mobility of passengers and goods. Airport efficiency concerns not only operational aspects but also economic sustainability as increased productivity can reduce airline operating costs and passenger fees, increasing airport competitiveness.

Research in this area is crucial to improving airport performance while minimizing costs and environmental impact [19, 5]. Data envelopment analysis (DEA) allows accurate measurement of airport service quality and profitability [21, 12, 31]. Furthermore, airports with better operational management make greater economic contributions to the regions in which they operate, enabling greater connectivity, attracting investment, and supporting employment [25]. Given the increasing global demand for air transport and the need for sustainable solutions, research in this area is becoming essential for the future development of the aviation industry. Moreover, the digital transformation and digitalisation of airports globally is a topical and very contemporary subject. The digital transformation process of airports is still predominantly viewed from the aspect of technology application. At the same time, the organizational impact, environment and passenger experience are often neglected and insufficiently researched in the literature [26].

The efficiency of Croatian airports in the domestic literature has received insufficient attention. To the best of the authors' knowledge, the efficiency of Croatian airports via the DEA methodology was analyzed only in [3] for the period 2004-2008 and in [24] for the period 2009-2014. This paper aims to assess and compare the relative efficiency of certified international airports in Croatia via DEA for the more recent period (2019-2023), thereby filling a gap in the domestic literature and providing data-driven guidance for improving airport performance and competitiveness. The contribution of this research lies in presenting a comprehensive and updated assessment of airport efficiency in Croatia for the recent period, filling a significant empirical gap in the national literature. The study highlights current performance trends across airports by applying a dynamic benchmarking approach. It offers valuable, evidence-based guidance for enhancing operational efficiency, resource allocation, and strategic competitiveness in an increasingly demanding and globally interconnected air transport sector. The remainder of the paper is organized as follows: Section 2 reviews relevant literature on airport efficiency and DEA; Section 3 explains the methodology and data; Section 4 presents the results; and Section 5 discusses the findings and concludes the study.

2. Literature review

The aviation industry has flourished in recent years and airport management styles have evolved to apply various benchmarking techniques [20]. As critical infrastructure to the aviation industry, airports are essential for their success and overall performance [31].

Operational efficiency refers to an airport's ability to optimally utilize its resources to ensure a safe, fast, and reliable flow of passengers, cargo, and aircraft. Increasing airport operational efficiency reduces costs, improves the passenger experience and increases airports' competitiveness in the global marketplace. Analysis of the operational efficiency of airports uses different methods to identify opportunities for improvement and optimization of resources. According to [14] and [6], parametric and nonparametric methods are mainly applied, with an emphasis on the nonparametric method of data envelopment analysis (DEA).

DEA is a frequently applied methodology for assessing airport efficiency and can be used to improve airport efficiency [16]. The authors noted that there is no systematic approach to selecting input and output variables, but research depends on available data. A literature review is focused on relevant studies that use the DEA methodology. While DEA forms the core of this review, it is important to note that airports have also been benchmarked using alternative efficiency measurement techniques. Parametric approaches such as stochastic frontier analysis (SFA) have been applied to assess cost and technical efficiency and to compare results with nonparametric models, as shown in [20]. Earlier studies also used Free Disposal Hull (FDH) models alongside DEA, providing a useful comparison between deterministic and nonparametric frontiers. These alternative methods offer complementary insights and underscore the importance of methodological choice when interpreting airport efficiency outcomes. Against this broader methodological background, the present review focuses on DEA-based research, including different DEA model variants, the selection of input and output variables, and policy recommendations for improving airport efficiency proposed in previous empirical studies. Findings from previous studies help to formulate an appropriate DEA framework for assessing airport efficiency in Croatia, identify feasible input-output variables based on data availability, and derive policy recommendations tailored to national conditions.

Empirical studies based on the DEA methodology have been conducted in many countries. The findings of recent and relevant empirical studies emphasizing input and output variable selection are summarized in Table 1 and described in continuation. International studies provide useful benchmarks for Croatia, though their contexts differ. Research on large European hubs [14] highlights the role of user satisfaction, relevant for Zagreb as the main gateway airport. In contrast, analyses of Indian and Colombian regional airports [7, 22] emphasize how smaller airports can achieve high efficiency despite resource constraints, offering insights for Osijek or Rijeka. German airports [27], operating within the EU framework, provide the closest regulatory parallel to Croatia. These comparisons underscore that efficiency does not solely depend on size but on resource alignment, an especially critical factor for Croatia's tourism-driven airports with strong seasonal variation.

The evolution of the DEA methodology in analyzing airport efficiency, with methodological limitations and the complexities of airport operations included, is described in the following papers. [28] proposed a stepwise efficiency improvement DEA model incorporating fixed factors such as runways, addressing the need for objective efficiency analysis of Japanese airports. [29] presented an extended DEA model by combining the distance friction minimization (DFM) method to improve airport efficiency projections for both input reduction and output increase. [30] combined DFM with fixed factor analysis for a few airports in Europe, including the influence of commercial facilities.

By analyzing the operational efficiency of 37 Chinese airports for the period of 2016-2019 via a multi-input/output slack-based DEA model, [8] determined that all 37 airports have very low levels of operational efficiency. On the basis of these results, the authors concluded that airports are not self-sustaining and that more investment should be made in their infrastructure. The bound adjusted measurement (BAM)-DEA model with variable returns to scale [22] is considered the most suitable for measuring the efficiency of Colombian airports. The authors suggest introducing a larger number of variables to gain a more realistic insight into their effectiveness and enhance the competitiveness of the most important variables. The efficiency of Croatian airports was the subject of research by [3] and [24]. [3] reported that only Dubrovnik Airport was relatively efficient in 2008, whereas the average efficiency ratings of almost all Croatian airports increased during the five years analyzed. [24] identified Split, Pula, and Zadar as efficient airports over four years, and Zagreb and Osijek airports over one year. The research results indicate that the relative performance of airports in the analyzed period progressively declined.

Interestingly, even though the application of DEA in airport efficiency is increasing globally,

its application in Croatia is quite modest. These two studies [3, 24] are the only studies in the Croatian scientific bibliography CROSBI. This was the main motivation for this study: to examine and assess the relative efficiency of Croatian airports in recent years. Owing to their vital role in traffic and tourism in Croatia, airport efficiency is extremely important to regulatory bodies, tourism offices, local governments, and scholars.

Author(s)	Period	Variables	Libraries and country	Methodology	DEA model orientation	Results
[7]	-	Inputs: average spending of the airport, the size, the number of parking bays, and the runways; Outputs: average passenger, average revenue, average aircraft movements, and freight	20 airports in India	DEA methodology	Output-oriented DEA model	Increasing returns to scale in 10 airports, decreasing in 2 airports, and constant in 8 airports
[22]	2018, 2019, 2020	Inputs: the length of the runway, the number of workers, the operational costs; Outputs: the number of passengers and tons of cargo.	26 airports in Colombia	Bounded Adjusted Measurement (BAM)-DEA method	Non-oriented DEA model	5 airports were detected as the most effective regional airports; variables runway length and number of employees noticed as the most important.
[14]	2018	Inputs: The number of runways at the airport and the total length of the airport runways; Outputs: the number of passengers	103 European airports	DEA CCR model	Output-oriented DEA model	9 airports detected as perfectly efficient
[15]	1992, 1993	Inputs: Employees, capital costs and other costs; Outputs: terminal passenger numbers and cargo	25 European and 12 Australian airports	DEA methodology and Free Disposal Hull Analysis (FDH)	n.a.	higher levels of efficiency detected in Australian airports
[27]	2016	Inputs: number of employees, number of runways, and airport area; Outputs: number of aircraft movements and the amount of cargo	27 airports in Germany	DEA methodology, CCR, and BCC models	Input-oriented DEA model	13 airports marked as efficient, 5 with an optimal and most productive size

[24]	2009-2014	Inputs: personnel costs, total expenditures excluding personnel costs, and total assets; Output: Total revenue/ATU.	7 airports in Croatia	DEA methodology, CCR and BCC window analysis	Input-oriented DEA model	relative performance of Croatian airports in the analyzed period is progressively declining
[30]	2003	Inputs: terminal space; the number of runways, the number of gates, the number of employees; Outputs: the number of passengers, aircraft movements	19 European airports	Extended DEA model by combining distance friction minimization (DFM)	The extended approach using DFM is non-radial	An extended DEA model is proposed to analyze and compare the efficiency of airports
[18]	2006-2010	Inputs: the airport size, the number of runways, the size of passenger terminals, and the size of cargo terminal; Outputs: the annual number of passengers and the annual cargo volume	20 major airports in East Asia, Europe, and North America	DEA-CCR, DEA-BCC, and DEA-Malmquist production index	Output-oriented DEA model	the most efficient were Gatwick, Zurich, Vienna, Leonardo da Vinci Fiumicino, Los Angeles International, Seattle-Tacoma, San Francisco, Hong Kong, Beijing Capital International, and Shanghai Pudong Airports
[3]	2004-2008	Inputs: expenditures and number of employees; Output: total number of passengers	7 Croatian airports	DEA CCR model, extended DEA method - window analysis	Input-oriented DEA model	The average efficiency ratings of almost all Croatian airports have increased over the observed period

Table 1: *Applications of the DEA methodology in airports' efficiency evaluation.*

Although many studies adopt input-oriented specifications, recent research increasingly supports output-oriented models when revenue generation is the primary performance objective.

3. Methodology and data

There are two basic approaches for assessing efficiency and performance in general, namely the parametric method (stochastic frontier) and the nonparametric method based on linear mathematical programming (which is most commonly the DEA methodology) [1]. The DEA approach is the leading nonparametric approach for the evaluation of airport efficiency, and compared with the stochastic frontier, the literature and the application of the DEA method in the hospitality and tourism sector are "significantly larger" [2].

DEA is one of the most prominent nonparametric methodologies in the field of operations research (OR) and is very popular among academic members, analysts, and consultants looking for ways to evaluate the efficiency of homogeneous units called decision-making units (DMU) [4]. This resulted in its continuous development and application in both theoretical and practical terms, as evidenced by numerous prestigious publications [11, 17].

The choice of DEA, rather than a parametric approach such as stochastic frontier analysis (SFA), is motivated by several methodological considerations. Airport operations involve

complex and heterogeneous production processes for which specifying an explicit functional form is difficult and often restrictive; DEA avoids this by requiring no prior assumptions about the production technology. In addition, the relatively small number of Croatian airports makes nonparametric approaches more suitable, as DEA performs well with small samples where parametric estimations may lack statistical reliability. Although alternatives, such as the Malmquist productivity index, are frequently used for productivity analysis, they rely on stronger assumptions, including stable technology and consistent input-output structures across periods. Given the small number of airports and the structural break caused by COVID-19, Malmquist results would be unstable and potentially misleading. Window DEA is therefore more appropriate, as it avoids intertemporal frontier comparability and remains robust under crisis conditions.

Window DEA is a widely adopted mathematical programming extension that enhances discriminatory power and allows dynamic efficiency assessment in panel data settings [23] and treats each DMU as distinct across reporting periods, enabling dynamic efficiency assessments that better reflect reality than static DEA models do. Subdividing data into overlapping time windows analyzed separately addresses limitations tied to variable-DMU ratios and introduces a time dimension to efficiency tracking [24]. By treating the same airport in different years as distinct units, window analysis increases the number of observations, stabilizes efficiency estimates, and makes it possible to track performance dynamics. This is particularly useful in the Croatian context, where only seven certified airports exist, and where efficiency is strongly influenced by seasonal and cyclical fluctuations.

A standard DEA framework requires a sufficiently large number of decision-making units (DMUs) relative to the number of inputs and outputs to ensure reliable and discriminating efficiency results. One widely accepted condition is that the number of DMUs should satisfy:

$$N \geq 3(m+s) \tag{1}$$

where m is the number of input variables and s is the number of output variables. In the present study, three inputs and one output are used, implying a minimum requirement of $3(3+1) = 12$ DMUs. However, the Croatian airport system consists of only seven airports, which violates this criterion in a conventional DEA setting and would lead to inflated efficiency scores.

To overcome this limitation, a Window DEA framework is employed. Window DEA extends static DEA by converting panel data into overlapping sub-samples ("windows"), treating each DMU in each year as a separate unit. Let $i = 1, \ldots, 7$ index airports, $t = 1, \ldots, 5$ denote years (2019–2023), and $w = 1, \ldots, 3$ represent analysis windows. Each window contains $7 \cdot 3 = 21$ DMUs, thus satisfying the required discrimination threshold:

$$N_w = 21 \geq 12 \tag{2}$$

The adopted windows are: (2019–2021), (2020–2022), and (2021–2023). Within each window, the output-oriented BCC model under variable returns to scale is solved.

For a given airport-year observation j in window w, the BCC model is defined as:

$$max_{\theta,\lambda}\theta \tag{3}$$

subject to:

$$\sum_{j=1}^{N_w} \lambda_j y_{rj} \geq \theta y_{rj_0}, r = 1, \ldots, s \tag{4}$$

$$\sum_{j=1}^{N_w} \lambda_j x_{ij} \leq \ x_{ij_0}, i = 1, \ldots, m \tag{5}$$

$$\sum_{j=1}^{N_w} \lambda_j = 1 \tag{6}$$

$$\lambda_j \geq 0, j = 1, \ldots, N_w \tag{7}$$

where x_{ij} and y_{rj} denote input and output values, respectively, and λ_j are intensity variables constructing the reference frontier. The convexity constraint ensures the variable returns to scale (BCC) specification, while the efficiency score θ represents the proportional expansion in outputs required for the evaluated airport to become efficient, given its current input levels.

In this formulation, θ is the scalar efficiency score of the evaluated airport-year observation (j_0). In the output-oriented model, $\theta \geq 1$ and $\theta = 1$ indicate full efficiency. The λ-weights represent intensity parameters that construct the reference (benchmark) frontier as a convex combination of observed DMUs. The constraint $\sum \lambda_j = 1$ imposes variable returns to scale, ensuring the BCC specification, where e denotes a row vector of ones. The BCC model requires two efficiency conditions: (1) $\theta = 1$ and (2) all constraints satisfied without excess input or output shortfalls.

Implementation Steps:

1. Form three-year moving windows,
2. Treat each airport-year as a DMU,
3. Apply BCC output-oriented DEA per window,
4. Aggregate efficiency results by airport and year,
5. Interpret trends across overlapping windows.

The model is specified under variable returns to scale (VRS), reflecting differences in airport size, capacity, and scale effects, large hubs such as Zagreb operate under fundamentally different conditions than small regional airports such as Osijek or Rijeka. VRS therefore provides a more realistic representation of the Croatian airport system than constant returns to scale. An output-oriented model was chosen because the primary managerial objective of airports is to maximise outputs (such as revenue generation) given largely fixed short-term inputs, especially infrastructure and labour. Other orientations are possible, but output orientation aligns best with both the operational reality of airports and the available dataset.

The main aim of this research is to measure and compare the relative efficiency of international airports in the Republic of Croatia by using the DEA window analysis technique for the period 2019-2023. Although the DEA model specification follows the structure used in [24], our choice of inputs and outputs is not based solely on that study. The broader literature review (shown in Table 1) reveals that financial variables, such as personnel costs, operating expenditures, assets, and revenues are among the most frequently used indicators in airport efficiency studies, particularly when the objective is to measure economic or financial performance (e.g., [3, 24, 7, 22, 8]). Airports differ substantially in size, traffic composition, seasonality, and commercial structure, which makes financial variables suitable since they capture resource utilization and performance in a way that is comparable across airports regardless of their operational mix.

Passenger traffic is indeed commonly used as an output variable in many DEA studies. However, its inclusion in our model was constrained by data consistency requirements. The annual passenger numbers presented in Table 7 originate from the Croatian Bureau of Statistics and represent realised traffic volumes. In contrast, the financial variables used in the DEA model were obtained directly from the annual financial statements of the airport operators. These two datasets are not fully synchronized: some airports classify transit passengers differently, some report combined domestic and international totals while others separate them, and cargo flows are inconsistently reported across years.

More importantly, the financial reporting framework remained consistent across all airports for the full 2019-2023 period, while the traffic dataset contains gaps in monthly and segment-level data for some airports during the COVID-19 years, making it unsuitable for constructing

a uniform operational output across the entire sample. Another important methodological consideration is the ratio of sample size to the number of variables in DEA. With only seven airports in the sample (i.e. DMUs), expanding the model with several additional outputs (e.g., passenger volumes, aircraft movements, or cargo tonnage) would risk breaching established DEA guidelines that recommend maintaining a minimum of three times the total number of inputs and outputs to preserve discriminatory power and prevent an artificially high proportion of units from appearing fully efficient. Although the window analysis framework increases the total number of observations by treating each airport-year combination as a distinct unit, the effective number of DMUs within each window remains relatively small ($n = 21$), which constrains the feasible number of variables that can be included without compromising the robustness of the efficiency estimates.

To maintain discriminatory power and avoid model overspecification, we restricted the model to three inputs and one output. Given these constraints, total revenue was chosen as the output variable because it reflects the aggregate economic outcome of airport operations, encompassing both aeronautical and non-aeronautical activities. This aligns with the financial nature of the selected inputs (costs and assets), creating a coherent and internally consistent economic efficiency model. While passenger numbers could add a complementary operational dimension, the objective of this study is to assess economic efficiency rather than purely operational throughput. Future research may incorporate multiple operational outputs once fully harmonized traffic datasets become available.

Therefore, in this study, three input variables and one output variable are selected as follows: personnel costs, total expenditures excluding personnel costs and total assets as inputs and total revenue as output. The employed variables are presented in Table 2.

Character of variable	Variables
INPUTS	Personnel costs (I1) Total expenditures excluding personnel costs (I2) Total assets (I3)
OUTPUTS	Total revenue (O1)

Table 2: *Inputs and outputs of the DEA model.*

The empirical analysis was conducted using the dedicated window analysis module in MaxDEA 9, which automatically treats each airport-year observation as a separate DMU and ensures consistent window construction. The efficiency results for each airport help identify which Croatian airports are relatively efficient and which are not. Efficient airports are assigned a score of 1, i.e., 100%, whereas inefficient airports are assigned an efficiency score lower than 1.

We use a rolling window of three years, yielding three overlapping windows: 2019-2021, 2020-2022, and 2021-2023 (Table 3). Each airport appears in each window as a separate DMU for each year.

Window 1	2019	2020	2021		
Window 2		2020	2021	2022	
Window 3			2021	2022	2023

Table 3: *Length of the DEA windows used.*

The sample consists of 7 international airports that have been granted certificates in accordance with Commission Regulation no. 139/2014, i.e., Dubrovnik Airport (DU), Pula Airport (PU), Split Airport (ST), Zadar Airport (ZD), Franjo Tuđman Airport Zagreb (ZG), Osijek

Airport (OS), Rijeka Airport (RI). Brač Airport is excluded from the model since it is not an international airport. All the airports in the sample are "state or region/municipality owned with major state ownership" [24].

4. Results

The descriptive statistics offer critical insights into the operational diversity among Croatian international airports from 2019 to 2023 (Table 4). Across the dataset, variables such as personnel costs, other expenditures, total assets, and total revenues reveal wide dispersion and skewed distributions.

	Personnel costs	Total expenditures, excluding personnel costs	Total assets	Total revenue
Mean	5506917	16500335	129861906	24477038
Standard Error	878593	2848683	23355578	4339171
Median	3338275	6314172	24693255	11706093
Standard Deviation	5197826	16853036	138173464	25670879
Kurtosis	-0.3670	-0.7170	-1.1149	-0.8012
Skewness	0.9789	0.7960	0.7274	0.8714
Range	16491892	52485042	389650758	76381796
Minimum	477402	830050	13045695	1342196
Maximum	16969294	53315092	402696453	77723992
Sum	192742093	577511729	4545166709	856696346
Count	35	35	35	35

Table 4: *Descriptive statistics.*

For example, total assets ranged from approximately 13 million EUR to over 402 million EUR, with a standard deviation exceeding 138 million EUR. A similar pattern is evident in total revenue, which varies from just over 1.3 million EUR to more than 77 million EUR, emphasizing the heterogeneity of airport operations. All the variables exhibit positive skewness, with values ranging from 0.73 to 0.98, suggesting that most airports operate below the mean, but a few large airports (like Zagreb and Split) drive the upper end of the distribution. The kurtosis values are all negative, confirming flat, light-tailed distributions.

These observations justify the application of DEA, which remains robust in non-normal, high-variance datasets. Unlike traditional parametric methods, DEA handles this variability without imposing restrictive assumptions, ensuring fair and scale-independent efficiency comparisons.

Outlier detection was conducted over the full five-year period, with the results showing 2019 as an illustrative year (Figure 1). No outliers were identified in any of the years. The lack of extreme values strengthens the validity of the DEA results, as it indicates consistency in airport input-output behavior and confirms the reliability of the sample. It also eliminates concerns of frontier distortion, ensuring a more stable and interpretable efficiency frontier.

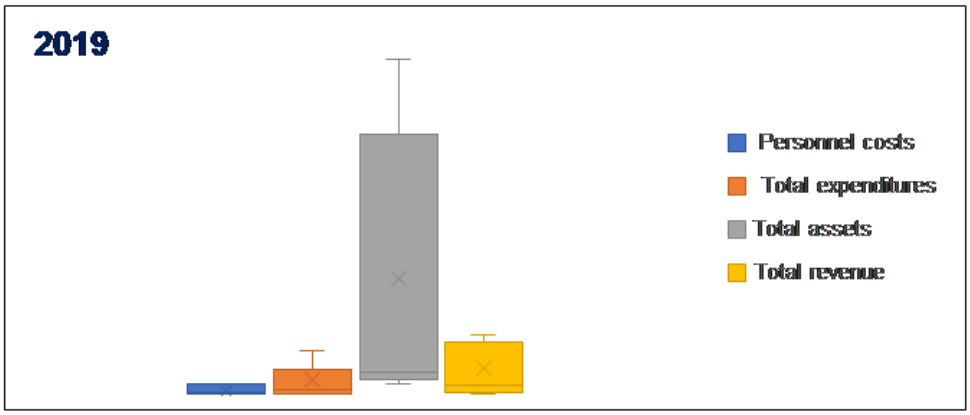


Figure 1: *Outlier diagnostics for input and output variables: evidence from the 2019 dataset.*

The output-oriented BCC DEA model, applied through window analysis with three-year overlapping windows, enabled a dynamic assessment of relative efficiency across the seven certified international airports in Croatia. Each airport was treated as a distinct decision-making unit in each window, allowing the model to capture both performance shifts over time and efficiency volatility. The analysis of the results is supported by efficiency scores across windows (Table 5) and average window DEA efficiency scores for each airport (Table 6), enabling a comprehensive interpretation of performance dynamics over time.

Osijek Airport (OS) demonstrated exceptional consistency and topped the efficiency rankings, with an average window DEA score of 0.9836. It maintained full efficiency in nearly all windows, showcasing highly effective resource management despite being one of the smaller airports in the system. Osijek's performance sets a benchmark for operational excellence and strategic agility in the sector.

Zadar Airport (ZD) ranked second, with an average score of 0.9008, indicating robust efficiency, particularly in the later windows. This indicates that Zadar has successfully adapted its operations over time, possibly benefiting from increased tourism flows and improved infrastructure.

Rijeka (RI) and Zagreb (ZG) followed closely, with average efficiency scores of 0.8627 and 0.8574, respectively. Rijeka achieved an efficiency score of 1.0000 in several later windows, suggesting improved cost control and better output generation. Zagreb, the capital airport and the country's largest airport, performed solidly overall, although not without fluctuations. Given its scale, a slightly lower average is not unexpected, as large airports often deal with more complex operational environments.

Split (ST), once reported as the most efficient airport in earlier studies [24], now ranks fifth, with an average efficiency of 0.7830. Although it reached full efficiency in later windows, its overall score reflects mid-range performance, which is likely affected by inconsistent cost efficiency or underutilization in certain years.

Pula Airport (PU) displayed moderate efficiency, with an average value of 0.6964. The gradual decline over successive windows points to either rising costs or stagnating revenues, which may require strategic investment and re-evaluation of service offerings.

Finally, Dubrovnik Airport (DU) recorded the lowest average score of 0.6695, despite the efficiency score of 1.0000 in the first year. Its persistent inefficiency across subsequent windows is concerning, especially considering Dubrovnik's prominent role in Croatian tourism. This drop might be attributed to seasonal imbalances, high fixed costs, or management challenges, and warrants a comprehensive performance review.

Window	Period	Dubrovnik (DU)	Osijek (OS)	Pula (PU)	Rijeka (RI)	Split (ST)	Zadar (ZD)	Zagreb (ZG)
1-3	1	1	1	0.9716	0.7783	1	1	1
1-3	2	0.3564	1	0.5162	0.5865	0.3577	0.63	0.5132
1-3	3	0.6043	1	0.7664	1	0.7238	1	0.6023
2-4	2	0.3706	1	0.5155	0.5264	0.3972	0.5809	0.7956
2-4	3	0.6033	1	0.7447	1	0.7841	0.9443	0.9357
2-4	4	0.8281	1	0.6795	1	1	1	1
3-5	3	0.6033	1	0.7447	1	0.7841	0.9517	0.9357
3-5	4	0.8145	1	0.6795	1	1	1	1
3-5	5	0.8454	0.8525	0.6497	0.8727	1	1	0.9337

Table 5: *Window DEA analysis score.*

DMU	Average Window DEA Efficiency Score
Osijek (OS)	0.9836
Zadar (ZD)	0.9008
Rijeka (RI)	0.8627
Zagreb (ZG)	0.8574
Split (ST)	0.7830
Pula (PU)	0.6964
Dubrovnik (DU)	0.6695

Table 6: *Average efficiency scores by airport (Window DEA).*

While Table 6 reports average efficiency scores aggregated across all windows, Table 6b presents year-specific averages, thereby restoring the dynamic dimension of airport performance.

Airport	2019	2020	2021	2022	2023
Dubrovnik (DU)	1.0000	0.3635	0.6036	0.8213	0.8454
Osijek (OS)	1.0000	1.0000	1.0000	1.0000	0.8525
Pula (PU)	0.9716	0.5159	0.7519	0.6795	0.6497
Rijeka (RI)	0.7783	0.5565	1.0000	1.0000	0.8727
Split (ST)	1.0000	0.3775	0.7640	1.0000	1.0000
Zadar (ZD)	1.0000	0.6055	0.9652	1.0000	1.0000
Zagreb (ZG)	1.0000	0.6544	0.8246	1.0000	0.9337

Table 6b: *Average efficiency scores by airport and year (Window DEA dynamic view).*

Dynamic efficiency trends become clearer when DEA scores are averaged by year. Table 6b reveals distinct efficiency trajectories across Croatian airports. Osijek displays exceptional consistency with near-perfect efficiency throughout the entire period, confirming superior resource optimization despite low traffic volumes. Zadar shows a strong upward trajectory, reaching full efficiency after 2021, which aligns with its expansion strategy and traffic growth. Rijeka exhibits significant improvement from inefficiency to sustained full efficiency in later years.

In contrast, Dubrovnik shows pronounced volatility. Despite full efficiency in 2019, it experienced a sharp drop during the COVID-19 period and recovered only partially by 2023, suggesting structural rigidity and high fixed costs. Pula similarly exhibits weakening economic performance. Zagreb and Split, while large hubs, display unstable patterns caused by scale complexity and recovery frictions. These year-wise averages confirm that airport efficiency is

dynamic rather than static and strongly influenced by crisis resilience, structural flexibility, and business model adaptation.

A comparative review with earlier studies offers a meaningful perspective. The current ranking diverges from the findings of [24], who identified Split as the top performer from 2009 to 2014. Moreover, the results contradict those of [3], who found Dubrovnik to be among the most efficient between 2004 and 2008. These discrepancies suggest that airport efficiency is not static but rather subject to shifts in policy, investment, demand dynamics, and managerial adaptation.

Window analysis proves invaluable in this context, allowing decision-makers to track performance trajectories rather than static snapshots. For example, Zadar, Rijeka, and Split all show increasing trends, whereas Pula and Dubrovnik reflect stagnation or decline. These patterns can inform resource allocation, strategic planning, and policy interventions at both the airport and national aviation levels.

According to data from the Croatian Bureau of Statistics (DZS), in terms of actual passenger traffic, Croatian airports experienced significant fluctuations during the 2019-2023 period, largely due to the COVID-19 pandemic and the uneven pace of recovery, as shown in Table 7. Zagreb (ZAG), the country's main hub, handled around 3.4 million passengers in 2019, saw traffic fall to only 0.9 million in 2020, and then gradually recovered to over 3.7 million by 2023. Split (SPU) followed a similar pattern, dropping from approximately 3.3 million passengers in 2019 to just 0.6 million in 2020, before rebounding to more than 3.0 million by 2023. Dubrovnik (DBV), heavily dependent on international tourism, was hit particularly hard: from nearly 2.9 million passengers in 2019, traffic collapsed to 0.3 million in 2020 and recovered more slowly compared to Split or Zagreb, reaching only 2.4 million by 2023. Regional airports reflected different dynamics. Zadar (ZD) grew from nearly 800,000 passengers in 2019 to more than 1.2 million in 2023, boosted by the expansion of Ryanair's base. Rijeka (RI), while smaller with around 150,000-200,000 passengers annually, remained stable with moderate growth in seasonal routes. Pula (PU), with approximately 770,000 passengers in 2019, showed a slower recovery to around 600,000 by 2023 due to its reliance on seasonal tourism flows. Finally, Osijek (OS) remained the smallest international airport, handling only 40,000-50,000 passengers annually, yet its DEA efficiency scores suggest that lean operations allowed it to perform relatively well despite limited traffic.

To examine the relationship between operational activity and economic efficiency, passenger traffic trends were compared with DEA efficiency trajectories. A clear association is observed for several airports: Zadar and Rijeka exhibit simultaneous growth in passenger volumes and efficiency scores, suggesting that rising traffic translated effectively into revenue generation and operational scale benefits. Osijek represents a distinctive case of lean efficiency, maintaining high DEA performance despite consistently low passenger numbers, which reflects strict cost control and flexible operations rather than volume-driven efficiency. In contrast, Dubrovnik demonstrates a mismatch between traffic recovery and efficiency outcomes. Although passenger volumes rebounded strongly after 2021, its DEA scores remained persistently low, implying that increased traffic did not proportionally convert into economic performance. This pattern indicates structural cost rigidity, seasonal dependence, and possibly inefficiencies in capacity utilization. Similarly, Pula's stagnant efficiency despite modest traffic recovery reflects vulnerability to seasonal operations and weak revenue diversification. Zagreb and Split, as dominant hubs, display greater efficiency volatility, consistent with scale complexity and delayed cost adjustment during post-pandemic recovery. Although formal statistical correlation is not the primary focus of DEA, the consistency of directional patterns across efficiency scores and traffic volumes supports the robustness of the benchmarking results.

Passenger flows at Croatian airports	2019	2020	2021	2022	2023
Osijek (OS)	45697	6382	10905	15174	36062
Zadar (ZD)	782574	111179	500286	1082672	1213247
Rijeka (RI)	197151	25460	52773	161873	152053
Zagreb (ZG)	3409977	913703	1388736	3102485	3696371
Split (ST)	3271045	659350	1559178	2885863	3336119
Pula (PU)	765508	78832	261647	383477	413439
Dubrovnik (DU)	2877801	322601	917666	2139382	2406674

Table 7: *Passenger flows at Croatian airports (2019-2023).*
Source: *Croatian Bureau of Statistics, TRAFFIC IN AIRPORTS.*

5. Discussion and conclusion

This study benchmarks the efficiency of Croatian international airports in 2019-2023, a period shaped by post-COVID recovery and rapid digital transformation in global aviation. While earlier research evaluated Croatian airports' efficiency from 2004-2014, this paper offers a much-needed update by capturing how airport performance has evolved. The DEA results reveal notable variations across airports, yet efficiency scores alone cannot explain performance gaps. Placing these findings in context highlights structural and managerial factors behind the results.

The applied methodology, incorporating overlapping time windows, allowed for a dynamic understanding of efficiency rather than a static snapshot. The results clearly reveal variations in airport performance. Osijek Airport consistently demonstrated the highest levels of efficiency, indicating that even smaller airports can lead to operational effectiveness when resources are well managed. On the other hand, Dubrovnik Airport, despite its prominent role in Croatian tourism, recorded the lowest average efficiency. This indicates the potential need for deeper strategic and operational adjustments. Zadar, Rijeka, and Split showed increasing trends, reflecting improvements in their efficiency over the observed period. Zagreb, though the largest hub, faces efficiency constraints due to complex operations and infrastructure bottlenecks.

This research offers several important contributions. First, it extensively reviews the global and European literature on airport efficiency, highlighting the strengths of nonparametric methods such as DEA. Second, it fills a significant empirical gap by evaluating the recent performance of Croatian airports, which have not been studied for nearly a decade. Third, and most importantly, the results have direct relevance for a wide range of stakeholders. Regulators, airport operators, and policymakers can use these insights to identify areas for improvement, develop tailored strategies, and inform investment and infrastructure decisions. From a managerial point of view, the study supports data-informed decision-making. Airports can benchmark their performance against that of their national peers and adjust their operations to improve output and cost efficiency. For policymakers, these findings can contribute to broader transport planning, regional development goals, and the design of performance-based support mechanisms.

This study advances the work of [24], in several important ways. While the earlier study focused on a stable pre-pandemic period (2009-2014), our research examines airport efficiency during a period characterized by unprecedented volatility, allowing us to observe how Croatian airports respond to external shocks and recovery trajectories. Moreover, this paper expands the methodological rigor by providing a more comprehensive explanation of the window DEA framework, its assumptions, suitability for small DMU sets, and advantages over static DEA and Malmquist indices under conditions of structural disruption. The study also incorporates a refined financial dataset that enables a cleaner comparison of economic efficiency, avoiding inconsistencies in traffic reporting observed in earlier work. Consequently, the findings contribute new empirical evidence on how Croatian airports adjust their resource utilization during periods

of crisis and thereafter, normalization, offering insights that were not available in the previous literature.

Despite offering valuable insights, this study has certain limitations that should be acknowledged. First, the analysis is limited to seven certified international airports in Croatia, excluding smaller regional or seasonal airports, which may also play important roles in local connectivity and tourism. Second, the choice of input and output variables and the use of only one output variable (total revenue) in the DEA model was guided by data availability and comparability across airports in the 2019-2023 period, which may not capture the full complexity of airport operations, particularly nonfinancial dimensions such as passenger satisfaction or environmental performance. While additional outputs such as passenger traffic or cargo volumes could provide a broader picture of business efficiency, comparable and complete datasets for all airports were not available for the entire period. Previous DEA studies in the aviation sector have demonstrated that financial outputs alone can yield meaningful and robust results when the selected inputs are also financial in nature. Third, the analysis covers the period 2019–2023, which includes the COVID-19 crisis. We recognize that the pandemic caused abnormal distortions in air traffic and efficiency scores. However, the inclusion of this period also represents a strength, as it provides insight into how airports managed efficiency during both crisis and recovery. This perspective is particularly relevant for policymakers and managers seeking to improve resilience against future demand shocks. Third, while window analysis adds a dynamic perspective, it still assumes consistent behavior of decision-making units within each window, which may overlook short-term shocks or rapid operational changes. Finally, the study is based on financial and operational data without integrating qualitative insights from airport management or governance structures, which could provide a deeper context behind efficiency scores. Future research could employ a multi-output DEA framework that incorporates operational measures such as passenger flows or cargo volumes, as well as by conducting sensitivity analyses that exclude the COVID-19 years. Such studies would allow a clearer distinction between structural efficiency patterns in normal conditions and performance under crisis circumstances.

For airport CEOs and policymakers, the results suggest several concrete actions: diversify revenues beyond seasonal passenger flows, expand digital solutions for cost and flow management, and invest in resilience against demand shocks. CFOs should use DEA benchmarking to refine cost efficiency strategies, while operations managers can adopt predictive analytics and automation to optimize resource allocation. Coordinated planning with national tourism authorities can further mitigate seasonal inefficiencies.

By combining DEA insights with traffic and operational realities, this study provides actionable guidance. Croatian airports can enhance competitiveness and sustainability by embracing digital transformation, strengthening cost discipline, and addressing seasonal imbalances, aligning local performance with global trends shaping the future of air transport.

Building on the findings of this study, future research should consider expanding the scope to include a broader set of airports, both within Croatia and across the region, to enable cross-country benchmarking and comparative policy analysis. Integrating parametric approaches such as stochastic frontier analysis (SFA) alongside DEA could offer complementary perspectives and strengthen the robustness of efficiency evaluations. Researchers may also explore multistage models that incorporate environmental and service quality indicators, aligning the assessment with sustainability and customer-centric goals in aviation. Additionally, extending the analysis period to include pre- and post-pandemic data would provide critical insights into how external shocks affect airport efficiency over time. Finally, future studies could incorporate qualitative methods, such as interviews with airport managers and stakeholders, to better understand the strategic and operational factors driving performance outcomes.

References

[1] Arbula Blecich, A. (2024). The performance of Croatian hotel companies–DEA window and Malmquist productivity index approach. Zbornik radova Ekonomskog fakulteta u Rijeci: časopis za ekonomsku teoriju i praksu, 42(1), 9-38. doi: 10.18045/zbefri.2024.1.9

[2] Assaf, A. and Cvelbar, K. L. (2011). Privatization, market competition, international attractiveness, management tenure and hotel performance: Evidence from Slovenia. International Journal of Hospitality Management, 30(2), 391-397. doi: 10.1016/j.ijhm.2010.03.012

[3] Bezić, H., Šegota, A. and Vojvodić, K. (2010). Measuring the efficiency of Croatian airports. In: Trivun, V., Djonlagic, Dz. and Mehic, E. (Eds.) *Economic Development Perspectives of SEE Region in the Global Recession Context – ICES 2010*. Sarajevo: University of Sarajevo, School of Economics and Business, 110–112.

[4] Braimllari, A. and Benga, A. (2019). Cost efficiency of banks in Albania: A Data Envelopment Analysis for the period 2015–2017. In: *Proceedings from the International Conference on Applied Analysis and Mathematical Modelling – ICAAMM19*, 10–13.

[5] Budd, T., Budd, L. and Ison, S. (2015). Environmentally sustainable practices at UK airports. Proceedings of the Institution of Civil Engineers: Transport, 168(2), 116–123. doi: 10.1680/tran.13.00076

[6] Cifuentes-Faura, J. and Faura-Martínez, U. (2021). Twenty Years of Airport Efficiency – A Bibliometric Analysis. Promet - TrafficTransportation, 33(4), 479-490. doi: 10.7307/ptt.v33i4.3790

[7] Chand, G., Mohapatra, D. R. and Jena, P. R. (2024). A DEA Approach to Efficiency Analysis of Major Indian Airports. Vision, 0(0). doi: 10.1177/09722629241255741

[8] Choi, Y., Wen, H., Lee, H. and Yang, H. (2020). Measuring operational performance of major Chinese airports based on SBM-DEA. Sustainability, 12(19), 8234. doi: 10.3390/su12198234

[9] Doganis, R. (2005). *Airline business in the 21st century*. Routledge.

[10] European Union Aviation Safety Agency (EASA) (2005). European aviation environmental report. doi: 10.2822/1537033

[11] Emrouznejad, A. and Yang, G. L. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. Socio-economic planning sciences, 61, 4-8. doi: 10.1016/j.seps.2017.01.008

[12] Fasone, V. and Zapata-Aguirre, S. (2016). Measuring business performance in the airport context: a critical review of literature. International Journal of Productivity and Performance Management, 65(8), 1137-1158. doi: 10.1108/IJPPM-06-2015-0090

[13] Graham, A. (2018). *Managing Airports: An International Perspective*. Routledge.

[14] Henke, I., Esposito, M., della Corte, V., del Gaudio, G. and Pagliara, F. (2021). Airport efficiency analysis in Europe including user satisfaction: A non-parametric analysis with DEA approach. Sustainability, 14(1), 283. doi: 10.3390/su14010283

[15] Holvad, T. and Graham, A. (2000). Efficiency measurement for airports. Proceedings from the Annual Transport Conference at Aalborg University, 7(1). doi: 10.5278/ojs.td.v7i1.4834

[16] Iyer, K. C. and Jain, S. (2019). Performance measurement of airports using data envelopment analysis: A review of methods and findings. Journal of Air Transport Management, 81, 101707. doi: 10.1016/j.jairtraman.2019.101707

[17] Jablonský, J., Emrouznejad, A. and Toloo, M. (2018). Special issue on data envelopment analysis. Central European Journal of Operations Research, 26, 809-812. doi: 10.1007/s10100-018-0584-1

[18] Kim, H-S and Park, J-R. (2013). An analysis of the operational efficiency of the major airports worldwide using DEA and Malmquist productivity indices. Journal of Distribution Science 11(8), 5-14. doi: 10.15722/jds.11.8.201308.5

[19] Lau, C. R., Stromgren, J. T. and Green, D. J. (2010). Airport energy efficiency and cost reduction. Airport Cooperative Research Program (ACRP) Synthesis 21. Transportation Research Board, Washington, DC. url: http://onlinepubs.trb.org/onlinepubs/acrp/acrp_syn_021.pdf [Accessed 19/07/2025]

[20] Matulová, M. and Rejentová, J. (2021). Efficiency of European airports: Parametric versus non-parametric approach. Croatian Operational Research Review, 12(1) 1-14. doi: 10.17535/crorr.2021.0001

[21] Merkert, R. and Assaf, A.G. (2015). Using DEA models to jointly estimate service quality per-

ception and profitability – Evidence from international airports. Transportation Research Part A: Policy and Practice, 75, 42-50. doi: 10.1016/j.tra.2015.03.008

[22] Montoya-Quintero, D. M., Larrea-Serna, O. L. and Jiménez-Builes, J. A. (2022). Evaluation of the efficiency of regional airports using data envelopment analysis. Informatics, 9(4), 90. doi: 10.3390/informatics9040090

[23] Peykani, P., Farzipoor Saen, R., Seyed Esmaeili, F. S. and Gheidar-Kheljani, J. (2021). Window data envelopment analysis approach: A review and bibliometric analysis. Expert Systems, 38(7), e12721. doi: 10.1111/exsy.12721

[24] Rabar, D., Zenzerović, R. and Šajrih, J. (2017). An empirical analysis of airport efficiency: the Croatian case. Croatian Operational Research Review, 8(2), 471–487. doi: 10.17535/crorr.2017.0030

[25] SEO Amsterdam Economics (2024). The Economic and social Impact of European Airports and Air Connectivity. url: aci-europe.org/downloads/resources/SEO [Accessed 19/07/2025]

[26] Spremić, H. (2025). Digitalna transformacija zračnih luka. Ekonomska misao i praksa, 34(2), 599-620. doi: 10.17818/EMIP/2025/30

[27] Stichhauerova, E. and Pelloneova, N. (2019). An Efficiency Assessment of Selected German Airports Using the DEA Model. Journal of Competitiveness, 11(1), 135–151. doi: 10.7441/joc.2019.01.09

[28] Suzuki, S. and Nijkamp, P. (2011). A stepwise efficiency improvement DEA model for airport operations with fixed production factors. European Regional Science Association Conference Paper, 1–18. url: hdl.handle.net/10419/120169 [Accessed 19/07/2025]

[29] Suzuki, S., Nijkamp, P., Rietveld, P. and Pels, E. (2010). A distance friction minimization approach in data envelopment analysis: A comparative study on airport efficiency. European Journal of Operational Research, 207(2), 1104-1115, doi: 10.1016/j.ejor.2010.05.049

[30] Suzuki, S., Nijkamp, P., Pels, E. and Rietveld, P. (2014). Comparative performance analysis of European airports by means of extended data envelopment analysis. Journal of Advanced Transportation, 48, 185-202. doi: 10.1002/atr.204

[31] Yu, C. (2023). Airport Performance-a multifarious review of literature. Journal of the Air Transport Research Society, 1(1), 22-39. doi: 10.59521/E7E8098D7A835864

Croatian Operational Research Review
CRORR **17:2**(2026), 237–252

Multimethod survival analysis for identifying predictors and forecasting mortality in a heart patient cohort study

Syed Wajahat Ali Bokhari [1,*], **and Nasir Ali**[1]

[1] *Department of Statistics, PMAS-Arid Agriculture University, Rawalpindi 46300, Pakistan*
E-mail: ⟨wajahatbokhari2@gmail.com, nasir_stat@uaar.edu.pk⟩*

Abstract. This study presents a multi-method survival analysis of 125 cardiac patients from IIMCT-Pakistan Railway Hospital in Rawalpindi, Pakistan. Parametric accelerated failure-time modeling identified the Weibull distribution as optimal for describing time-to-event data. Semi-parametric analyses, including Cox proportional hazards and Bayesian Cox regression, consistently identified hypertension, ischemic heart disease, and smoking as significant predictors of elevated mortality risk. Higher systolic blood pressure demonstrated a protective effect. Kaplan-Meier analysis revealed steadily declining survival rates up to 300 days with no significant gender differences. The random survival forest model achieved robust predictive accuracy, identifying ischemic heart disease, smoking, and age as the most influential predictors. Our multi-methodological approach demonstrates the value of integrating parametric, semi-parametric, Bayesian, and machine learning techniques for comprehensive risk assessment in cardiac patient cohorts, offering potential for enhanced clinical risk stratification and personalized prognosis.

Keywords: Survival analysis, heart failure, random survival forest, Bayesian Cox regression, personalized risk prediction.

Received: July 11, 2025; accepted: November 26, 2025; available online: February 23, 2026

DOI: 10.17535/crorr.2026.0018

Original scientific paper.

1. Introduction

Cardiovascular diseases are the highest causes of death worldwide accounting for 17.9 million deaths yearly with HF contributing highly to this burden [28]. In spite of the breakthroughs in pharmaceutical and device oriented management of HF, the outcome of the disease continues to be poor with the five-year survival rate of about 50% even now [30]. Identification of high-risk patients at an early stage and precise estimation of survival outcomes imperatively contribute to treatment strategy optimization and success of patients' outcomes [22, 10].

Survival analysis becomes essential in HF research, particularly if outcomes are time-dependent, and the data are right censored. The Kaplan–Meier (KM) estimator is still a frequently applied non-parametric tool to estimate survival functions, and compare groups with log-rank tests [16]. Its ease of interpretation has made it a standard for the clinical studies even though it does not have any ability to adjust for covariates and assumes homogeneity across groups. The variance of KM estimates is usually computed with the Greenwood's formula [9].

To control for the effects of covariates, the semi-parametric nature of the Cox proportional hazards (PH) model makes it widely used, and would provide interpretable hazard ratios without the need to estimate the baseline hazard function [7]. Nevertheless, it assumes proportional

*Corresponding author.

hazards which can be violated in heterogeneous clinical populations. Schoenfeld residuals among other tools help diagnose such violations [25].

When assumption of Proportional Hazards does not hold, parametric survival models like Accelerated Failure Time models (AFT): Weibull, exponential, log-normal, and log-logistic provide the flexibility as they directly model the distribution of survival time [17]The quality of these models can be enhanced to give more accurate estimates if the underlying distribution was well calibrated. Model selection criteria including Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) help identify the most appropriate model by balancing goodness of fit with complexity [5].

Bayesian survival analysis extends these techniques so as to accommodate prior information and give posterior distributions for the model parameters. This is especially beneficial to studies using small sample size or complex hierarchical structure [14, 2]. Bayesian Cox and AFT models have exhibited more flexibility of quantifying uncertainty, than their frequentist opponents.

Recently, machine learning techniques, including random survival forests (RSF) came out as strong methods for survival analysis. While RSF extends Breiman's Random Forests to censored data it does so automatically allowing nonlinear effects and interactions without stringent parametric assumptions [15]. In HF studies, RSF has been shown to provide excellent predictive capabilities and usefulness for discovering new prognostic factors (concordance indices often greater than 0.70) [21].

Considering the complexity of HF and the relevance of correct risk stratification, this study compares a wide array of survival analysis methods, such as KM estimation, Cox and Bayesian Cox models, AFT models and RSF, on a retrospective sample of 125 cardiac patients collected from IIMCT–Pakistan Railway Hospital in Rawalpindi, Pakistan. Our objectives are to: (i) evaluate distributional fit based on AIC and BIC, (ii) determine significant predictors derived from Cox and Bayesian models, (iii) visualize survival differences using KM curves, and (iv) analyze RSF metrics for risk prediction and feature importance. This combined approach is intended to support the choice of strong modeling strategies in individualized HF prognosis.

In Section 2, we describe our materials and methods for the study. We estimate Kaplan–Meier survival curves and carry out log-rank tests of the difference between groups, train a random survival forest (RSF) for prediction and feature importance ranking, and both, the accelerated failure time (AFT) models to determine the Weibull distribution as the most suitable parametric model and Cox regression under the proportional-hazard assumptions. Section 3 provides a methodological synthesis and analysis, specifically, across all methods: reaffirming the prognostic relevance of hypertension, ischemic heart disease, and smoking, checks and balance on model assumptions, and checking RSF's forecasting performance, lastly Section **??** offers general conclusions on the methods' performance, inference stability, and on the potential of individualized risk assessment for use in clinical practice.

2. Materials and methods

2.1. Data collection and ethical considerations

This study was a retrospective cohort design based on the patients' Electronic Health Record data from the Pakistan Railway Hospital, Islamic International Medical College Trust (IIMCT) with the permission from the hospital management. The participants recruited in the study included all those who were admitted to the hospital for cardiac evaluation from January, 2018 to December, 2023, and out of them, 125 patients only were included and followed up completely.

We enlisted research personnel to abstract all the data from the paper and electronic medical records of the patients and input the information in the database. Some of variables included in the study were age, gender, hypertension, ischemic heart disease, diabetes mellitus, anemia, smoking, water-filter use, education level, systolic arterial blood pressure at admission, event

status and, time to event. Patient identifiers were stripped off in consonance with ethical considerations and the patients were assigned unique identification numbers for analysis purposes. Hypertension was categorized as binary (1 = has high blood pressure, 0 = no). In every regression model, the non-hypertensive patients will be the reference category. Therefore, when the coefficient is positive (or HR > 1), the level of risk is increased with hypertension. Systolic blood pressure (SBP) was measured as a continuous variable in millimeters of mercury (mmHg) and all effects reported were in millimeters of mercury per 10 mmHg (mmHg) so that all models could have consistent results.

2.2. Non-parametric methods

2.2.1. Kaplan–Meier estimator

Kaplan–Meier test was used to evaluate the survival function from a time-to-event data format. This non-parametric method takes into consideration of right-censored data and complements the previous one with displaying probability of survival in time.

Kaplan–Meier survival function is defined as

$$\hat{S}(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right), \tag{1}$$

where t_i is the time of the i-th event, d_i is the number of events at t_i, and n_i is the number at risk just prior to t_i.

The Kaplan–Meier survival curve based on the total number of 125 patients is depicted in Figure 1. This curve offers information in estimating the survival probability of the whole population at a given period.

To determine the gender related differences in terms of survival, Kaplan–Meier curves depicted in Figure 2 were constructed for males and females separately. In order to determine statistical differences between these two groups, the log-rank test was applied. The null hypothesis of the log-rank test is that there is no difference in the survival distribution between the male and female patients.

2.2.2. Random survival forest (RSF) model

The random survival forest (RSF) is an extension of the bagged decision tree model proposed by Breiman for right censoring survival data, in that an ensemble of survival trees is constructed to estimate the survival functions for each subject without giving any specific functional form for the hazard or the survival distribution [15]. Since every of the B trees is built with a bootstrap sample of the initial data set, about one third of observations are rejected from the bootstrap (out-of-bag, OOB) and are used to calculate non-biased internal estimates of prediction error. At each node, a random selection of some features out of candidate is performed, and a split that provides the maximum log-rank test statistic that means that the splitting aims at differences in survival rates of daughters, is selected [19]. Only trees are grown to maximal depth so that no pruning is done to eliminate splitting of a variable and the ensembles are used to reduce variability [1].

At each terminal node of tree b, one estimate $\hat{S}_b(t \mid X)$ is computed with the aid of Kaplan–Meier estimator on the OOB samples. The overall RSF survival function for a new covariate vector X is then the simple average

$$\hat{S}(t \mid X) \;=\; \frac{1}{B} \sum_{b=1}^{B} \hat{S}_b(t \mid X), \tag{2}$$

This results in a fully non-parametric estimator of the survival curve that is built by averaging across all the trees in the forest [11].

Model performance is evaluated on the basis of the test data set which is different from out-of-bag (OOB) data. It enables obtaining the survival curves outside the training data, which provides the estimates of the prediction error, including the concordance index, without the need for the test set [29]. To determine variable importance, all the covariates in the OOB is permuted and the increase in the prediction error measured; predictors are ranked according to the extent of contributing to survival discrimination [23].

Key advantages of RSF include:

- No distributional assumptions – the model does not specifically assume any functional form for the hazard and survival functions [15].
- High-dimensional robustness – it can accommodate many high-dimensional correlated covariates by using the randomized splitting and bagging strategies [3].
- *Built-in imputation* – where the missing covariate values are dealt within by using tree-based imputation techniques [25].

Terminology. The results of cox regression are expressed as hazard ratio (HRs), whereas the results of parametric accelerated-failure-time (AFT) model are expressed as acceleration factors (AFs). A positive AF has been known to correspond to increased survival time (protective) and a negative AF corresponded to reduced survival time (increased risk). This difference introduces uniformity between the time-based and hazard-based model interpretations.

2.3. Parametric methods

In survival analysis the method most commonly applied for assessing the impact of different factors on the time to a specific event is the CPH. This method has no links with any specific survival distribution but with a basic assumption that predictor variables affect survival equally at all times. Nevertheless, it may not give a good result if the distribution of the random variable is based on normal distribution. To overcome this weakness, various forms of parametric survival models including Weibull, exponential, log-normal and log-logistic models are used. These models make certain assumptions about the distribution of survival times and are appropriately suited for a range of applications when chosen. The mathematical formulations of these models are as follows:

- *Weibull model:* The Weibull distribution model can fit different shape of hazard function. Its probability density function (PDF) is given by:

$$f(x;\sigma,p) = p \cdot \frac{x^{p-1}}{\sigma} e^{(-x^p/\sigma)}, \quad x > 0,\ \sigma > 0,\ p > 0. \tag{3}$$

  In model, $\sigma > 0$ is known as the scale parameter, which controls the spread of the distribution and $p > 0$ is the shape parameter that governing behaviour of the hazard function. It is particularly useful in modelling experiments where hazards may change with time such as wear-out failures or improvements with time [17].

- *Exponential model:* The exponential distribution is a special case of the Weibull distribution in which $p = 1$. Its PDF is:

$$f(x;\sigma) = \frac{1}{\sigma} e^{-x/\sigma}, \quad x > 0,\ \sigma > 0. \tag{4}$$

  This distribution involves a constant hazard rate, therefore work well in memory less processes including failure processes in systems which have no aging impact [17].

- *Log-normal model:* The log-normal distribution is derived when distribution of logarithm of the survivable life duration is normal. Its PDF is expressed as:

$$f(x;\mu,\sigma^2) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, \quad x \in (0,\infty),\ \mu > 0,\ \sigma > 0. \tag{5}$$

  where the parameters which determine the location and spread of the distribution are termed as mean (μ) and standard deviation (σ), respectively. This model is appropriate when the data is positively skewed as are many survival times that are affected by multiply factors [17].

- *Log-logistic model:* Another distribution from the class of parametric models, useful for skewed data on survival, is the log-logistic distribution. Its PDF is given by:

$$f(x;\alpha,\beta) = \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{[1+(x/\alpha)^{\beta}]^2}, \quad x > 0,\ \alpha > 0,\ \beta > 0. \tag{6}$$

  whereas α stands for scale parameter and β stands for shape parameter. This model is useful when the hazard function rises and then falls over time, and it is good for time-to-event data with fleshy tails [17].

All the parametric models that had been fitted in accelerated failure time (AFT) framework with the coefficients of fit, denoted by β, which act on $logT$. We therefore present the acceleration factors (AFs) in lieu of hazard ratios, in which the acceleration factor (AF) is given as $\exp(\beta\,\Delta x)$. The value of AF below 1 represents a more harmful and the value of AF above 1 represents a more protective. In continuous predictors the effects are expressed as a clinically significant change (e.g. SBP per 10 mmHg): $\mathrm{AF}_{10} = \exp(10\beta)$, and in binary ones the AF is the level "1 = Yes" and "0 = No." In the parameterization of the covariate effects shown below under the AFT, we have covariate effect by describing it as: $\mathrm{AF} = \exp(\beta)$ (or $\exp(\beta\,\Delta x)$).

The selection of these parametric models depends on characteristics of the data and the assumptions made on the form of the hazard function. They offer a large benefit when CPH's assumption of proportional hazards doesn't apply or if a parametric strategy is likely to provide greater estimation accuracy. The use of these models provides more flexible computation and more accurate estimates when used in survival analysis.

2.4. Semi-parametric methods

2.4.1. Cox proportional hazards model

The Cox Proportional Hazards (PH) model was used in order to evaluate effect of more than one variable in hazard rate. The semi-parametric version of the model can define the hazard ratios without estimating the baseline hazard rate:

$$h(t|X) = h_0(t)\exp(\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p) \tag{7}$$

where $h(t|X)$ is the hazard at time t assuming covariates set to be X and $h_0(t)$ is the unknown baseline hazard. The partial likelihood method was used for fitting the models using the Cox PH function from the `coxph()` in R. To evaluate the PH assumption, the Schoenfeld residuals test was done using the function `cox.zph()`.

2.4.2. Bayesian Cox regression analysis

A Bayesian version of the Cox model was also fitted, with the baseline distribution specified as Weibull. This model takes into account the prior information as well as offers the probability of arriving at the hazard ratios. The model specification was:

$$\log h(t|X) = \log\lambda + \gamma\log t + \beta_1 X_1 + \cdots + \beta_p X_p \tag{8}$$

where λ and γ are scale and shape parameter of the Weibull distribution respectively. The model was fitted using the `survregbayes()` function in R that fits a proportional hazards parameterisation and a Weibull baseline. Various convergence diagnostics and posterior summaries were calculated based on the results of the model.

3. Results and discussion

3.1. Results of non-parametric methods

3.1.1. Results of Kaplan–Meier estimator

Figures 1 and 2 display the Kaplan-Meier (K-M) survival curves. Even the overall survival rate of the cohort of 125 patients (Figure 1) reveals that, over a 300 day interval, there is gradual decline in survival. The 95% confidence interval (shaded region) gives an indication of the uncertainties associated with the survival estimate, and this is a function of the number at risk, which curtails as more members at risk at a given time and expands as the size of the sample shrinks over time. Relations of the survival distributions for gender, obtained from Figure 2, show slightly lower survival probabilities for the male patients in comparison to the female throughout the perceptible period of the survival, while the log-rank test has non-significant p-values ($p > 0.05$) and one could not identify any difference in survival between the genders. The number at the risk table that appears beneath each curve depicts, a diminishing number of participants under observation while time elapses.

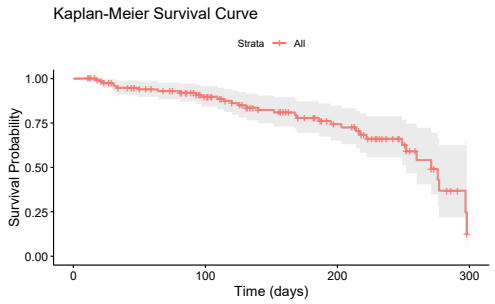


Figure 1: *Kaplan–Meier survival curve for the overall cohort. Time-to-event in days. Y-axis: survival probability; X-axis: follow-up (days).*

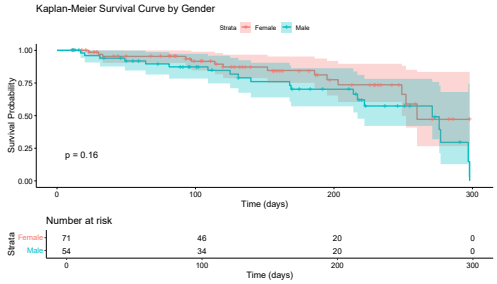


Figure 2: *Kaplan–Meier survival curves by gender. Time-to-event in days. Y-axis: survival probability; X-axis: follow-up (days). Numbers at risk indicated below each curve at each 50-day time point.*

3.1.2. Results of random survival forest (RSF) model

Table 1 shows that RSF was trained on n=125 patients of which 34 had the event of interest. The ensemble contained 1000 trees, each of which grew up to a terminal node size of 15 with the average number of terminal nodes in every tree being 4.691. At each split 4 of the 10 available predictors were chosen at random, and 10 random split points were examined per variable from a log–rank "random" splitting rule. Bootstrap sampling without replacement (SWOR) employed 79 observations per tree, with the OOB error estimated with about one third of the cohort. OOB CRPS (standardized CRPS 0.1315) was 39.1909, corresponding to an OOB prediction error = 0.3312, true OOB prediction error under minimal-depth variable selection 33.1169.

Minimal-depth analysis showed that three top predictors according to minimal-depth values (1.321, 1.728, and 1.752) were, respectively, ischemic heart disease, smoking status, and age, the latter with variable importance (VIMP) reaching –0.009 which in absolute values is low.

Metric	RSF summary	Minimal depth selection
Sample size	125	125
Number of deaths	34	—
Number of trees	1 000	1 000
Forest terminal node size	15	15
Average no. of terminal nodes	4.691	—
No. of variables tried per split	4	4
Total no. of variables	10	10
Resampling used to grow trees	SWOR	—
Resample size used to grow trees	79	—
Analysis	RSF	—
Family	surv	surv
Splitting rule	logrank *random*	logrank
Number of random split points	10	10
(OOB) CRPS	39.1909	—
(OOB) stand. CRPS	0.1315	—
(OOB) Requested performance error	0.3312	33.1169
Conservativeness	—	medium
X-weighting used?	—	TRUE
Dimension (no. of vars)	—	10
Refitted forest	—	FALSE
Model size	—	3
Depth threshold	—	1.7559
Top 3 variables	**Depth / VIMP**	
ischemic_heart_disease	1.321 / 0.244	
smoking	1.728 / 0.102	
age	1.752 / –0.009	

Table 1: *RSF model summary and minimal depth variable selection*

The minimal-depth procedure having the medium conservativeness and X-weighting enabled fitted an unrefitted forest of size 3 with a depth threshold of 1.7559.

The minimal-depth and VIMP measures in the RSF are worth considering, as they are arising at different factors that represent the influence of the predictor. Minimal depth measures the early use of a variable in splitting in the forest which is important to the structure of trees whereas VIMP measures how permuting a variable affects prediction error. The age is very high in our findings in terms of minimal depth, but zero or slightly negative to VIMP. The seeming mismatch suggests that age is a significant contributor to the partitioning of trees (non-linear or interaction effects), but does not significantly enhance the predictive power when other factors (including ischemic heart disease and smoking) are accounted in. The role of age should thus be viewed as being context dependent and secondary to superior clinical predictors.

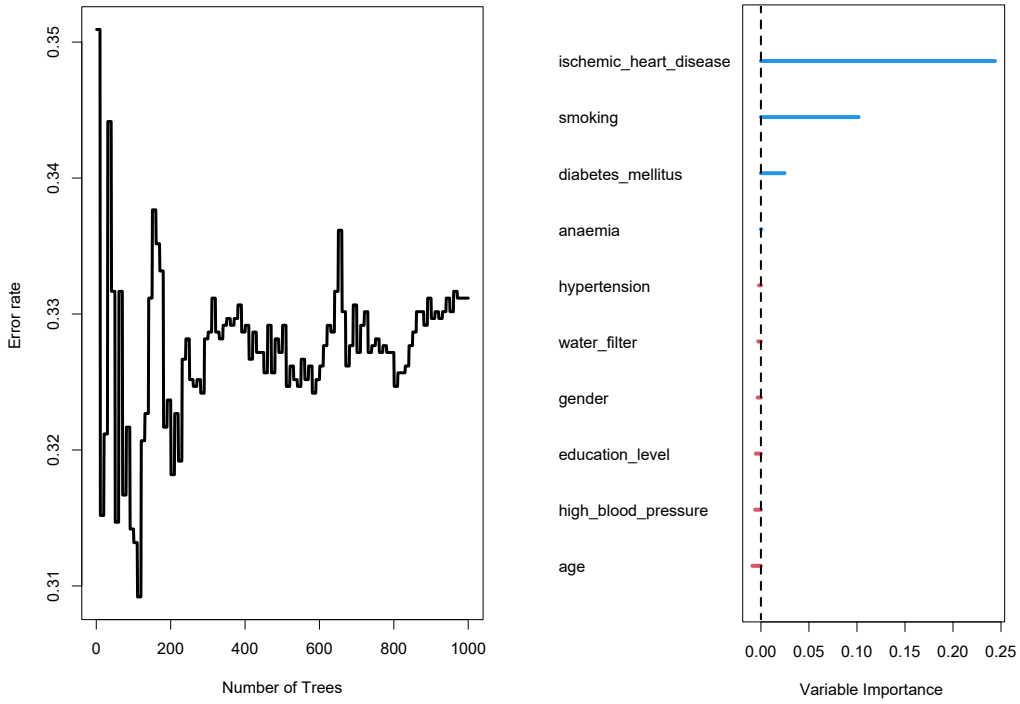
On the whole the RSF had strong performance and interpretability: it used an ensemble of fully mature survival trees for the nonparametric capturing of complex interactions, offered dependable OOB error estimates that do not require separate validation set, and was able to discuss the important covariates through minimal-depth selection.

The variable importance (Table 2) reveals that ischemic heart disease is the best predictor for patient's survival in our random survival forest model (RSF) (VIMP = 0.2437; excluding the station, relative importance = 1.0000), then by smoking status (VIMP = 0.1017; relative importance = 0.4173) and diabetes mellitus (VIMP = 0.0246; relative importance = 0.1008). Other covariates such as anemia have minor positive contributions (VIMP = 0.0006), while hypertension, water filter usage, gender, education level, high blood pressure, and age have

negligible or small negative uses (VIMP range: from –0.0029 to –0.0089), which reflects the weak influence on the model's performance in prediction (Figure 3). The ensemble survival curves (Figure 4) show the risk paths of the entire cohort, as the patient 81 showed the peak value of the predicted survival probability within the follow-up time; other patients' curve is superimposed over all 125 patient estimates and lets one see heterogeneity of the profiles and the ability of the model to generate individual survival estimates.

Variable	Importance	Relative Imp.
ischemic_heart_disease	0.2437	1.0000
smoking	0.1017	0.4173
diabetes_mellitus	0.0246	0.1008
anemia	0.0006	0.0025
hypertension	-0.0020	-0.0083
water_filter	-0.0029	-0.0119
gender	-0.0034	-0.0142
education_level	-0.0053	-0.0218
high_blood_pressure	-0.0060	-0.0245
age	-0.0089	-0.0364

Table 2: *Variable Importance from RSF Model*



Note: Left: OOB error and continuous ranked probability score (CRPS). Right: Minimal depth rankings and variable importance (VIMP). Best predictors: age, ischemic heart disease and smoking.

Figure 3: *Random survival forest (RSF) multimodal results.*

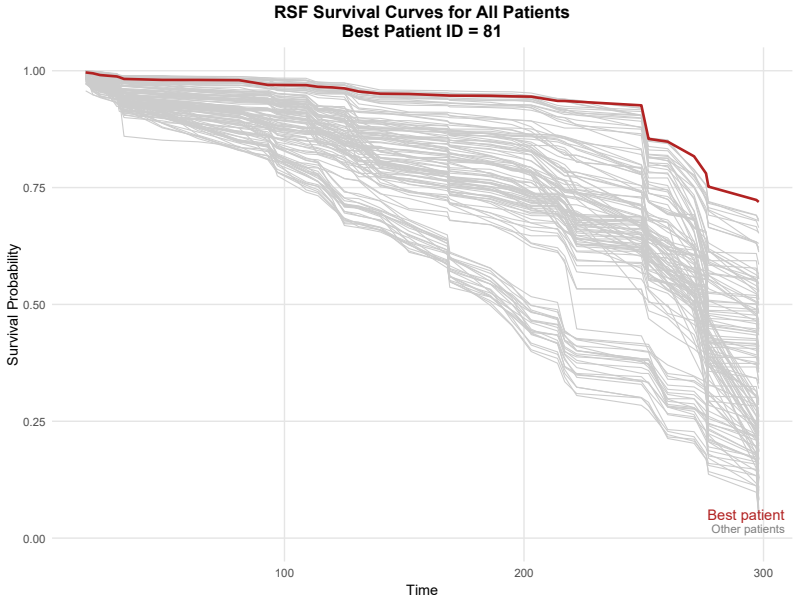


Figure 4: *RSF survival curves for all patients (best patient ID = 81).*

3.2. Results of parametric methods

Table 3 shows maximum-likelihood estimates for four candidate survival distributions: Weibull, Exponential, Log-Normal and Log-Logistic, adjusted for gender, age, hypertension, ischemic heart disease, smoking, diabetes mellitus, anemia, education level, water filter use, and high blood pressure. Across all models, hypertension (Weibull: $\beta = -0.9505$, $p = 0.0152$; Exponential: $\beta = -1.6581$, $p = 0.0122$), ischemic heart disease (Weibull: $\beta = -0.9050$, $p < 0.001$; Exponential: $\beta = -1.5044$, $p < 0.001$), and smoking (Weibull: $\beta = -0.6044$, $p = 0.0054$; Exponential: $\beta = -1.0580$, $p = 0.0067$) emerged as significant predictors of reduced survival time. There is a weak yet noticeable acceleration effect observed with high blood pressure (Weibull: $\beta = 0.0224, p = 0.0193$; Exponential: $\beta = 0.0393, p = 0.0230$) while other covariates such as gender, age, anemia, and education have weaker, non-significant influences in most specifications. It is necessary to note that negative coefficients of hypertension ($\beta < 0$) are associated with a reduced survival time, i.e. an increased hazard of hypertensive patients in the parametric AFT models when transformed into the hazard scale. Therefore, the coefficient sign is negative but the direction of the effect is similar to that of the Cox model where hypertension is a risk factor of mortality. In the case of systolic blood pressure (SBP), the positive coefficient ($\beta = 0.022$ in Weibull AFT model) represents an increment in the predicted survival time by 10 mmHg, and it proves the fact that increasing SBP is related to the decreased risk of mortality. Clinically, the results show that hypertension, ischemic heart disease, and smoking are significant predictors of shorter survival times, and that the patients with higher systolic blood pressure are more likely to have a little longer survival time.

Scale parameters' values differ by distribution (Weibull scale log-estimate -0.6555, $p < 0.0001$; Log-Logistic scale $\hat{\gamma} = 0.463$), whereby different baseline hazard shapes are used. The use of AFT coefficients as acceleration factors (AFs) simplifies the explanation of effects direction and strength of effects on time scale. According to the Weibull AFT model, the hypertension has a $\beta = -0.9505$ meaning the AF $= \exp(-0.9505) =$ **0.387** (the survival time of hypertensive people is about 61% shorter than of non-hypertensive ones).

	Weibull				Exponential				Log Normal				Log Logistic			
Variable	Coef	SE	z	p	Coef	SE	z	p	Coef	SE	z	p	Coef	SE	z	p
(Intercept)	4.136	1.120	3.69	0.0002	3.719	2.114	1.76	0.0785	3.602	1.502	2.40	0.016	3.879	1.394	2.78	0.0054
genderMale	-0.249	0.224	-1.11	0.266	-0.420	0.413	-1.02	0.310	-0.181	0.259	-0.70	0.486	-0.265	0.252	-1.05	0.294
age	0.001	0.006	0.19	0.850	-0.001	0.010	-0.05	0.961	0.004	0.006	0.69	0.490	0.003	0.006	0.51	0.612
hypertension1	-0.950	0.391	-2.43	0.015	-1.658	0.662	-2.51	0.012	-1.018	0.427	-2.38	0.017	-0.910	0.414	-2.20	0.028
ischemic_heart_disease	-0.905	0.250	-3.63	0.0003	-1.504	0.432	-3.48	0.001	-1.093	0.275	-3.98	<0.001	-0.991	0.274	-3.62	0.0003
smoking1	-0.604	0.217	-2.78	0.005	-1.058	0.390	-2.71	0.007	-0.602	0.259	-2.33	0.020	-0.619	0.249	-2.49	0.013
diabetes_mellitus	-0.333	0.230	-1.45	0.147	-0.837	0.405	-2.07	0.039	-0.632	0.263	-2.41	0.016	-0.445	0.251	-1.77	0.077
anemia1	0.124	0.241	0.51	0.609	0.191	0.440	0.43	0.664	0.293	0.275	1.06	0.288	0.225	0.265	0.85	0.394
education_levelPrimary	0.273	0.251	1.09	0.276	0.389	0.476	0.82	0.414	0.519	0.316	1.64	0.100	0.328	0.320	1.03	0.305
education_levelSecondary	-0.103	0.299	-0.34	0.732	-0.210	0.554	-0.38	0.704	0.248	0.371	0.67	0.504	0.106	0.360	0.30	0.767
water_filterNon-filter	0.289	0.211	1.37	0.172	0.309	0.392	0.79	0.430	0.189	0.266	0.71	0.478	0.230	0.250	0.92	0.358
high_blood_pressure	0.022	0.010	2.34	0.019	0.039	0.017	2.27	0.023	0.023	0.013	1.85	0.065	0.022	0.012	1.87	0.062
Log(scale)	-0.656	0.145	-4.53	<0.001	—	—	—	—	-0.123	0.125	-0.98	0.326	-0.770	0.144	-5.35	<0.001
Scale	0.519				1 (fixed)				0.884				0.463			

Note: Binary variables coded as (1 = Yes, 0 = No); reference category = absence of condition. Continuous variables in natural units (e.g., systolic blood pressure in mmHg per 10 mmHg increase). Acceleration factor (AF) = $\exp(\beta\Delta x)$*; AF > 1 indicates longer survival, AF < 1 indicates shorter survival.*

Table 3: *Parameter estimates of parametric AFT models (coefficients on* $\log T$*)*

Model	**Degrees of Freedom (df)**	**AIC**	**BIC**
Weibull	13	**466.5933**	**503.3614**
Exponential	12	480.5786	514.5184
Log-normal	13	476.1975	512.9656
Log-logistic	13	474.4945	511.2626

Table 4: *Model comparison using AIC and BIC for different survival distributions*

Ischemic heart disease has a $\beta = -0.9050$ that gives AF = **0.405** and smoking $\beta = -0.6044$ that gives AF = **0.546**, both reducing survival. In the case of systolic blood pressure (SBP), coefficient is in units per mmHg, the change in survival per unit of 10 mmHg increase in SBP results in the coefficient $AF_{10} = \exp(10 \times 0.0224) = \mathbf{1.251}$, or about 25 % longer survival with a +10 mmHg increase in SBP. These AFs are consistent with the protective direction of SBP in the Cox analysis and also parametric effects on the suitable time scale. Since the AFT model functions on the time scale and the Cox model functions on the hazard scale, the consequences of AFT are interpreted as acceleration factors (AF), whereas the consequences of Cox are interpreted as hazard ratios (HR). The raw coefficients might look inverted across models, but on their scale, both the raw coefficients of the models are the same, i.e. that the interpretation of the raw coefficients of the model using the scale of interpretation are equal. AIC and BIC comparison of models (Table 4) reveals that Weibull model achieves the lowest value of information criteria (AIC = 466.59; BIC = 503.36), is closely followed by the Log-Logistic (AIC = 474.49; BIC = 511.26) suggesting that the Weibull distribution is optimal model in terms of fit and parsimony for these data.

3.3. Results of semi-parametric methods

3.3.1. Results of Cox proportional hazards model

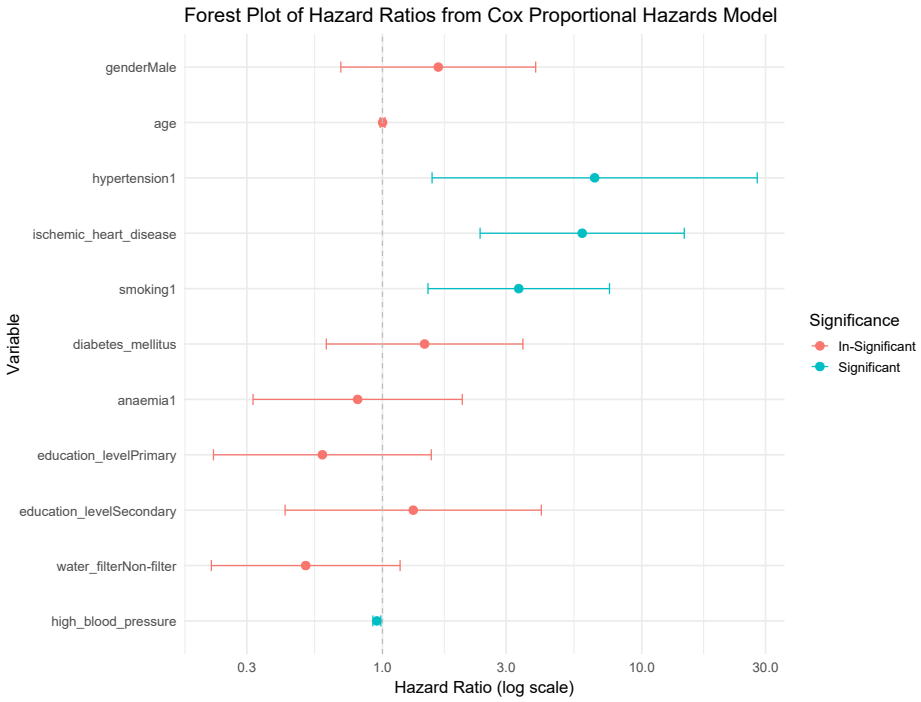
The multivariable Cox PH analysis (Table 5, Figure 5) determined that hypertension (HR = 6.58, 95% CI 1.55–27.87, $p = 0.0105$), ischemic heart disease (HR = 5.89, 95% CI 2.38–14.58, $p < 0.001$), and smoking (HR = 3.35, 95% CI 1.50-7.51, $p = 0.0033$) were significant predictors of mortality. Gender, age, and diabetes mellitus, anemia, education level, and water-filter use did not achieve statistical significance levels since the hazard ratios for these parameters were close to unity with wide confidence intervals crossing 1.0. The concordance, 0.807 (SE = 0.033), indicated excellent discrimination in the model, while the likelihood ratio $p < 0.0001$ reflected a good fit of the model.

Variable	Coef	exp(Coef)	SE(Coef)	z	$Pr(>\|z\|)$	95% CI
genderMale	0.4956	1.6415	0.4413	1.123	0.2614	(0.6912, 3.8983)
age	0.0010	1.0010	0.0108	0.090	0.9283	(0.9800, 1.0223)
hypertension1	1.8842	6.5810	0.7365	2.558	0.0105*	(1.5537, 27.8746)
ischemic_heart_disease	1.7734	5.8907	0.4625	3.834	0.0001***	(2.3795, 14.5829)
smoking1	1.2100	3.3536	0.4112	2.942	0.0033**	(1.4979, 7.5085)
diabetes_mellitus	0.3741	1.4537	0.4453	0.840	0.4009	(0.6073, 3.4798)
anemia1	-0.2203	0.8023	0.4740	-0.465	0.6421	(0.3169, 2.0313)
education_levelPrimary	-0.5335	0.5866	0.4934	-1.081	0.2796	(0.2230, 1.5428)
education_levelSecondary	0.2732	1.3142	0.5804	0.471	0.6379	(0.4213, 4.0993)
water_filterNon-filter	-0.6812	0.5060	0.4275	-1.593	0.1111	(0.2189, 1.1696)
high_blood_pressure	-0.0497	0.9515	0.0180	-2.769	0.0056**	(0.9186, 0.9856)

Note: Binary covariates (1 = Yes, 0 = No); reference category = absence of condition. Continuous variables SBP (mmHg) reported per 10 mmHg increase. Hazard ratio (HR) > 1 indicates higher hazard (shorter survival), HR < 1 indicates lower hazard (longer survival). Concordance = 0.807 (SE = 0.033); Likelihood ratio test = 39.68 on 11 df, p = 4e-05. ***Significance codes:*** ****$p < 0.001$, ** $p < 0.01$, * $p < 0.05$*

Table 5: *Summary of Cox proportional hazards model*

No systematic departures from proportionality were observed for any covariate on Schoenfeld residual plots (Figure 6), thus supporting that the model has valid PH assumption. In all the analyses, hypertensive group (coded as 1) had a consistently high risk over non-hypertensive patients. The differences between the signs of the coefficient in the AFT and Cox frameworks are due to the fact that the parameterizations of the two models are different: the coefficients of AFT are associated with the acceleration of time, whereas Cox coefficients are associated with the increase of hazards. The effect of SBP increase per 10 mmHg was used in the treatment; the negative coefficient (β = -0.0497) and the related HR = 0.95 which show that more SBP



Note: HR > 1 indicates higher hazard (risk); HR < 1 indicates protective effect. Covariates coded as 1 = Yes, 0 = No; SBP expressed per 10 mmHg increase.

Figure 5: *Forest plots of hazard ratios (HRs) and 95% confidence/credible intervals for the Cox model.*

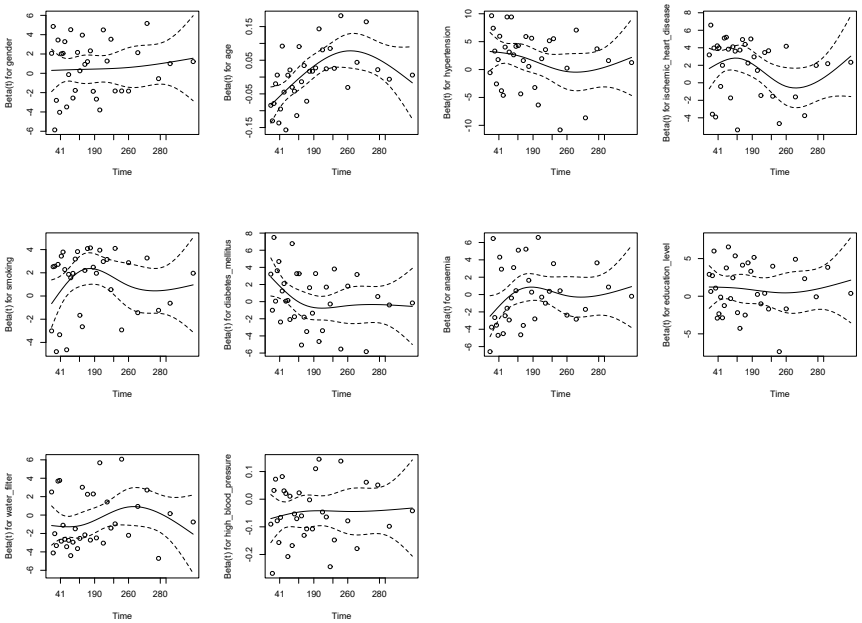


Figure 6: *Schoenfeld residuals for all variables.*

decreased hazard and therefore it was a protective effect. Clinically, these risk ratios indicate that patients with hypertension are at a higher risk of death (more than six times) than patients with normal blood pressure, and that ischemic heart disease and smoking are at a higher risk (three to five times, respectively). Conversely, a 10 mmHg elevation in systolic blood pressure lowers the risk of mortality by approximately 5 % points, which is of a slight protective effect.

It is also important to note that diabetes mellitus exhibited statistically significant effects in certain parametric AFT models (notably the log-normal and log-logistic specifications) but lost significance in the semi-parametric Cox regression framework. These discrepancies stem from the distinct assumptions underlying the two approaches: AFT models directly estimate effects on survival time, while Cox regression models effects on hazard rates and assumes proportional hazards. Because of this, covariates that influence survival time non-proportionally or interact with time may appear significant in one model and not in another. Hence, the diabetes effect should be interpreted cautiously, emphasizing convergence across multiple modeling frameworks rather than any single specification.

3.3.2. Results of Bayesian cox regression model

The Bayesian Cox regression (Table 6, Figure 7) gives estimates of posterior hazard ratio that largely agree with the frequentist Cox results: These include hypertension (posterior mean log-HR = 2.01; 95% CI: 0.62–3.44) and ischemic heart disease (mean = 1.89; 95% CI: 0.94–2.85) as the strongest risk factors, smoking (mean = 1.26; 95% CI: 0.44–2.17) taking second place, and all with credible intervals excluding zero. A consistent protective effect was observed for each 10 mmHg increase in systolic blood pressure (posterior mean log-HR = –0.0476, 95% credible interval: –0.0849 to –0.0097), but other variables, including gender, age, diabetes, anemia, education, and water-filter use have posterior intervals across zero, indicating mixed evidence on the effects of the independent variable based on these parameters. Model fit indices (LPML = –233.46, DIC = 460.94, WAIC = 465.38) indicated adequate predictive performance. The forest plot (Figure 7) displays strong associations for key clinical covariates while appropriately quantifying uncertainty for less influential predictors. Clinically, these Bayesian estimates corroborate the frequentist interpretation: hypertension, ischemic heart disease, and smoking remain key mortality risk factors, whereas higher systolic blood pressure contributes to improved survival probabilities. The case study shows the benefit of integrating classical and machine learning survival models to make risk predictions in underexplored populations. The RSF model enhanced predictive accuracy rates, but it also determined that there were other critical variables like age, which might have a nonlinear influence on mortality risk. Risk factors, as ischemic heart disease and smoking, have even stronger predictive power on mortality in our Pakistani sample, contrary to that typical of Western populations. These results hold significant clinical significance to risk assessment and resource planning in South Asia. Methodologically, the study will reflect the way that RSF can augment customary models through revealing complex associations in the data. The findings of the study have enabled us to make a case in favor of applying machine learning methods to survival analysis to diverse clinical scenarios where data can be scarce or risk characteristics can be dissimilar to historical cohorts.

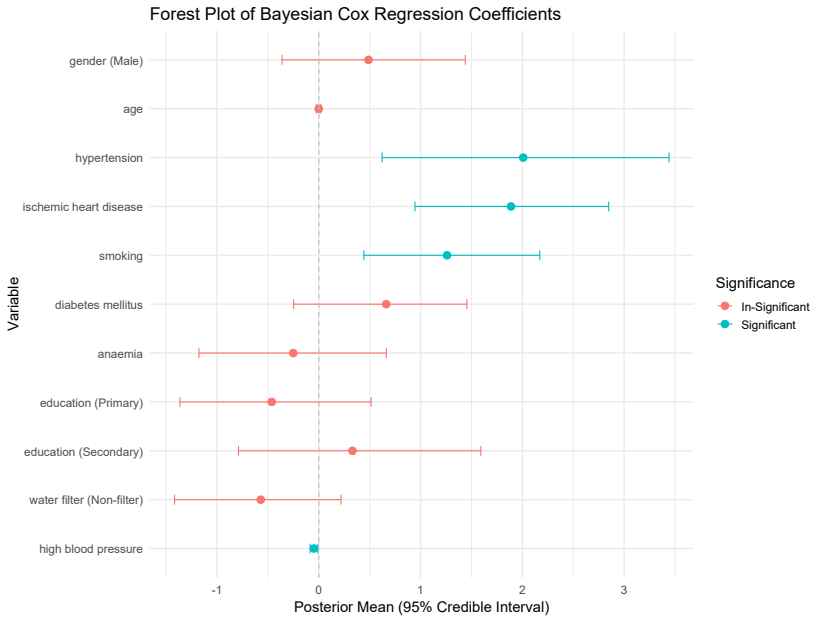
4. Conclusion

This study employed a multi-method survival analysis—integrating parametric, semi-parametric, Bayesian, and machine learning frameworks—to comprehensively identify mortality predictors in a Pakistani cardiac cohort. Our results provide critical insights into risk factors in South Asia and demonstrate the significant potential of combined modeling strategies, particularly random survival forests, for enhancing clinical risk prediction and personalizing patient care in this population. Across all analytical frameworks, hypertension, ischemic heart disease,

Variable	Mean	Median	Std. Dev.	95% CI (Low)	95% CI (Upper)
gender (Male)	0.4881	0.4679	0.4615	-0.3610	1.4397
age	-0.0016	-0.0016	0.0112	-0.0243	0.0209
hypertension	2.0084	2.0253	0.7155	0.6220	3.4407
ischemic heart disease	1.8880	1.8892	0.4831	0.9449	2.8478
smoking	1.2597	1.2606	0.4361	0.4424	2.1710
diabetes mellitus	0.6623	0.6780	0.4331	-0.2473	1.4549
anemia	-0.2513	-0.2576	0.4739	-1.1782	0.6630
education level (Primary)	-0.4629	-0.4591	0.4834	-1.3643	0.5135
education level (Secondary)	0.3305	0.3122	0.5955	-0.7891	1.5917
water filter (Non-filter)	-0.5707	-0.5650	0.4117	-1.4175	0.2179
high blood pressure	-0.0476	-0.0478	0.0192	-0.0849	-0.0097

Note: Binary variables coded as (1 = Yes, 0 = No); reference = absence of condition. Continuous predictors (e.g., SBP) reported per 10 mmHg increase. Posterior means and 95% credible intervals correspond to log-hazard ratios (log-HR). Log Pseudo Marginal Likelihood (LPML) = -233.456, Deviance Information Criterion (DIC) = 460.941, Watanabe-Akaike Information Criterion (WAIC) = 465.378, Number of Subjects = 125

Table 6: *Summary of posterior inference from the Bayesian Cox model*



Note: HR > 1 indicates higher hazard (risk); HR < 1 indicates protective effect. Covariates coded as 1 = Yes, 0 = No; SBP expressed per 10 mmHg increase.

Figure 7: *Forest plots of hazard ratios (HRs) and 95% confidence/credible intervals for the Bayesian Cox model.*

and smoking were consistently significant predictors of mortality. Parametric AFT models confirmed the same direction of effect with acceleration factors indicating shorter survival for these conditions. In contrast, higher systolic blood pressure demonstrated a modest protective effect against mortality.

Semi-parametric Cox proportional hazards and Bayesian Cox regression models identified hypertension, ischemic heart disease, and smoking as the strongest predictors of death, with hazard ratios consistently above unity. An increased systolic blood pressure, reported per 10 mmHg increase, remained a consistent protective factor in both frequentist and Bayesian analyses. Non-parametric Kaplan–Meier analysis revealed steadily declining survival probabilities

over 300 days, with no statistically significant differences between male and female patients. The random survival forest model achieved robust predictive performance and highlighted ischemic heart disease, smoking, and age as the most influential predictors.

Combining these complementary methods provides a comprehensive toolkit for survival analysis: parametric models for direct hazard acceleration inference, Cox and Bayesian Cox regression for robust estimation with minimal assumptions, Kaplan–Meier curves for non-parametric survival estimation, and random survival forests for high-accuracy prediction and variable importance assessment. This multi-methodological approach demonstrates the value of integrating diverse analytical techniques for comprehensive risk assessment in cardiac patient cohorts, offering enhanced potential for clinical risk stratification and personalized prognosis in resource-constrained settings.

Acknowledgments

The authors gratefully acknowledge the institutional support provided by PMAS-Arid Agriculture University, Rawalpindi. We extend our sincere appreciation to **Deputy Superintendent of Police (DSP) Naila Ashraf** of the Pakistan Railways Police for her authoritative support and facilitation during the data collection process. We are particularly grateful to **Dr. Ali Tahir Rana**, Director of Hospitals at Riphah International University and Riphah Healthcare Services, for his significant contributions to data collection. Special thanks are also due to **Ms. Khushboo Khurshid** for her meticulous work in data preparation and cleaning, as well as for her valuable encouragement during this research endeavor. The contributions of all individuals were instrumental in the successful completion of this study.

References

[1] Aalen, O. O., Borgan, Ø., and Gjessing, H. K. (2008). *Survival and Event History Analysis: A Process Point of View*. New York: Springer. doi: 10.1007/978-0-387-68560-1

[2] Ashine, T., Muleta, G., and Tadesse, K. (2021). Assessing survival time of heart failure patients: using Bayesian approach. Journal of Big Data, 8. doi: 10.1186/s40537-021-00537-4

[3] Biau, G., and Scornet, E. (2016). A random forest guided tour. TEST, 25, 197–227. doi: 10.1007/s11749-016-0481-7

[4] Brard, C., Le Teuff, G., Le Deley, M.-C., and Hampson, L. V. (2017). Bayesian survival analysis in clinical trials: What methods are used in practice? Clinical Trials, 14(1), 78–87. doi: 10.1177/1740774516673362

[5] Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.* New York: Springer-Verlag. doi: 10.1007/b97636

[6] Collett, D. (2014). *Modelling Survival Data in Medical Research.* New York: Chapman and Hall/CRC. doi: 10.1201/b18041

[7] Cox, D. R. (1972). Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological), 34(2), 187–220. doi: 10.1111/j.2517-6161.1972.tb00899.x

[8] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis*. New York: Chapman and Hall/CRC. doi: 10.1201/b16018

[9] Greenwood, M. (1931). On the statistical measure of infectiousness. The Journal of Hygiene, 31(3), 336–351. doi: 10.1017/S002217240001086X

[10] Harrell, F. E., Jr. (2015). *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis.* Springer Cham. doi: 10.1007/978-3-319-19425-7

[11] Heagerty, P. J., Lumley, T., and Pepe, M. S. (2021). Random survival forests for dynamic predictions of a time-to-event outcome. BMC Medical Research Methodology, 21. doi: 10.1186/s12874-021-01375-x

[12] Hosseinnataj, A., RezaBaneshi, M., and Bahrampour, A. (2020). Mortality risk factors in patients with gastric cancer using Bayesian and ordinary Lasso logistic models: a study in the

Southeast of Iran. Gastroenterology and Hepatology from Bed to Bench, 13(1), 31–36. url: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7069537/

[13] Hosmer, D. W., Lemeshow, S., and May, S. (2008). *Applied Survival Analysis: Regression Modeling of Time-to-Event Data*. Hoboken, New Jersey: John Wiley & Sons, Inc. doi: 10.1002/9780470258019

[14] Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2001). *Bayesian Survival Analysis*. New York, NY: Springer. doi: 10.1007/978-1-4757-3447-8

[15] Ishwaran, H., and Kogalur, U. B. (2008). Random survival forests for R. R News, 7(2), 25–31. url: https://journal.r-project.org/articles/RN-2007-015/RN-2007-015.pdf

[16] Kaplan, E. L., and Meier, P. (1958). Nonparametric estimation from incomplete observations. Journal of the American Statistical Association, 53(282), 457–481. doi: 10.1080/01621459.1958.10501452

[17] Klein, J. P., and Moeschberger, M. L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*. New York, NY: Springer. doi: 10.1007/b97377

[18] Kleinbaum, D. G. and Klein, M. (2012). *Survival Analysis: A Self-Learning Text, Third Edition* New York, NY: Springer. doi: 10.1007/978-1-4419-6646-9

[19] LeBlanc, M., and Crowley, J. (1992). Relative risk trees for censored survival data. Biometrics, 48(2), 411–425. doi: 10.2307/2532300

[20] Omurlu, I. K., Ozdamar, K., and Ture, M. (2009). Comparison of Bayesian survival analysis and Cox regression analysis in simulated and breast cancer data sets. Expert Systems with Applications, 36(8), 11341–11346. doi: 10.1016/j.eswa.2009.03.058

[21] Penny-Dimri, J. C. , Bergmeir, C., Reid, C. M., Williams-Spence, J., Perry, L. A., and Smith, J. A. (2023). Tree-based survival analysis improves mortality prediction in cardiac surgery. Frontiers in Cardiovascular Medicine, 10, 1211600. doi: 10.3389/fcvm.2023.1211600

[22] Ponikowski, P., Voors, A. A., Anker, S. D., Bueno, H., Cleland, J. G., Coats, A. J., Falk, V., Gonzalez-Juanatey, J. R., Harjola, V. P., Jankowska, E. A., Jessup, M., Linde, C., Nihoyannopoulos, P., Parissis, J. T., Pieske, B., Riley, J. P., Rosano, G. M. C., Ruilope, L. M., Ruschitzka, F., Rutten, F. H., and van der Meer, P. (2016). ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. European Heart Journal, 37(27), 2129–2200. doi: 10.1093/eurheartj/ehw128

[23] Strobl, C., Boulesteix, A.-L., Zeileis, A., and Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformatics, 8. doi: 10.1186/1471-2105-8-25

[24] Therneau, T. M. and Grambsch, P. M. (2000). *Modeling Survival Data: Extending the Cox Model*. New York, NY: Springer. doi: 10.1007/978-1-4757-3294-8

[25] Therneau, T. M. (2020). *A Package for Survival Analysis in R (version 3.2-11)*. url: https://cran.r-project.org/package=survival

[26] Thomas, C., Ye, F.Q., Irfanoglu, M.O., Modi, P., Saleem, K.S., Leopold, D.A., and Pierpaoli, C. (2014). Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. Proc. Natl. Acad. Sci. U.S.A., 111(46), 16574–16579. doi: 10.1073/pnas.1405672111

[27] Van De Schoot, R., Broere, J. J., Perryck, K. H., Zondervan-Zwijnenburg, M., and Van Loey, N. E. (2015). Analyzing small data sets using Bayesian estimation: the case of posttraumatic stress symptoms following mechanical ventilation in burn survivors. European Journal of Psychotraumatology, 6(1). doi: 10.3402/ejpt.v6.25216

[28] World Health Organization. (2021). Cardiovascular diseases (CVDs) – Fact sheet. World Health Organization. url: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[29] Wright, M. N., and Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software, 77(1), 1–17. doi: 10.18637/jss.v077.i01

[30] Yancy, C. W., Jessup, M., Bozkurt, B., et al. (2013). 2013 ACCF/AHA guideline for the management of heart failure. Journal of the American College of Cardiology, 62(16). doi: 10.1016/j.jacc.2013.05.019

Croatian Operational Research Review
CRORR **17:2**(2026), 253–269

Linking dynamic managerial capabilities and organizational agility: The role of technological innovation

Hau Van Nguyen[1,*]

[1] *Faculty of Business Administration, HUTECH University, 475A Dien Bien Phu, Thanh My Tay Ward, Ho Chi Minh City, Vietnam*
E-mail: nv.hau@hutech.edu.vn

Abstract. In an era of rapid technological change and increasing environmental uncertainty, organizational agility has become a key driver of firm competitiveness, underscoring the importance of understanding how dynamic managerial capabilities enable agile organizational responses. The current study investigates the relationship between dynamic managerial capabilities and organizational agility, with a particular focus on the mediating role of technological innovation. It employs a quantitative approach, analyzing survey data from SMEs in Vietnam using partial least squares structural equation modeling (PLS-SEM). The findings reveal that dynamic managerial capabilities positively affect organizational agility, and that technological innovation serves as a mediator in this relationship. Managers with strong dynamic capabilities enable firms to enhance their agility through continuous technological advancements. This study contributes to the literature on dynamic managerial capabilities and innovation by empirically validating the role of managerial capabilities in fostering organizational agility through technological innovation. In practice, it provides valuable insights for business leaders on leveraging managerial capabilities and adopting technology to navigate uncertain market conditions effectively.

Keywords: dynamic managerial capabilities, organizational agility, partial least squares structural equation modeling (PLS-SEM), SMEs, technological innovation

Received: August 15, 2025; accepted: January 21, 2026; available online: February 23, 2026

DOI: 10.17535/crorr.2026.0019

Original scientific paper.

1. Introduction

In the contemporary business landscape, characterized by technological advancements and digitalization, businesses face numerous challenges in maintaining competitiveness and ensuring long-term sustainability, especially small and medium-sized enterprises (SMEs) in developing countries [17]. Organizational agility, the capacity of an organization to swiftly adapt to changes in the external environment, reconfigure its internal structures and operations, and respond effectively to emerging opportunities and threats, plays a pivotal role in sustaining competitive advantage in today's turbulent markets [5]. Organizational agility empowers organizations to quickly detect and respond to technological, economic, and social shifts, ensuring adaptability and resilience [36]. It also supports the pursuit of competitive advantage by allowing firms to capitalize on emerging opportunities, innovate rapidly, and comply effectively with evolving regulations [36]. However, achieving and sustaining such agility requires not only structural flexibility or an advanced information technology infrastructure but also the strategic competencies and decision-making abilities of top managers. Recent research underscores the pivotal role

*Corresponding author.

of strategic capabilities for managers, known as dynamic managerial capabilities, in fostering agility and increasing digital readiness, facilitating innovation in rapidly evolving environments [52].

Dynamic managerial capabilities are defined as the capabilities through which managers create, extend, and modify how firms operate, impacting strategic change and organizational performance [25]. They are an extension of dynamic capabilities, focusing on how managers transform the firm's resources to sustain and enhance its competitive advantage and performance [25]. This capability is a key determinant of organizational agility, as it empowers managers to respond proactively to environmental changes [51, 25]. Managers with excellent dynamic managerial capabilities facilitate strategic decision-making, foster innovation, and effectively allocate resources to enhance agility [1, 25]. Although substantial evidence links dynamic capabilities to organizational agility, the specific contribution of dynamic managerial capabilities remains underexplored, as most studies fail to address the managerial aspect [6]. Further, researchers often combine managerial capabilities with other dynamic capabilities, complicating the isolation of the effects of managerial actions and decisions [52].

Technological innovation refers to the implementation of novel ideas in the form of new products or services, or the incorporation of new elements into an organization's production processes or service operations [11]. Technological innovation mediates the relationship between dynamic managerial capabilities and organizational agility. Managers with better dynamic managerial capabilities are more likely to foster an innovation-oriented culture, mobilize cross-functional teams, and allocate resources to develop new products, services, and processes that meet evolving market demands [24]. In turn, firms that engage in technological innovation are better positioned to anticipate and respond to market changes, thereby enhancing their organizational agility through improved tools and adaptive frameworks [3]. The interplay of dynamic managerial capabilities, technological innovation, and organizational agility is crucial for firms aiming to maintain competitive advantage in rapidly changing environments. There is a need for more empirical studies that explore the specific mechanisms through which technological innovation mediates the relationship between dynamic managerial capabilities and organizational agility. While prior studies have examined the role of information technology capabilities [6], there is limited research on other forms of technological innovation and their impact on organizational agility. A clearer understanding of these relationships is crucial for several reasons. First, it provides insights into how organizations can better leverage dynamic managerial capabilities to foster innovation and agility, an essential advantage in industries marked by rapid technological change and intense competition, where adaptability and innovation are critical to success [18]. Second, it informs strategic decision-making by highlighting the importance of investing in technological capabilities and fostering an innovation-oriented organizational culture [3].

Vietnam provides a helpful context for examining the relationships among dynamic managerial capabilities, technological innovation, and organizational agility. According to The Vietnam Enterprise White Book 2025 [39], while the number of enterprises in Vietnam continues to grow (as of 2024, the number of active enterprises reached 940,069, marking a 2% increase compared to 2023), firms are facing increasing challenges in operational efficiency, profitability, and productivity, which highlights persistent weaknesses in managerial effectiveness and competitiveness. In 2023, SMEs accounted for 97.3% of all enterprises that published financial statements but showed relatively modest financial performance, with 30.2% of medium-sized enterprises (3.7% of the total), 37.5% of small enterprises (26.5%), and 53.6% of micro-sized enterprises (67.1%) reporting financial losses [39]. These patterns indicate limited financial resilience and heavy exposure to environmental uncertainty. In addition, enterprises are highly concentrated in major urban centers such as Ho Chi Minh City (34.9% of the total) and Hanoi (20.9% of the total), while firms in peripheral regions face infrastructure and human capital constraints [39]. Although Vietnam has made substantial progress in digital infrastructure and innovation capacity, these improvements remain unevenly distributed across firms and regions [41]. These

institutional and structural conditions imply that managerial capabilities play a central role in shaping organizational innovation and agility in Vietnamese SMEs, and the empirical findings of this study should be interpreted in light of this specific ecosystem.

The objective of this study is to examine the relationship between dynamic managerial capabilities and organizational agility among SMEs in Vietnam, clarifying the role of technological innovation in this relationship. To achieve this objective, the study adopts a quantitative research design and employs a survey-based methodology using a snowball sampling approach, collecting primary data from senior managers of SMEs across multiple industries in Vietnam. The proposed research model examines the direct and indirect effects of dynamic managerial capabilities on organizational agility, with technological innovation as a mediating mechanism. The hypothesized relationships are empirically tested using the PLS-SEM technique. In doing so, the study contributes to the literature by clarifying how dynamic managerial capabilities are associated with organizational strategic change, particularly organizational agility in volatile environments, and by providing empirical evidence on the mediating role of technological innovation in translating dynamic managerial capabilities into organizational outcomes. Finally, by focusing on SMEs in a developing economy, the study provides contextual evidence to assess the applicability of existing theoretical relationships in settings characterized by higher uncertainty, resource constraints, and managerial discretion.

2. Review of the literature and hypothesis development

2.1. Dynamic managerial capabilities and organizational agility

Adner and Helfat introduced the theory of dynamic managerial capabilities to refine and complement the theory of dynamic capabilities [1]. The authors defined dynamic managerial capabilities as managers' abilities to build, integrate, and reconfigure organizational resources and competences [1]. In broader terms, these capabilities refer to a manager's ability to develop, expand, or adjust the firm's approach to generating value by influencing both internal and external factors [24]. Dynamic managerial capabilities theory emphasizes the managerial influence on strategic change, expanding previous frameworks by explicitly incorporating the role of individual actors, their capabilities, social interactions, and agency in strategic decision-making [27]. Dynamic managerial capabilities, as individual-level capabilities with significant organizational impact, stem from managers' diverse skills, experiences, networks, and mental models, relying on their authority over resources to enable effective action [24]. Although dynamic managerial capabilities are conceptualized at the individual level, this study draws on upper echelons theory [22] and the micro-foundations perspective of dynamic capabilities [51, 15] to explain how individual-level capabilities translate into organizational-level outcomes. Both perspectives emphasize that strategic choices, resource allocations, and organizational change are shaped by the cognition, experiences, and capabilities of top managers. In this sense, managers' dynamic capabilities serve as micro-level mechanisms through which strategic decisions, investment priorities, and organizational routines are formed, thereby influencing firm-level innovation and organizational agility. This link is particularly evident in SMEs, where decision-making authority is typically concentrated at the top and managers tend to exert a more direct, disproportionate influence on organizational outcomes.

The role of dynamic managerial capabilities can be analyzed through the three key capabilities, involving the capacity to (1) sense and shape opportunities and threats, (2) seize opportunities, and (3) transform the resource base [51], enabling managers to guide their organizations through rapidly changing competitive environments. In the light of dynamic managerial capabilities theory, managerial human capital, social capital, and cognition are the three key antecedents of dynamic managerial capabilities [1]. These elements are deeply interconnected, influencing and reinforcing one another to shape a manager's ability to adapt

and respond to a changing business environment [37]. Variations contribute to disparities in outcomes, as the uneven distribution of these antecedents among managers results in differing levels of dynamic managerial capabilities [25]. For instance, dynamic managerial capabilities could promote proactive innovation behaviors, enabling managers to identify emerging technological trends and align internal processes and resources to exploit these opportunities swiftly. This agility is particularly evident when firms face market turbulence or disruptive innovations. Agility refers to the ability to respond to unpredictable and uncertain changes arising from external and internal changes [55]. Organizational agility is a firm-wide capability to swiftly and innovatively respond to unexpected changes in the business environment, leveraging them as opportunities for growth and success [55]. It comprises market capitalization agility and operational adjustment agility, both of which entail a continual openness to change, with the former focusing on an entrepreneurial mindset and the latter emphasizing speedy execution/implementation [35]. In a dynamic environment, organizations survive by adapting to evolving conditions, requiring quick, flexible responses and agility to operate efficiently amid turbulence and uncertainty [8]. Dynamic managerial capabilities shape organizational agility by enabling managers to sense opportunities, seize them effectively, and reconfigure resources to maintain competitive advantage. Managers with strong dynamic managerial capabilities can identify emerging trends, foster collaboration across organizational boundaries, and implement strategic changes that enhance organizational responsiveness and flexibility [25].

Prior studies offer both conceptual and empirical support for a close link between dynamic capabilities and organizational agility. Teece et al. (2016) argued that dynamic capabilities are a prerequisite for achieving the level of organizational agility required to cope with deep uncertainty [50], and several empirical studies have indicated a positive association between the two constructs across different contexts e.g., [60, 23, 30]. Akkaya and Qaisar (2021) further highlighted the strategic relevance of organizational agility, demonstrating that it moderates the relationship between dynamic capabilities and firm performance [2]. Overall, this stream of research suggests that dynamic capabilities are an important antecedent of organizational agility. At the same time, these studies conceptualize dynamic capabilities at the organizational level. They are conducted outside transition-economy contexts, leaving relatively little insight into how dynamic managerial capabilities operate as micro-foundational drivers of organizational agility in SMEs operating in transition economies. Recently, Tenggono et al. (2025) examined the positive relationship between dynamic managerial capabilities and strategic agility in SMEs within a transition economy [53]. However, this study did not incorporate technological innovation as a mediating mechanism in the relationship between dynamic managerial capabilities and organizational agility. In addition, dynamic managerial capabilities were operationalized as a set of resources (managerial human capital, social capital, and cognition), rather than as a set of behaviors reflected in sensing, seizing, and reconfiguring capabilities [53].

Building on this literature, dynamic managerial capabilities can be understood as the manager-level mechanisms through which dynamic capabilities are enacted in organizations. They are particularly relevant in SMEs, where strategic decision-making and resource allocation are typically concentrated among a small number of senior managers. In such settings, managers' abilities to sense opportunities and threats, handle them appropriately, and transform resources are likely to translate directly into how quickly and flexibly the organization can respond to its environment. Therefore, it seems reasonable to expect that firms led by managers with stronger dynamic managerial capabilities will exhibit higher levels of organizational agility. Based on the above discussion, we propose the following hypothesis:

H1: Dynamic managerial capabilities positively impact organizational agility.

2.2. Dynamic managerial capabilities and technological innovation

Technological innovation refers to the implementation of ideas that result in new products or new services, or the introduction of new elements into an organization's production or service processes [11]. It encompasses any modification to an existing product or process, as well as the development of entirely new technologies [24]. Managers play a critical role in technological innovation, particularly those overseeing research and development and new product development, as they significantly contribute to idea generation and may even propose new concepts [22]. Additionally, they influence the outcome through managerial processes and strategic decisions, such as determining funding levels and shaping the composition of the development team, guiding the innovation process [24].

Senior managers also play a pivotal role in shaping technological innovation by scanning the external environment for emerging technologies relevant to their organizations, by investing in specific new technologies, establishing an organizational structure conducive to innovation, such as implementing cross-functional teams or determining the centralization or decentralization of research and development units, or by modifying existing structures to support technological advancement better [24]. Managers with strong dynamic managerial capabilities are more likely to scan the environment for innovative technologies and integrate these insights into strategic decision-making. They play a pivotal role in overcoming internal inertia and fostering a culture that embraces change, thereby enhancing the organization's innovation capabilities [24]. Based on the above arguments, we propose the following hypothesis:

H2: Dynamic managerial capabilities positively impact technological innovation.

2.3. Technological Innovation and Organizational Agility

As mentioned above, technological innovation involves the development and application of new technologies to improve processes, modify existing or new products, and enhance competitive advantage [24]. Some studies indicate that investing in and adopting information technology and digital technologies enhances organizational agility [47, 42, 6]. First, technological innovation enhances agility by enabling real-time data processing and faster decision-making through advanced technologies such as cloud computing, big data analytics, and artificial intelligence, allowing organizations to detect emerging trends and adapt their operations swiftly [47]. Second, these technological advancements enable flexible operational frameworks, making it easier for organizations to reallocate resources and restructure processes in response to market disruptions [42]. Organizations that invest in technology innovation are better equipped to enhance their agility, enabling them to remain competitive in rapidly changing environments. Thus, we propose the following hypothesis:

H3: Technological innovation positively impacts organizational agility.

2.4. The mediating effect of technological innovation

Dynamic managerial capabilities enable managers to adjust the ways an organization earns a living in the present [24]. They focus on transforming how an organization sustains itself by building, integrating, and reconfiguring its resources and competencies [1]. Thus, these capabilities play a pivotal role in shaping an organization's strategic direction [25]. As mentioned above, managers with strong dynamic managerial capabilities will enhance the organization's agility by sensing environmental changes and responding more quickly. Managers with strong dynamic managerial capabilities are also equipped to identify and capitalize on technological opportunities by driving investments in new technologies and fostering an innovation-friendly culture essential for developing and implementing technological innovations [24]. The integration and adoption of technological innovations enhances an organization's ability to process real-time information, streamline decision-making, and rapidly reconfigure operations, thereby

directly increasing organizational agility and enabling a more effective response to environmental changes [6, 47, 42]. Thus, we propose the following hypothesis:

H4: Technological innovation mediates the relationship between dynamic managerial capabilities and organizational agility.

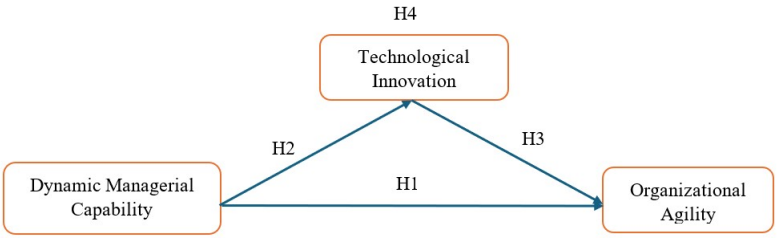


Figure 1: *Research framework.*

3. Methodology and data

3.1. Design

The current study employed two phases. First, a preliminary quantitative study was conducted to ensure that respondents understand the questionnaire, assess the reliability and validity of the measurement items, and identify any potential errors before launching the main study. It consisted of two main steps: a pretest and a pilot test. The pretest study was conducted through in-depth interviews with three academics currently teaching management-related courses and five top managers of SMEs. The pretest aimed to confirm the clarity and relevance of the questions and, if necessary, adjust the terminology for the upcoming survey. The pilot test involved a preliminary survey with 30 senior SME managers to assess the reliability and validity of the measurement items. The main study was also conducted with senior SME managers.

3.2. Sample and Procedure

Data were collected from senior managers of SMEs across diverse industries in Vietnam between December 2024 and May 2025. According to The Vietnam Enterprise White Book 2025 [39], as of the end of 2024, there were 940,078 active enterprises nationwide. In terms of economic sectors, service enterprises accounted for 68.5% of the total, followed by industry and construction at 30.3%, while agriculture, forestry, and fisheries accounted for only 1.3%. Regarding geographical distribution, enterprises located in the Red River Delta accounted for 32.7%, the Northern Midlands and Mountainous Areas for 4.7%, the North Central and Central Coastal region for 13.1%, the Central Highlands for 2.9%, the Southeast region for 39.1%, and the Mekong River Delta for 7.5%. The two cities with the highest concentrations of enterprises are Ho Chi Minh City (after the administrative merger), with 34.9%, and Hanoi, with 20.9%. Together, these two cities represent as much as 55.8% of the total number of enterprises in operation nationwide in 2024. In terms of firm size among operating enterprises with business results as of the end of 2023, micro-sized enterprises accounted for 67.1%, small-sized enterprises 26.5%, medium-sized enterprises 3.7%, and large enterprises 2.7%. As of the end of 2023, SMEs represented 97.3% of all enterprises reporting financial statements. SMEs account for the majority of enterprises in Vietnam's economy and make significant contributions to the country's socio-economic development and its competitive advantage in the global market [38, 57]. Due to the absence of a comprehensive sampling frame and contact lists, this study used the snowball sampling method, which has been widely used in management research [12, 40]. Initially, we contacted

senior managers and provided a clear explanation of the study's scientific purpose, inviting them to participate voluntarily in a self-administered questionnaire with mixed delivery modes. We conducted face-to-face meetings with most survey participants to introduce the study and distribute the questionnaire, but respondents completed the questionnaire themselves. In cases where a face-to-face meeting could not be arranged, we contacted potential respondents by phone to explain the purpose of the study and invite them to participate. Upon obtaining their consent, the questionnaire was sent either in paper or electronic form, and the completed questionnaire was returned later. No personal information, company names, or contact details were collected, and no audio recordings were made. Respondents were also informed that there were no right or wrong answers and that their responses would be treated confidentially and used solely for academic purposes, in aggregated form. These procedures were intended to reduce the risk of socially desirable responses and response bias. After the first round, we asked the senior managers surveyed for referrals or connections to other senior managers. Then we contacted the next senior managers to request a meeting for the survey.

All questions were rated on a 7-point Likert scale from 1 (totally disagree) to 7 (totally agree). To estimate the minimum sample size, we employed the 10-times rule as a heuristic guideline commonly used in research utilizing the partial least squares structural equation modeling approach [33]. According to this rule, the minimum sample size should be determined by the greater of the following two values: (1) ten times the maximum number of formative indicators used to measure a single construct, or (2) ten times the highest number of structural paths directed toward any endogenous latent construct in the model [19]. The construct of organizational agility has the highest number of structural paths pointing toward it, which is two. Therefore, based on the 10-times rule, the minimum sample size for this study is 20. In addition to this heuristic benchmark, we conducted a statistical power analysis [10] using G*Power (with parameters, including alpha = 0.05; power = 0.80; effect size $f^2 = 0.15$; number of predictors = 2). The results indicate that a minimum sample size of 68 is required. With 119 valid responses, the present study exceeds this requirement and therefore has adequate statistical power to detect medium-sized effects.

This study surveyed 119 senior managers, each representing a distinct SME, resulting in one respondent per firm. Among them, 35 were female (29.4%), 83 were male (69.7%), and 1 respondent (0.8%) preferred not to specify their gender. In terms of age, 20.2% were 25–35, 52.1% were 36–45, 25.2% were 46–55, and 2.5% were over 55. Regarding education, 4.2% had a high school education, 66.4% held a university/college degree, 27.2% had a postgraduate degree, and 1.7% fell into other categories. In terms of company size, according to Decree 80/2021/ND-CP of the Government of Vietnam, SMEs are classified based on employee headcount and financial criteria: micro enterprises have fewer than 10 employees and total annual revenue or capital not exceeding 3 billion VND; small enterprises have 10 to fewer than 50 employees and total annual revenue or capital not exceeding 50 billion VND; medium enterprises have 50 to fewer than 200 employees and total annual revenue or capital not exceeding 200 billion VND [56]. Based on this classification, 19.3% are micro enterprises, 66.4% are small enterprises, and 14.3% are medium-sized enterprises. Regarding firm age, 0.8% had been established for less than a year, 26.9% for 1–5 years, 37% for 6–10 years, and 35.3% for more than 10 years. By industry, 18.5% operated in manufacturing, 16% in construction, 56.3% in trade and services, 7.6% in agriculture, forestry, and fisheries, and 1.7% in other sectors.

3.3. Measurement

Three constructs were analyzed in this study: dynamic managerial capabilities, organizational agility, and technological innovation. All of these constructs were derived from existing studies across various contexts worldwide, which may lead to differences in meaning and terminology. Therefore, Schaffer and Riordan's back-translation technique was applied to ensure that the

Vietnamese version of the scale items accurately conveyed the corresponding original meaning [48]. Specifically, the questionnaire was initially prepared in English and then translated into Vietnamese by the author. This procedure was carried out because managers in this market, especially at SMEs, have limited proficiency in English. Another bilingual business administration scholar subsequently translated it back into English. Both English versions of the questionnaire were subsequently compared to verify semantic equivalence. All items in the questionnaire were assessed using a seven-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree).

Dynamic managerial capabilities were measured using a six-item scale adopted from Roth et al. (2023) [46]. Example items include "I systematically identify opportunities from changes in customer needs, new technologies, and the activities of other companies" and "I frequently take the risk of championing investments in new service solutions." Organizational agility was measured by six items adopted from Cegarra-Navarro et al. (2016) [9]. Example items are "We have the ability to rapidly respond to customers' needs" and "We rapidly implement decisions to face market changes." Technological innovation was measured by four items based on Donbesuur et al. (2020) [13]. Example items are "The firm can extend its range of products" and "The firm can replace products that are obsolete."

3.4. Data analysis

The study employed the partial least squares structural equation modeling (PLS-SEM) approach with Smart PLS4 for data analysis. PLS-SEM is well-suited for this study because it is designed for prediction-oriented and theory-building research and allows the simultaneous estimation of measurement and structural models, including direct and indirect relationships [19, 21]. This methodological choice aligns with recent comparative research, which emphasizes that PLS-SEM is especially appropriate for predictive and exploratory research, whereas covariance-based SEM (CB-SEM) is more suitable for purely confirmatory purposes [58]. Moreover, PLS-SEM follows a composite-based logic in which indicators are aggregated to form proxies of abstract constructs [34, 44], which is appropriate for survey-based operationalization of managerial and organizational capabilities and for cross-level designs linking individual-level perceptions to organizational-level outcomes. Finally, PLS-SEM performs well in contexts with relatively small samples and complex models, which is particularly relevant in SME settings [19].

At the same time, we acknowledge that PLS-SEM does not provide the same global fit measures as CB-SEM and is therefore applied here as a method suited to predictive and complex modeling rather than as a general substitute for covariance-based approaches [19, 21]. Following Hair et al.'s (2021) recommendations, both the measurement and structural models were systematically evaluated [20]. The assessment of the measurement model included examinations of internal consistency, convergent validity, and discriminant validity. Then, a bootstrapping procedure with 5,000 subsamples was conducted to test hypotheses in the structural model estimation stage.

3.5. Common method bias

This study used cross-sectional data obtained from a single respondent pool, specifically senior managers of firms, which may include the risk of common method variance. To mitigate common method variance, this study employed multiple techniques [43]. First, the study used enhancement techniques to encourage participants' engagement [14] by incorporating statements into the questionnaire such as: "There is no right or wrong answer; all of your responses are valuable to the study, answering the questionnaire is completely voluntary, and we have no authority to force anyone to participate in the research," and "Your responses will be kept

confidential, only researchers of this project can access the data for the sole purpose of science." Second, the study conducted Harman's one-factor test, and the single factor accounted for less than 50% of the variance [43, 32]. The results indicate that the single-factor model accounted for 37.77% (below the 50% threshold), suggesting that common method variance did not bias the results of this study.

4. Results and discussion

4.1. Measurement model

The internal consistency of the reflective measurement model was evaluated using Cronbach's alpha and composite reliability. To guarantee internal consistency, Hair et al. (2017) have indicated that both Cronbach's alpha and composite reliability should exceed 0.7 but remain below 0.95 to prevent measurement redundancy, as the values above this threshold suggest that all indicators measure the same phenomenon [19]. The results in Table 1 indicate a satisfactory level of internal consistency, with Cronbach's alpha values ranging from 0.717 (dynamic managerial capabilities) to 0.862 (organizational agility). Similarly, composite reliability values vary from 0.824 (dynamic managerial capabilities) to 0.897 (organizational agility).

Next, the study evaluated the convergent validity of all constructs by analyzing outer loadings and average variance extracted (AVE). The outer loadings value indicates the reliability of the indicators and is suggested to exceed the threshold of 0.708 [19]. For values between 0.4 and 0.7, the researchers should consider removing the indicator only if its removal enhances composite reliability or AVE [19]. In the first analysis of convergent validity, the outer loadings and AVEs of constructs such as organizational agility and technological innovation met the required criteria; however, the AVE for the dynamic managerial capabilities construct was only 0.434, which is below the threshold of 0.5 [19]. In the subsequent analysis, after removing two items (DMC2 and DMC6) with outer loadings below 0.708, the AVE for this construct increased to 0.540, surpassing the threshold. In Table 1, although DMC1 had an outer loading below 0.708, it was retained because the AVE of the dynamic managerial capabilities construct had already exceeded 0.5. Table 1 below presents the final results on indicator reliability, showing that the AVEs exceeded the required threshold of 0.5 (dynamic managerial capabilities = 0.54; organizational agility = 0.593; technological innovation = 0.649).

Research constructs	**Items**	**Outer loadings**	**Construct reliability and validity**		
			Cronbach's α	**Composite reliability**	**Average variance extracted (AVE)**
Dynamic managerial capabilities – DMC [46]	DMC1 DMC3 DMC4 DMC5	0.678 0.727 0.731 0.799	0.717	0.824	0.540
Organizational agility – ORA [9]	ORA1 ORA2 ORA3 ORA4 ORA5 ORA6	0.707 0.826 0.749 0.826 0.764 0.740	0.862	0.897	0.593
Technological innovation – TEI [13]	TEI1 TEI2 TEI3 TEI4	0.777 0.864 0.789 0.790	0.821	0.881	0.649

Table 1: *Convergent validity*

Another stage in evaluating the measurement model was assessing discriminant validity. This study employed the Fornell–Larcker criterion [16] and the heterotrait-monotrait ratio [26]. According to the Fornell–Larcker criterion, the square root of the AVE for each construct should be greater than its correlations with other constructs. Additionally, the heterotrait-monotrait ratio should not exceed 0.9 to ensure the discriminant validity. As shown in Table 2, our measurement model met the discriminant validity requirements based on the Fornell–Larcker criterion and heterotrait-monotrait ratio.

Variables	**Fornell–Larcker criterion**			**HTMT**		
	DMC	**ORA**	**TEI**	**DMC**	**ORA**	**TEI**
DMC	**0.735***			–		
ORA	0.661	**0.770***		0.825	–	
TEI	0.526	0.659	**0.806***	0.649	0.759	–

Note: * Square root of AVE value.

Table 2: *Discriminant validity.*

4.2. Structural model

To ensure a reliable evaluation of the structural model, the current study examined collinearity issues between two sets of constructs. Hair et al. (2017) have suggested that the ideal variance inflation factor (VIF) value should be below 5 [19]. Based on this criterion, collinearity was not an issue in this study (Table 3). Next, the study assessed the in-sample predictive power using the R^2 value. According to Hair et al. (2017), R^2 values of 0.75, 0.5, and 0.25 are considered substantial, moderate, and weak, respectively [19]. The R^2 value of organizational agility is 0.595, indicating a moderate effect of predictors (Table 4). We then evaluated the model's out-of-sample predictive power by using the Q^2_{predict} value. To estimate this criterion, the study applied the partial least squares predictive procedure. The Q^2_{predict} value for the organizational agility was 0.417, which is greater than 0, meaning that the model has predictive power [19].

	ORA	**TEI**
DMC	1.382	1.000
TEI	1.382	

Table 3: *VIF values.*

	R^2 **value**	$Q^2_{\textbf{predict}}$ **value**
TEI	0.288	0.256
ORA	0.595	0.417

Table 4: *Model fit indices.*

This study examined the direct and indirect effects of dynamic managerial capabilities on organizational agility using a bootstrapping procedure with 5,000 subsamples to validate the proposed hypotheses (Table 5). Analyses supported H1, indicating that dynamic managerial capabilities directly impact organizational agility ($\beta = 0.435$, $p = 0.000$, LL/UL=0.295/0.577). Additionally, dynamic managerial capabilities directly impact organizational innovation ($\beta = 0.526$, $p = 0.000$, LL/UL=0.347/0.670), supporting H2. Hypothesis H3 is supported, indicating that technological innovation directly impacts organizational agility ($\beta = 0.43$, $t = 5.253$,

LL/UL=0.265/0.587). Finally, technological innovation plays a mediating role in the relationship between dynamic managerial capabilities and organizational agility ($\beta = 0.226$, $p = 0.000$, LL/UL=0.130/0.351).

Path description	**Original sample (O)**	**Sample mean (M)**	**Standard deviation (STDEV)**	**LL/UL**	**t**	**p**	**Hypotheses**
Direct effects							
Dynamic managerial capabilities → Organizational agility	0.435	0.437	0.073	[0.295; 0.577]	5.973	0.000	H1: Supported
Dynamic managerial capabilities → Technological innovation	0.526	0.533	0.076	[0.347; 0.670]	6.912	0.000	H2: Supported
Technological innovation → Organizational agility	0.430	0.432	0.082	[0.265; 0.587]	5.253	0.000	H3: Supported
Mediating effects							
Dynamic managerial capabilities → Technological innovation → Organizational agility	0.226	0.230	0.056	[0.130; 0.351]	4.007	0.000	H4: Supported

Table 5: *Results of the hypotheses.*

4.3. Discussion

The study found a direct, positive impact of dynamic managerial capabilities on organizational agility (H1), demonstrating that managers who can identify opportunities, restructure internally, and adapt strategies flexibly enhance the organization's responsiveness to environmental changes. This finding is consistent with the dynamic capabilities perspective, which conceptualizes organizational agility as an outcome of a firm's abilities to sense, seize, and reconfigure resources and competences in response to environmental change [50]. This study is also consistent with previous research on the impact of dynamic managerial capabilities on firm agility. Specifically, dynamic managerial capabilities have been shown to have a direct and positive effect on strategic agility [53], an indirect and positive impact on supply chain responsiveness [45], and a direct and positive impact on coordination flexibility [29]. However, unlike prior studies that conceptualize dynamic managerial capabilities primarily as resources (e.g., managerial human capital, social capital, cognition) [53, 29], this study adopts a behavioral perspective that focuses on what managers actually do. This perspective provides a more process-oriented explanation of how agility is generated. Moreover, prior studies focused on firms of different sizes rather than exclusively on SMEs and were mainly conducted in developed economies [29, 45]. By contrast, the current study extends this line of research by providing empirical evidence from SMEs in a transition economy, where organizational structures tend to be more flexible, but resources are more constrained.

The study also found that dynamic managerial capabilities have a substantial impact on technological innovation. These results are consistent with the arguments of Helfat and Martin (2015), who emphasize the role of dynamic managerial capabilities in shaping organizational innovation [24]. We extend this work by empirically testing this relationship in the specific context of SMEs and by demonstrating that dynamic managerial capabilities influence not only strategic change (organizational agility) but also concrete technological investment and adoption decisions. This study also aligns with the previous empirical research on the impact of dynamic managerial capabilities on innovation activities in organizations. Specifically, dynamic managerial capabilities have a direct and positive impact on green product innovation [54], SMEs' innovation performance [31], and a digital firm's innovativeness [28]. However, while

these studies focus mainly on general innovation outcomes and dynamic managerial capabilities conceptualized as managerial resources, the present study explicitly examines technological innovation as a distinct and critical organizational process and dynamic managerial capabilities conceptualized as managerial behaviors. This empirical validation is critical in emerging economies, where innovation depends heavily on managerial initiative and discretion [4].

Data analysis from the current study revealed that technological innovation positively affects organizational agility (H3). This finding indicates that adopting and developing new technologies enhances organizational agility, enabling businesses to process information more efficiently. Prior studies have shown that investments in information technology and the adoption of digital technologies enhance organizational agility [42, 6]. However, our findings go beyond these studies by showing that technological innovation does not merely act as an independent driver of organizational agility but also serves as a mediating mechanism linking dynamic managerial capabilities to organizational agility. Specifically, the current study highlights the mediating role of technological innovation in reinforcing the relationship between dynamic managerial capabilities and organizational agility (H4). This finding suggests that dynamic managerial capabilities are not automatically translated into organizational agility; rather, they operate through innovation-related decisions and implementation processes that enable organizations to respond more effectively to environmental change.

As described earlier, the majority of Vietnamese firms can be classified as microenterprises or SMEs (93.6% in 2023). They are financially constrained and operate under conditions of high environmental uncertainty, uneven digital infrastructure, and limited access to skilled human capital [39, 41]. At SMEs, high-level managers and owner-managers often hold centralized decision-making authority and shape organizational culture; thus, their vision, knowledge, and strategic decision-making can shape how proactively and effectively an SME pursues digital initiatives [59]. Top executives must not only support technological investments but also guide their organization through change, fostering new mindsets, capabilities, and agility [49]. This organizational structure strengthens the influence of dynamic managerial capabilities in driving innovation and enhancing organizational agility.

Furthermore, the uneven distribution of digital infrastructure, skills, and innovation capacity across SMEs in Vietnam [41] implies that technological innovation is a critical yet fragile mechanism through which managers' capabilities are translated into organizational outcomes. For many SMEs, innovation is not a routine process but rather a selective, manager-driven investment decision shaped by resource scarcity and risk considerations. Accordingly, the mediating role of innovation reflects both its strategic importance and the practical limitations faced in emerging economies. These results highlight that the effects of dynamic managerial capabilities on organizational agility are likely to be stronger and more visible in environments characterized by institutional uncertainty, resource limitations, and centralized decision-making authority, such as Vietnam [41, 39, 59]

5. Conclusion

5.1. Theoretical contributions

The study makes several theoretical contributions. First, it complements the theory of dynamic managerial capabilities by elucidating the mechanism through which managers' capabilities influence strategic change, specifically the impact on organizational agility in response to environmental volatility. While previous studies have often focused on the underpinnings of dynamic managerial capabilities and their impact on strategic change in volatile environments [29], our study emphasizes how sensing, seizing, and transforming capabilities influence strategic change.

Second, the current study extends the theory of dynamic managerial capabilities by integrating technological innovation as a mediating variable in the relationship between dynamic

managerial capabilities and organizational agility. While many studies have emphasized the direct impact of dynamic managerial capabilities on strategic change and performance outcomes, intermediary mechanisms such as technological innovation have not been thoroughly examined [7]. The current research demonstrates that technological innovation is not merely a byproduct of dynamic managerial capabilities but rather a crucial bridge that transforms managers' capabilities into competitive advantage and organizational agility.

Finally, our study confirms the positive relationship between dynamic managerial capabilities and technological innovation, clarifying a dual mechanism in which leaders not only directly influence the innovation process but also indirectly foster a sustainable innovation environment that enhances organizational agility. While previous studies highlighted the link between dynamic managerial capabilities and innovation [25], no research has thoroughly analyzed both the direct and indirect impacts of these capabilities on technological innovation.

5.2. Practical contributions

From a practical perspective, the research findings show that an organization can improve its agility by developing dynamic managerial capabilities. These capabilities enable businesses to quickly identify opportunities, reconfigure resources, and innovate, but also enhance their ability to respond flexibly to unexpected changes. Thus, managers should proactively build effective dynamic managerial strategies, focusing on improving flexible decision-making, enhancing an adaptive organizational culture, and encouraging continuous innovation. At the same time, businesses should focus on training and developing their workforce to enhance flexible thinking and the ability to cope with change at all levels of the organization. Moreover, the study provides policymakers with a scientific basis for designing support programs for businesses, particularly SMEs, helping them develop dynamic managerial capabilities, increase competitiveness, and achieve sustainable growth in a challenging, risk-prone environment.

Second, the research finding that dynamic managerial capabilities positively affect technological innovation confirms that these capabilities play a key role in helping businesses seize technological opportunities, modify strategies, and optimize resources for effective innovation implementation. Consequently, businesses, including SMEs, should focus on developing dynamic managerial capabilities, including the ability to perceive technological trends, restructure processes, and flexibly adapt to a changing environment. Additionally, the study offers practical insights for managers to formulate appropriate innovation strategies, thereby enhancing competitiveness and achieving sustainable growth. Business support organizations and policymakers can leverage these findings to design training, consulting, and innovation support programs, enabling businesses to maximize their dynamic managerial capabilities and drive technological innovation.

Third, technological innovation makes significant practical contributions to enhancing organizational agility, especially for SMEs, which often have limited resources but must quickly adapt to the market to maintain a competitive edge. It not only helps SMEs optimize their operations and improve efficiency but also enables them to swiftly identify opportunities, reallocate resources, and adjust strategies in response to fluctuations in the business environment. More importantly, technology establishes a foundation for internal connectivity, enhances coordination between departments, automates processes, and strengthens data-driven decision-making, allowing SMEs to respond more flexibly to continuous market changes. Moreover, technological innovation serves as a crucial mediator between dynamic managerial capabilities and organizational flexibility, enabling SMEs to maximize their managerial potential – not only reacting quickly to risks but also proactively driving change. Given their smaller scale, SMEs can leverage technology to bridge the gap with larger enterprises, enhancing competitiveness and expanding the market. This study provides essential practical insights, encouraging SMEs to invest strategically in technological innovation to enhance adaptability, optimize resource use,

and achieve sustainable growth in an increasingly volatile business environment.

5.3. Limitations and directions for future research

Although this study contributes to the literature on dynamic managerial capabilities from both theoretical and practical perspectives, it still has certain unavoidable limitations that should be acknowledged and addressed in future research. First, it is based on a relatively small sample of 119 observations, which may limit the generalizability of the findings. A small sample size can lead to higher margins of error and reduce the reliability of statistical analyses. Future studies should aim to increase the sample size to enhance the representativeness and reliability of the results. Given the lack of comprehensive sampling frames and contact lists, particularly in surveys of senior managers, researchers may consider extending data collection using structured snowball sampling approaches across different industries and regions. Further, future studies using probability-based sampling techniques could enable systematic comparisons between responding and non-responding firms, thereby strengthening the methodological rigor of the findings. A larger and more heterogeneous sample would allow for more robust validation of the findings and improve the generalizability of the conclusions. Second, this study focuses exclusively on SMEs in Vietnam, a developing country in Asia. Because the data are drawn from Vietnamese SMEs, the findings reflect the institutional and economic characteristics of a transitioning economy and may not be directly generalizable to firms in more developed or structurally different contexts. This limitation may affect the generalizability of the findings to other types of businesses or countries with different economic and social conditions. Further studies should consider expanding the scope to include larger enterprises or multinational corporations. Additionally, conducting similar studies in other countries would help assess whether the findings hold across different economic and cultural contexts. Third, the study relies on self-reported survey data, which can introduce biases due to personal subjectivity, inaccuracies, or social desirability effects. Subsequent research could adopt a mixed-methods approach by incorporating observational data, in-depth interviews, or secondary data sources to complement and validate self-reported findings. Fourth, the current study may not have fully accounted for all potential factors influencing organizational agility, such as corporate culture, technological advancements, and leadership styles. Future studies should expand the theoretical framework to incorporate these factors. Finally, this study does not examine the role of moderating variables, such as the competitive environment and innovation level. To gain deeper insight, future research should introduce moderating variables to examine how the relationship between key factors changes across different conditions.

References

[1] Adner, R. and Helfat, C. E. (2003). Corporate effects and dynamic managerial capabilities. Strategic Management Journal, 24(10), 1011-1025. doi: 10.1002/smj.331

[2] Akkaya, B. and Qaısar, I. (2021). Linking dynamic capabilities and market performance of SMEs: The moderating role of organizational agility. Istanbul Business Research, 50(2), 197-214. doi: 10.26650/ibr.2021.50.961237

[3] AlTaweel, I. R. and Al-Hawary, S. I. (2021). The mediating role of innovation capability on the relationship between strategic agility and organizational performance. Sustainability (Switzerland), 13(14), 7564. doi: 10.3390/su13147564

[4] Anand, J., McDermott, G., Mudambi, R. and Narula, R. (2021). Innovation in and from emerging economies: New insights and lessons for international business research. Journal of International Business Studies, 52(4), 545-559. doi: 10.1057/s41267-021-00426-1

[5] Asghar, J., Kanbach, D. K. and Kraus, S. (2025). Toward a multidimensional concept of organizational agility: A systematic literature review. Management Review Quarterly, 1-27. doi: 10.1007/s11301-025-00497-6

[6] Baloch, M. A., Meng, F. and Bari, M. W. (2018). Moderated mediation between IT capability and organizational agility. Human Systems Management, 37(2), 195-206. doi: 10.3233/HSM-17150

[7] Bassam Madi-Odeh, R. and Yousef Obeidat, B. (2024). The moderating role of managerial discretion: The impact of dynamic managerial capabilities on established firms' response strategies to disruptive innovation. International Journal of Innovation Science, 18(1), 130–165. doi: 10.1108/IJIS-11-2023-0258

[8] Budiman, M. A., Soetjipto, B. W., Kusumastuti, R. D. and Wijanto, S. (2023). Organizational agility, dynamic managerial capability, stakeholder management, discretion, and project success: Evidence from the upstream oil and gas sectors. Heliyon, 9(9), e19198. doi: 10.1016/j.heliyon.2023.e19198

[9] Cegarra-Navarro, J.-G., Soto-Acosta, P. and Wensley, A. K. P. (2016). Structured knowledge processes and firm performance: The role of organizational agility. Journal of Business Research, 69(5), 1544-1549. doi: 10.1016/j.jbusres.2015.10.014

[10] Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Mahwah, NJ: Lawrence Erlbaum. doi: 10.4324/9780203771587

[11] Damanpour, F. and Evan, W. M. (1984). Organizational innovation and performance: The problem of "organizational lag." Administrative Science Quarterly, 29(3), 392-409. doi: 10.2307/2393031

[12] Dhar, R. L. (2016). Ethical leadership and its impact on service innovative behavior: The role of LMX and job autonomy. Tourism Management, 57, 139-148. doi: 10.1016/j.tourman.2016.05.011

[13] Donbesuur, F., Ampong, G. O. A., Owusu-Yirenkyi, D. and Chu, I. (2020). Technological innovation, organizational innovation and international performance of SMEs: The moderating role of domestic institutional environment. Technological Forecasting and Social Change, 161, 120252. doi: 10.1016/j.techfore.2020.120252

[14] Eluwole, K. K., Karatepe, O. M. and Avci, T. (2022). Ethical leadership, trust in organization and their impacts on critical hotel employee outcomes. International Journal of Hospitality Management, 102(2022), 103153. doi: 10.1016/j.ijhm.2022.103153

[15] Felin, T., Foss, N. J., Heimeriks, K. H. and Madsen, T. L. (2012). Microfoundations of routines and capabilities: Individuals, processes, and structure. Journal of Management Studies, 49(8), 1351-1374. doi: 10.1111/j.1467-6486.2012.01052.x

[16] Fornell, C. and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18(1), 39-50. doi: 10.2307/3151312

[17] Gamage, S. K. N., Ekanayake, E. M. S., Abeyrathne, G. A. K. N. J., Prasanna, R. P. I. R., Jayasundara, J. M. S. B. and Rajapakshe, P. S. K. (2020). A review of global challenges and survival strategies of small and medium enterprises (SMEs). Economies, 8(4), 79. doi: 10.3390/ECONOMIES8040079

[18] Gyemang, M. D. and Emeagwali, O. L. (2020). The roles of dynamic capabilities, innovation, organizational agility and knowledge management on competitive performance in the telecommunication industry. Management Science Letters, 10(7), 1533-1542. doi: 10.5267/j.msl.2019.12.013

[19] Hair, J. F., Hult, G. T. M., Ringle, C. M. and Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. 2nd ed. Los Angeles: Sage Publications.

[20] Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., Ray, S., et al. (2021). *Partial least squares structural equation modeling (PLS-SEM) using R: a workbook*. Cham, Switzerland: Springer. doi: 10.1007/978-3-030-80519-7

[21] Hair, J. F., Sarstedt, M., Ringle, C. M. and Gudergan, S. P. (2018). *Advanced Issues in Partial Least Squares Structural Equation Modeling*. Los Angeles: Sage Publications.

[22] Hambrick, D. C. and Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. Academy of Management Review, 9(2), 193-206. doi: 10.5465/amr.1984.4277628

[23] Harsch, K. and Festing, M. (2020). Dynamic talent management capabilities and organizational agility - A qualitative exploration. Human Resource Management, 59(1), 43-61. doi: 10.1002/hrm.21972

[24] Helfat, C. E. and Martin, J. A. (2015). Dynamic managerial capabilities: A perspective on the relationship between managers, creativity, and innovation. In Shalley, C.E., Hitt, M.A. and Zhou, J. (Eds.) The Oxford Handbook of Creativity, Innovation, and Entrepreneurship (421-429). New

York: Oxford University Press. doi: 10.1093/oxfordhb/9780199927678.001.0001

[25] Helfat, C. E. and Martin, J. A. (2015). Dynamic managerial capabilities: Review and assessment of managerial impact on strategic change. Journal of Management, 41(5), 1281-1312. doi: 10.1177/0149206314561301

[26] Henseler, J., Ringle, C. M. and Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. Journal of the Academy of Marketing Science, 43, 115-135. doi: 10.1007/s11747-014-0403-8

[27] Heubeck, T. (2023a). Looking back to look forward: a systematic review of and research agenda for dynamic managerial capabilities. Management Review Quarterly, 74, 2243–2287. doi: 10.1007/s11301-023-00359-z

[28] Heubeck, T. (2023b). Managerial capabilities as facilitators of digital transformation? Dynamic managerial capabilities as antecedents to digital business model transformation and firm performance. Digital Business, 3(1), 100053. doi: 10.1016/j.digbus.2023.100053

[29] Hock-Doepgen, M., Heaton, S., Clauss, T. and Block, J. (2024). Identifying microfoundations of dynamic managerial capabilities for business model innovation. Strategic Management Journal, 46(1), 470-501. doi: 10.1002/smj.3663

[30] Kanten, P., Kanten, S., Keceli, M. and Zaimoglu, Z. (2017). The antecedents of organizational agility: organizational structure, dynamic capabilities and customer orientation. PressAcademia Procedia, 3(1), 697-706. doi: 10.17261/Pressacademia.2017.646

[31] Khan, K. U., Atlas, F., Ghani, U., Akhtar, S. and Khan, F. (2021). Impact of intangible resources (dominant logic) on SMEs innovation performance, the mediating role of dynamic managerial capabilities: Evidence from China. European Journal of Innovation Management, 24(5), 1679-1699. doi: 10.1108/EJIM-07-2020-0276

[32] Kock, N. (2021). Harman's single factor test in PLS-SEM: Checking for common method bias. Data Analysis Perspectives Journal, 2(2), 1-6.

[33] Kock, N. and Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. Information Systems Journal, 28(1), 227-261. doi: 10.1111/isj.12131

[34] Lohmöller, J.-B. (2013). *Latent Variable Path Modeling with Partial Least Squares*. Heidelberg: Springer Science & Business Media. doi: 10.1007/978-3-642-52512-4

[35] Lu, Y. and Ramamurthy, K. (2011). Understanding the link between information technology capability and organizational agility: An empirical examination. MIS Quarterly, 35(4), 931-954. doi: 10.2307/41409967

[36] Marhraoui, M.A. and El Manouar, A. (2020). Organizational Agility and the Complementary Enabling Role of IT and Human Resources: Proposition of a New Framework. In Baghdadi, Y., Harfouche, A., Musso, M. (Eds.) ICT for an Inclusive World. Lecture Notes in Information Systems and Organisation, vol 35. Cham: Springer. doi: 10.1007/978-3-030-34269-2_4

[37] Martin, J. A. (2011). Dynamic managerial capabilities and the multibusiness team: The role of episodic teams in executive leadership groups. Organization Science, 22(1), 118-140. doi: 10.1287/orsc.1090.0515

[38] National Statistics Office. (2025). The Prime Minister instructs the promotion of the development of small and medium-sized enterprises. url: https://www.nso.gov.vn/tin-tuc-khac/2025/03/thu-tuong-chi-thi-thuc-day-phat-trien-doanh-nghiep-nho-va-vua/ [Accessed 26/7/2025]

[39] National Statistics Office. (2025). *Vietnam Enterprise White Book Year 2025*. Hanoi: Statistical Publishing House. url: https://www.nso.gov.vn/wp-content/uploads/2025/11/Sach-trang-DN-2025-PDF.pdf [Accessed 26/7/2025]

[40] Nguyen, H. V. and Nguyen, L. L. H. (2024). Linking ethical leadership to employees' prohibitive voice: The role of reflective moral attentiveness and leader identification. International Journal of Ethics and Systems, 42(1). doi: 10.1108/IJOES-11-2023-0252

[41] Open Development Vietnam. (2023). Vietnam Digital Transformation Agenda. url: https://vietnam.opendevelopmentmekong.net/topics/vietnam-digital-transformation-agenda/ [Accessed 12/4/2025]

[42] Overby, E., Bharadwaj, A. and Sambamurthy, V. (2006). Enterprise agility and the enabling role of information technology. European Journal of Information Systems, 15(2), 120-131. doi: 10.1057/palgrave.ejis.3000600

[43] Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y. and Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. Journal of Applied Psychology, 88(5), 879. doi: 10.1037/0021-9010.88.5.879

[44] Rigdon, E. E. (2016). Choosing PLS path modeling as an analytical method in European management research: A realist perspective. European Management Journal, 34(6), 598-605. doi: 10.1016/j.emj.2016.05.006

[45] Roh, J., Swink, M. and Kovach, J. (2022). Linking organization design to supply chain responsiveness: The role of dynamic managerial capabilities. International Journal of Operations & Production Management, 42(6), 826-851. doi: 10.1108/IJOPM-08-2021-0526

[46] Roth, K., Rau, C. and Neyer, A. K. (2023). Design thinking and dynamic managerial capabilities: A quasi-experimental field study in the aviation industry. R&D Management, 53(5), 801-818. doi: 10.1111/radm.12600

[47] Sambamurthy, V., Bharadwaj, A. and Grover, V. (2003). Shaping agility through digital options: Reconceptualizing the role of information technology in contemporary firms. MIS Quarterly, 27(2), 237-263. doi: 10.2307/30036530

[48] Schaffer, B. S. and Riordan, C. M. (2003). A review of cross-cultural methodologies for organizational research: A best-practices approach. Organizational Research Methods, 6(2), 169-215. doi: 10.1177/1094428103251542

[49] Schrage, M., Pring, B., Kiron, D. and Dickerson, D. (2021). Leadership's digital transformation: Leading purposefully in an era of context collapse. https://sloanreview.mit.edu/projects/leaderships-digital-transformation/ [Accessed 16/2/2025]

[50] Teece, D., Peteraf, M. and Leih, S. (2016). Dynamic capabilities and organizational agility: Risk, uncertainty, and strategy in the innovation economy. California Management Review, 58(4), 13-35. doi: 10.1525/cmr.2016.58.4.13

[51] Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. Strategic Management Journal, 28(13), 1319-1350. doi: 10.1002/smj.640

[52] Tenggono, E., Soetjipto, B. W. and Sudhartio, L. (2024). Navigating institutional pressure: Role of dynamic managerial capabilities and strategic agility in healthcare organizations' renewal. International Journal of Healthcare Management, 18(3), 502-511. doi: 10.1080/20479700.2024.2323846

[53] Tenggono, E., Soetjipto, B. W. and Sudhartio, L. (2025). Dynamic managerial capabilities in action: advancing strategic agility and digital readiness in healthcare organizations. Journal of Organizational Change Management, 38(3), 548-568. doi: 10.1108/JOCM-12-2024-0772

[54] Tran, M. L. and Vo Thai, H. C. (2025). Future-ready strategies: Dynamic managerial capabilities, digitalization, and green product innovation in building firm resilience. Business Ethics, the Environment & Responsibility. doi: 10.1111/beer.70007

[55] Van Oosterhout, M., Waarts, E. and Van Hillegersberg, J. (2006). Change factors requiring agility and implications for IT. European Journal of Information Systems, 15(2), 132-145. doi: 10.1057/palgrave.ejis.3000601

[56] Vietnamese Government. (2021). Decree No. 80/2021/ND-CP detailing and guiding the implementation of certain provisions of the Law on Support for Small and Medium-Sized Enterprises. url: https://thuvienphapluat.vn/van-ban/Doanh-nghiep/Nghi-dinh-80-2021-ND-CP-huong-dan-Luat-Ho-tro-doanh-nghiep-nho-va-vua-486147.aspx [Accessed 12/1/2025]

[57] Vu, M. H. and Tran, H. C. (2022). The factors of financial performance of SMEs (case of Vietnam). Contemporary Economics, 16(3), 346-360. doi: 10.5709/ce.1897-9254.486

[58] Vuković, M. (2024). CB-SEM vs PLS-SEM comparison in estimating the predictors of investment intention. Croatian Operational Research Review, 15(2), 131-144. doi: 10.17535/crorr.2024.0011

[59] Wu, W., Song, J., Lu, L. and Guo, H. (2024). Is managerial ability a catalyst for driving digital transformation in enterprises? An empirical analysis from internal and external pressure perspectives. Plos one, 19(2), e0293454. doi: 10.1371/journal.pone.0293454

[60] Yi, H.-T., Oh, D. and Amenuvor, F. E. (2023). The effect of SMEs' dynamic capability on operational capabilities and organisational agility. South African Journal of Business Management, 54(1), a3696. doi: 10.4102/sajbm.v54i1.3696

Country-specific financial distress prediction using decision trees: Empirical evidence from Slovakia and Croatia

Dominika Gajdošíková[1,*], **Katarína Valášková**[2] **and Tomislava Pavić Kramarić**[2]

[1] *University of Žilina, Faculty of Operation and Economics of Transport and Communications, Department of Economics, Univerzitná 1, 010 26 Žilina, Slovakia*
E-mail: ⟨dominika.gajdosikova@uniza.sk, katarina.valaskova@uniza.sk⟩

[2] *University of Split, Department of Forensic Sciences, R. Boškovića 33, 21 000 Split, Croatia*
E-mail: ⟨tpkramaric@forenzika.unist.hr⟩

Abstract. This paper develops and compares financial distress (FD) prediction models for Slovak and Croatian firms based on a decision tree (DT) approach. The study aims to identify the crucial financial variables that predict firm FD in these two post-transition economies that share historical and institutional similarities but differ in the structure of the economy. A database of Slovak and Croatian enterprises was analysed, focusing on firms being classified as prosperous or non-prosperous based on standardized accounting parameters from financial reports. DTs were employed since they can model complex, nonlinear relationships without diminishing interpretability. The results indicate that high levels of total indebtedness, equity leverage, and debt-to-equity ratios are cross-sectionally associated with an increased likelihood of FD. In the Slovak context, liquidity, particularly the current ratio, showed greater importance than profitability measures, whereas in Croatia profitability ratios, especially return on assets, gained relatively higher relevance compared to liquidity. The findings close a research gap concerning country-specific differences in Central and Eastern European (CEE) FD predictors. The paper provides theoretical contributions by refining knowledge on insolvency drivers in post-transition economies and practical contributions by understanding tailored early warning systems. The study is limited by sample restrictions and the omission of macroeconomic factors. Future research may incorporate wider contextual determinants and alternative machine learning (ML) approaches.

Keywords: financial distress prediction, decision trees, financial distress, firm-level analysis, post-transition economies

Received: July 15, 2025; accepted: February 20, 2026; available online: May 13, 2026

DOI: 10.17535/crorr.2026.0020

Original scientific paper.

1. Introduction

Financial distress (FD) prediction remains a central topic in financial and economic research due to its implications for corporate governance, financial stability, and regulatory oversight [18, 38]. The increasing demand for reliable early-warning systems has intensified the search for accurate yet interpretable predictive models. This is particularly relevant for countries in Central and Eastern Europe (CEE), where the legacy of centrally planned economies, the post-transition reforms, and the gradual integration into European institutional frameworks have created unique patterns of corporate FD. Among these countries, Slovakia and Croatia offer a compelling comparison due to their shared post-socialist history and European Union (EU) membership, yet also notable differences in macroeconomic trajectories, sectoral structures,

*Corresponding author.

and institutional maturity. Slovakia is more manufacturing- and export-oriented, with deeper integration into European value chains and higher credit penetration, whereas Croatia has a larger share of small and medium-sized enterprise (SMEs) and relies more heavily on service-based sectors such as tourism. These structural differences may influence the role of specific predictors, as liquidity shortages are often more pressing in SME-dominated economies, while profitability indicators tend to be more relevant in investment- and export-driven contexts. In both Slovakia and Croatia, the prediction and management of corporate FD have gained increasing policy relevance over the past two decades. Following the transition to market economies, insolvency frameworks and financial reporting standards have gradually aligned with EU regulations, improving transparency and enabling more systematic firm-level analysis [30, 13]. However, institutional maturity, credit availability, and sectoral dynamics continue to differ, suggesting that the explanatory power of financial ratios may not be uniform across national contexts.

Despite the progress in financial infrastructure, empirical research on FD prediction in these two countries, especially from a comparative data-driven, perspective remains limited. Much of the existing literature in the CEE region either focuses on single country analyses or applies generic models borrowed from Western contexts that may not account for the institutional and economic particularities of post-transition economies (e.g. [36, 41, 5, 43], etc). Moreover, traditional statistical techniques such as logistic regression or discriminant analysis may struggle to capture nonlinear interactions among financial variables [3]. There is thus a growing need for models that are not only empirically robust but also context-sensitive and interpretable for practical use. Recent comparative research confirms that the selection of an appropriate prediction method significantly influences classification performance and interpretability, particularly in SME and transition-economy contexts [32, 28]. Empirical evidence increasingly highlights the competitive performance of DT-based approaches in bankruptcy and FD modelling. Ďurica et al. [10] developed a decision tree (DT) model for business failure prediction in Poland and demonstrated its practical applicability in identifying high-risk firms. European evidence beyond the CEE region also indicates that DT models achieve strong predictive performance while maintaining interpretability, especially when compared with more complex artificial intelligence approaches [15, 11, 33]. These findings suggest that DT techniques represent a suitable compromise between predictive accuracy and transparency, which is particularly relevant for regulatory bodies, credit institutions, and policymakers operating in post-transition economies.

This study addresses this research gap by developing and comparing DT models for FD prediction in Slovakia and Croatia using large-scale firm-level financial data. DTs offer a compelling methodological advantage. They are capable of handling nonlinearity and interactions among predictors while preserving interpretability, an essential feature for practitioners, financial analysts, and regulatory authorities. By applying the same modelling framework to both national datasets, the study ensures comparability while allowing for the identification of both common and country-specific FD indicators. The findings demonstrate that while certain financial indicators consistently emerge as strong predictors of insolvency across different national settings, there are also meaningful variations that point to the influence of broader contextual factors. These differences likely stem from variations in financial systems, regulatory frameworks, access to credit, tax policies, and business practices, as well as sectoral compositions and macroeconomic conditions. For instance, the relative importance of liquidity, leverage, or profitability ratios may vary depending on how firms are financed, how risk is assessed by creditors, or the extent to which FD procedures are enforced. Such cross-contextual divergence underscores the importance of not relying on a one-size-fits-all model for FD prediction. Even when applying a consistent methodological framework, the predictive accuracy and relevance of specific indicators depend on the institutional and economic environment in which firms operate. This comparative approach reveals the necessity of adapting analytical models to reflect local conditions, thereby enhancing their practical applicability and robustness. It also offers

valuable insights for policymakers and regulators, who may use these findings to better understand systemic risks, identify early warning signals, and design more targeted interventions to promote financial stability and corporate resilience. Ultimately, the study contributes to the broader literature on FD by demonstrating how data-driven approaches can uncover both universal patterns and context-specific nuances in financial vulnerability. To operationalize this contribution, the empirical analysis is structured around two interrelated research questions. The first addresses the identification and comparative relevance of financial indicators associated with FD risk in Slovak and Croatian firms. The second evaluates the effectiveness of DT models in classifying firms as prosperous or non-prosperous across distinct economic and institutional settings.

By focusing on Slovakia and Croatia, two EU member states at different stages of economic development within the same regional and historical framework, this paper contributes to a deeper understanding of how national contexts influence the nature and predictability of corporate FD. It extends the literature on CEE insolvency prediction by showing that the explanatory power of financial ratios is not uniform across countries, and that tailored country-specific models are necessary to capture the complexity of business failure in diverse economic systems. This study also responds to a broader theoretical and empirical gap, the underrepresentation of smaller, post-transition economies in the global FD prediction literature. While extensive research has been devoted to large economies or mature markets, CEE countries often remain peripheral in comparative modelling studies. The lack of attention to national heterogeneity not only limits the generalizability of widely used models but also underestimates the role of institutional and structural differences in shaping financial behaviour and corporate risk. By generating empirical evidence from Slovakia and Croatia and comparing results through a unified framework, this study enhances the regional understanding of insolvency drivers and offers a pathway toward more precise and effective risk assessment strategies.

In addition to its academic contributions, the findings have significant practical implications for financial institutions, credit rating agencies, policymakers, and firms themselves. Understanding the most relevant predictors of financial distress within a given national context enables the development of more accurate early warning systems, supports risk-based credit allocation, and fosters proactive regulatory oversight. As both Slovakia and Croatia continue to integrate into broader European financial markets, the ability to identify at-risk firms through data-driven interpretable models becomes increasingly important for economic resilience and sustainable growth. In summary, this paper aims to advance both the methodological frontier and the empirical understanding of FD prediction in post-transition economies. By leveraging DT models and applying them to the cases of Slovakia and Croatia, it demonstrates the importance of national specificity in predictive modelling and contributes new insights to the broader discourse on financial risk in the CEE region.

The paper is structured as follows. After the introductory part, the methodology and data section presents the development and evaluation of a predictive model for estimating corporate FD using DT algorithms. The results and discussion section evaluates the performance of DT models for predicting FD among Slovak and Croatian firms. It compares classification accuracy, variable importance, and predictive metrics across the two countries, demonstrating that while both models are highly effective, the Croatian model exhibits slightly better balance and precision, confirming the strength of machine learning (ML) approaches in national insolvency prediction. Finally, the conclusion summarizes the key findings, discusses their practical implications and suggests directions for future research.

2. Methodology and data

The empirical design relies on firm-level financial data, clearly specified classification criteria, and a unified modelling framework applied separately to Slovakia and Croatia to ensure cross-

country comparability. All financial ratios were constructed in a consistent manner across both samples, and identical modelling settings were retained to preserve methodological symmetry. DT models were selected for their interpretability and capacity to capture nonlinear relationships and interaction effects among financial indicators.

2.1. Data and preprocessing

The firm-level financial data for Slovak and Croatian companies were used, extracted from the ORBIS database containing standardized financial and ownership records on enterprises worldwide. The ORBIS database was queried using predefined filters to ensure economic relevance and data consistency. The search was restricted to active companies (including those with unknown current status), excluding public authorities and government entities. Only firms with total assets exceeding EUR 200,000 and with available financial statements for recent fiscal years were retained. Companies lacking standardized financial reporting were excluded. After applying these criteria, dataset contained 27,616 Slovak and 23,565 Croatian firms. The sample includes all industries in the economy, and companies were sampled regardless of industry type, thereby ensuring adequate representation of the corporate sector in both countries. Financial ratios during 2022 were used as independent variables, whereas the dependent variable, financial stability in 2023, captures FD. To ensure data quality and consistency, firms with incomplete or missing financial data were excluded from the analysis. To minimize the effect of extreme values and the asymmetry of the distribution of financial indicators, the dataset was winsorized by the Z-score method with cutoff points at ±3 standard deviations from the mean [48]. After preprocessing and data-cleaning procedures, the final analytical sample consisted of 4,761 Slovak and 1,186 Croatian firms, each classified as either prosperous or non-prosperous according to the defined criteria.

For model development and validation, the data were randomly divided into training (70 %) and test (30 %) sets using stratified sampling to ensure that the original distribution of class labels was preserved. The DT approach was utilized for its ability to model non-linear interactions and detect intricate interactions between financial variables with high interpretability. The choice of predictors is derived from existing empirical studies (e.g., [35, 47, 11, 8, 30, 12]). The summarized formulas of selected financial indicators are presented in Table 1.

Abbreviation	Indicator	Formula
QR	Quick ratio	Current assets minus inventories to current liabilities
CurrR	Current ratio	Current assets to current liabilities
Ins	Insolvency ratio	Total liabilities to total assets and other current assets
IT	Inventory turnover ratio	Sales to inventories
CollP	Collection period ratio	Receivables to sales (multiplied by 365 days)
CredP	Credit period ratio	Creditors to sales (multiplied by 365 days)
ROA	Return on assets ratio	Profit after tax to total assets
ROE	Return on equity ratio	Profit after tax to shareholders' funds
ROS	Return on sales ratio	Profit after tax to sales
TI	Total indebtedness ratio	Total liabilities to total assets
DE	Debt-to-equity ratio	Total liabilities to shareholders' funds
EL	Equity leverage ratio	Total assets to shareholders' funds

Table 1: *Summarized formulas of financial indicators.*

In several ratios, the denominator can take on values close to zero or even negative, most notably in the case of equity-based measures such as the equity leverage or the debt-to-equity ratios. Instead of excluding these observations or applying winsorization, such cases were retained in the dataset, as they represent genuine FD situations (e.g., firms with negative equity

or excessive short-term liabilities). This approach is consistent with prior research, which emphasizes that extreme or non-standard financial ratios may carry critical information about the deteriorating financial condition of a firm. While this choice increases variance in the predictor space and may generate outliers, it was considered preferable to preserve the economic reality of the data and ensure that distress signals were not artificially suppressed. To mitigate potential issues of numerical instability, all ratios were computed consistently across firms, and their distributional properties were reviewed prior to modelling.

2.2. Classification criteria

To predict FD, the enterprises included in the model formation are categorized into two groups: prosperous enterprises with sustainable debt and non-prosperous enterprises with substantial indebtedness. The classification of firms into prosperous and non-prosperous categories follows the accounting-based distress identification framework proposed by Klieštik et al. [26], which is widely applied in Central and Eastern European insolvency research.

Condition (1): The firm reports negative equity, where the value of total liabilities (including accrued liabilities) is greater than total assets, defined as:

$$\text{Assets} - (\text{Liabilities} + \text{Accrued liabilities}) < 0 \tag{1}$$

Condition (2): If Condition (1) is violated, the firm is not thriving whenever its bonity index, an adjusted liquidity measure, is less than 1. The index is calculated as the percentage ratio of adjusted current assets to adjusted current liabilities:

$$\text{Bonity index} = \frac{\text{Adjusted Current Assets}}{\text{Adjusted Current Liabilities}} < 1 \tag{2}$$

Condition (3): When data for the bonity index are incomplete or unavailable, a simplified liquidity ratio L_3 is used instead, which sums inventories, receivables, financial assets, cash, and prepaid expenses, divided by the sum of current liabilities and accrued liabilities. The firm is non-prosperous if this ratio is less than 1:

$$L_3 = \frac{\begin{array}{c}\text{Inventories} + \text{Current Receivables} + \text{Financial Assets} \\ +\text{Cash and Equivalents} + \text{Prepaid Expenses}\end{array}}{\text{Current Liabilities} + \text{Accrued Liabilities}} < 1 \tag{3}$$

The two ratios both capture liquidity, but with the difference that the bonity index is more precise, while L_3 ensures practicability even in the instance of incomplete information.

Condition (4): The firm recorded a net loss during the observed accounting period, an indication of insufficient revenues to meet financial obligations:

$$\text{Net Profit After Tax} \leq 0 \tag{4}$$

If a firm does not meet Condition (1) of negative equity, it will still be labeled as non-prosperous if, simultaneously, it experiences inadequate liquidity, as shown by either Condition (2) or Condition (3), and reports a net loss based on Condition (4). Based on these criteria, out of 4,761 Slovak firms in the sample, 560 were classified as non-prosperous and 4,201 as prosperous. In Croatia, out of 1,186 firms, 57 were classified as non-prosperous and 1,129 as prosperous. This distribution reveals a pronounced class imbalance in both datasets, with distressed firms representing only around 12% of Slovak firms and 5% of Croatian firms.

Table 2 illustrates the distribution of firms according to the applied criteria. In Slovakia, the combination of liquidity shortfall and net loss was the most frequent driver of classification, while negative equity accounted for a smaller but still substantial share of distressed cases.

By contrast, in Croatia, negative equity was more prevalent, representing the majority of non-prosperous firms, with the remainder identified through the combined liquidity and profitability criteria. This breakdown illustrates how different mechanisms of financial fragility dominate across the two countries and confirms the complementary role of the applied conditions in capturing firm-level distress.

Classification criterion	**SK**	**HR**
Negative equity (Condition 1)	206	32
Liquidity shortfall combined with net loss (Conditions 2/3 + 4)	354	25
Total non-prosperous firms	560	57

Table 2: *Classification of non-prosperous enterprises based on criteria.*

The application of a multi-criteria approach makes it easier to identify financially distressed firms even in the absence of a court-driven FD process, thereby maintaining consistency and comparability across countries. The output Y of the model is binary classification, with 0 standing for a prosperous firm and 1 for a non-prosperous firm:

$$Y = \begin{cases} 0 & \text{for prosperous firm} \\ 1 & \text{for non-prosperous firm} \end{cases} \tag{5}$$

Although the ORBIS database provides bankruptcy event indicators, these were not used as the outcome variable. In the Slovak and Croatian context, bankruptcy filings are often underreported, delayed, or affected by institutional factors that may not adequately reflect firms' financial condition. For consistency and comparability, FD was instead operationalized through accounting-based criteria, which are commonly applied proxies in the literature. It should be noted that the classification may be sensitive to threshold choices: lowering the liquidity cutoff to 0.9 would restrict the definition of distress to more extreme cases, while raising it to 1.1 would broaden the pool of firms at risk. While such changes could affect absolute counts of distressed firms, the overall class imbalance would remain, and the relative interpretation of minority vs. majority classes would not materially change.

While these criteria are conceptually related to some predictor variables, lagged financial ratios were employed to ensure temporal separation between predictors and outcomes and to reduce concerns of circularity. The analysis was restricted to financial ratios in order to maintain methodological comparability with prior research.

2.3. Predictive modelling method

This study employs the DT model to classify enterprises into prosperous and non-prosperous ones based on their financial characteristics. DTs represent a class of supervised learning methods that recursively partition the predictor space to maximize class homogeneity. Compared to traditional linear statistical models, DTs do not require distributional assumptions, can naturally accommodate nonlinear and interaction effects among financial ratios, and remain robust in the presence of multicollinearity [40]. In addition, their hierarchical structure ensures high interpretability, allowing explicit identification of decision rules and threshold values, which is particularly relevant for early-warning applications in FD prediction.

In comparison with ensemble approaches such as random forests, single-tree models may exhibit lower predictive variance but can be more sensitive to overfitting if not properly pruned. Random forest models often achieve higher predictive accuracy due to aggregation across multiple trees. However, this comes at the cost of reduced interpretability, as the internal decision structure becomes opaque [45]. Given that the objective of this study is not only to achieve

high classification performance but also to identify and compare country-specific FD determinants, interpretability was considered a primary methodological criterion. Therefore, a single DT framework was preferred over more complex ensemble techniques.

Model development is reliant on the classification and regression trees (C&RT) algorithm, where a binary recursive partitioning approach is used to split the dataset into increasingly homogeneous subgroups based on the dependent variable. In this study, the DTs were implemented in SPSS using the CRT growing method with the Gini index as the splitting criterion. The dependent variable was the binary classification of firms with or without financial difficulties, while the independent variables included a set of financial ratios.

For each node, the algorithm examines all possible splits on each predictor and selects the one that results in the most dissimilar child nodes in terms of class membership. In the SPSS C&RT procedure, default hyperparameters were retained: a maximum tree depth of 5, a minimum of 10 cases in parent nodes, and a minimum of 5 cases in child nodes. Pruning was performed automatically by the software to avoid overfitting.

To make the best splits, the algorithm attempts to minimize impurity and to perform this minimization, and it applies the Gini index, a standard measure that approximates node heterogeneity. The Gini index is defined as:

$$\text{Gini} = 1 - \sum_{i=1}^{C} (p_i)^2 \tag{6}$$

where p_i is the likelihood that an observation at random is in category i, and C is the number of categories. A node can be considered pure when all observations in it are from one class, indicating that the Gini index equals zero. The algorithm seeks to create nodes of minimum impurity [29]. Prior comparative studies in bankruptcy prediction demonstrate that DT-based models achieve competitive predictive performance relative to neural networks and support vector machines, particularly when interpretability is required [6, 11]. In several Central and Eastern European applications, DT models have reported classification accuracies exceeding 90%, confirming their empirical robustness in insolvency modelling contexts [12]. One of the main advantages of DTs is their ability to automate variable selection in tree construction. Irrelevant and poor predictors are automatically eliminated, reducing model complexity and making it easier to interpret. This is particularly valuable in high-dimensional financial data, where multicollinearity and noise would stop optimal performance in conventional statistical models. Predictor importance was derived from the reduction in Gini impurity aggregated across all splits where a variable was used. These importance values were later reported in the results to identify the financial ratios most relevant to predicting FD.

To mitigate overfitting risk, pruning procedures were applied within the CRT framework to ensure generalizability of the final model. The DT is trained on a 70 % stratified training sample, while predictive performance was evaluated on the remaining 30 % test set using standard classification metrics.

2.4. Model evaluation metrics

The predictive performance of DT models was assessed through three popular classification measures: AUC, accuracy, and F1 score. AUC indicates how well the model can discriminate between prosperous and non-prosperous firms at various classification thresholds, with values closer to 1 indicating higher discrimination power. Accuracy computes the overall ratio of correctly predicted instances and provides a straightforward measure of correctness. The F1 score, the harmonic mean of recall and precision, is particularly beneficial in imbalanced datasets, balancing the false positive and true negative trade-offs with more sensitivity than accuracy by itself. According to Berrada et al. [7], these metrics provide an integrated measurement

framework for FD models. All classification performance metrics (precision, recall, F1 score, and AUC) are evaluated with respect to the minority class of non-prosperous firms, which represents the event of interest in FD prediction.

3. Results and discussion

Table 3 summarises the classification performance of the DT models for Slovakia and Croatia. In both countries, the models classify prosperous firms with very high accuracy. In Slovakia, 99.6% of prosperous firms are correctly classified in the training sample and 99.3% in the test sample, while the identification of non-prosperous firms reaches only 15.4% and 24.2%, respectively. The overall test accuracy is 91.26%. In Croatia, the model correctly classifies 98.5% of prosperous firms in training and 99.7% in testing, whereas the identification of distressed firms improves to 46.3% and 50.0%. The overall test accuracy reaches 97.38%.

Classification							
		SK			**HR**		
Sample	**Observed**	**Predicted**		**Percent**	**Predicted**		**Percent**
		0	**1**	**Correct**	**0**	**1**	**Correct**
Training	0	2,928	13	99.6%	790	12	98.5%
	1	330	60	15.4%	22	19	46.3%
	Overall Percent	97.8%	2.2%	89.7%	96.3%	3.7%	96.0%
Testing	0	1,268	9	99.3%	326	1	99.7%
	1	116	37	24.2%	8	8	50.0%
	Overall Percent	96.8%	3.2%	91.3%	97.4%	2.6%	97.4%

Table 3: *Classification performance of DT models for Slovakia and Croatia.*

Given the pronounced class imbalance, where non-prosperous firms represent only around 5–11% of observations, overall accuracy is largely driven by the majority class (Table 4). The Croatian model achieves a higher discriminatory power, with an AUC of 0.922 compared to 0.861 in Slovakia. Precision remains high in both countries, however, recall is substantially stronger in Croatia, resulting in a higher F1 score. Overall, the Croatian model provides a more balanced early-warning performance, whereas the Slovak model remains more biased toward the majority class.

Country	**AUC**	**Accuracy**	**Precision**	**Recall**	**F1 Score**
SK	0.861	0.9126	0.8043	0.2418	0.3719
HR	0.922	0.9738	0.8889	0.5000	0.6400

Table 4: *Prediction results and models accuracy for Slovakia and Croatia.*

The relatively low recall confirms that model performance is constrained by the minority representation of non-prosperous firms. While overall accuracy and precision remain high, these metrics are largely driven by the dominant class of prosperous enterprises. This pattern is consistent with prior FD prediction research and reflects the inherent difficulty of early-warning modelling under imbalanced data conditions. Future methodological refinements, such as resampling strategies or cost-sensitive learning approaches, may help improve the detection rate of distressed firms.

These findings compare to or surpass those of comparable benchmarks in earlier Slovak research. Adamko and Švábová [1] constructed Altman-based logistic models with 0.81-0.88 AUC scores, reiterating high discriminatory power. Conversely, Gregová et al. [19] and Horváthová et al. [22] confirmed that ANN operated with 94.37% and 95.45% accuracy, respectively. Ďurica

et al. [11] were slightly more precise in their findings, employing ANN (96.5%) and DT (93.2%) models, although their data horizon was shorter. Horák et al. [21], employing a mixed evaluation approach with both financial and non-financial performance indicators, applied artificial neural networks and logistic regression models. Their results showed overall accuracy close to 81%, with higher accuracy (almost 90%) for firms surviving FD, but considerably lower accuracy (around 55%) for identifying firms at risk of bankruptcy. In this case, the performance of the developed model reaches 91.26% accuracy and an AUC of 0.861. While competitive in terms of overall accuracy compared to past Slovak studies, its recall is modest (24.18%), underscoring the challenge of early-warning performance under class imbalance.

In the Croatian context as well, the DT model of this study is much improved over past national endeavors. Such past models used to perform inadequately at lower-quality categorizations or possess limited generalizability. Zenzerović [46] achieved 95.3% accuracy using multiple discriminant analysis (MDA), but from the estimation sample only, limiting its external validity. Bogdan [9] applied logistic regression to the restaurant industry to achieve 82.8% accuracy, whereas Pervan et al. [34] achieved merely 76.3% on a larger dataset. In that context, the Croatian DT model achieved 97.38% accuracy and an AUC of 0.922, with a precision of 88.89% and recall of 50.00% for distressed firms. Some recent Croatian studies extend FD prediction beyond traditional financial ratios. In addition, Galant and Zenzerović [17] examined the role of corporate social responsibility, highlighting the potential relevance of non-financial determinants in distress modelling. Although such variables are not incorporated in the present study, their inclusion may represent a promising direction for future research aimed at enriching firm-level prediction frameworks. Arnerić and Moćan [4] applied Bayesian logistic regression to predict bankruptcy in trading companies under uncertainty, reporting high predictive accuracy and strong identification of distressed firms compared to conventional approaches.

The structure of the trained DT models reflects the hierarchical importance of selected financial indicators identified during model estimation. A graphical representation of the full tree structures is provided in the online supplementary material [16]. The relative importance of independent variables to predict FD using DT models exhibits both consistent patterns and profound national differences (Table 5). For both countries, total indebtedness, equity leverage, and debt-to-equity ratio are rated as the most significant predictors with importance values of close to or exactly 100%. These ratios collectively capture the leverage position and corporate capital structure, underlining their central role in FD prediction. In the Slovak model, among liquidity measures, the current ratio emerges as one of the most important variables, exceeding the predictive relevance of profitability indicators. Profitability ratios, especially return on assets, followed by return on equity and return on sales, also contribute meaningfully to classification performance, although to a lesser extent. Additional liquidity and solvency indicators, including the quick ratio and insolvency ratio, further support the differentiation between prosperous and non-prosperous firms. These findings suggest that capital structure and short-term liquidity conditions play a dominant role in explaining FD risk in Slovakia, while profitability measures provide complementary explanatory power. In contrast, the Croatian model is primarily driven by leverage-related indicators, with total indebtedness, debt-to-equity, and equity leverage emerging as the most influential predictors. Beyond leverage, return on assets and the current ratio show substantial predictive importance, followed by return on sales and return on equity. The insolvency ratio exhibits only marginal relevance, while the quick ratio and turnover-based variables contribute negligibly to classification performance. Overall, while leverage-related metrics are consistently central across both countries, the Slovak model places relatively stronger emphasis on liquidity conditions, particularly the current ratio, whereas the Croatian model assigns comparatively greater relevance to profitability indicators.

The findings of FD prediction models for Croatian and Slovak enterprises corroborate the established applicability of leverage ratios as primary drivers of FD, consistent with earlier empirical research in these countries and Europe. In the Slovak context, the total indebted-

Independent variable importance				
	SK		HR	
Variable	Importance	Percent	Importance	Percent
Quick ratio	0.008	21.7%	0.000	0.0%
Current ratio	0.019	50.3%	0.006	26.2%
Insolvency ratio	0.007	18.5%	0.000	1.3%
Inventory turnover ratio	0.002	5.2%	0.000	0.0%
Collection period ratio	0.000	1.1%	0.000	0.0%
Credit period ratio	0.002	4.7%	0.000	0.1%
Return on assets ratio	0.017	45.1%	0.010	41.8%
Return on equity ratio	0.013	35.0%	0.003	10.6%
Return on sales ratio	0.013	35.1%	0.004	16.4%
Total indebtedness ratio	0.037	100.0%	0.024	100.0%
Debt-to-equity ratio	0.037	97.7%	0.023	96.5%
Equity leverage ratio	0.037	97.7%	0.023	96.5%

Table 5: *Independent variable importance in DT models for Slovakia and Croatia.*

ness ratio is the most prevailing variable, consistent with Mihalovič [31] and Horváthová et al. [23], both of which analysed Slovak companies, and similarly identified leverage measures as primary in artificial neural network (ANN) models. The extremely high importance of the debt-to-equity ratio is also in agreement with findings by Korol [27], focused on Central and Eastern Europe, and Behr and Weinblat [6], who examined German enterprises using machine learning methods including DTs and pointed to its central role in ML models of FD prediction. Concerning liquidity, the importance of current ratio in Slovakia validates research by Radovanovic and Haas [37] on European enterprises and Slovak-specific research by Weiss et al. [44], and Jenčová et al. [24] that liquidity is a significant but secondary factor after leverage. However, when comparing individual indicators, profitability ratios rank higher than quick and insolvency ratios, which suggests that profitability plays a more consistent role than some liquidity measures. This pattern indicates that liquidity ratios are crucial in ascertaining the short-term ability of a firm to settle its immediate financial obligations, but profitability also provides an important complementary signal of financial health. Firms with low liquidity are more vulnerable to cash flow disruptions that can rapidly aggravate FD. The moderate importance of profitability ratios echoes the same research and regional studies, such as Tudor et al. [42] focused on Romania, implying a complementary yet less determining role in prediction. Operating and insolvency-based ratios were not influential, consistent with the prior literature that these variables generally have limited contributions compared to basic financial health indicators.

For Croatian firms, leverage ratios again emerge as premier predictors, consistent with overall European results presented by Behr and Weinblat [6] as well as by Pervan et al. [35], who used logit models on Croatian enterprises, and Žiković [47], who employed discrete-time hazard models. The relatively weaker importance of liquidity and profitability ratios compared to Slovakia can be attributed to structural and economic differences within the Croatian corporate sector. Croatia's economy, having a larger share of SMEs and industries with distinctive financial profiles, may experience different patterns of FD where liquidity shortages are less prominent or manifest in unique ways. Specifically, according to European Commission data [14], the SME share of value added is higher in Croatia than in Slovakia, while the credit-to-GDP ratio is markedly lower, indicating more limited access to bank financing. These structural conditions help explain why leverage indicators dominate in Croatia, whereas liquidity indicators retain greater importance in Slovakia. Institutional settings further reinforce these differences, as variations in restructuring and liquidation frameworks across CEE significantly influence the detection and management of FD [25]. Similarly, comparative research on SME-led growth in

Adriatic economies highlights the distinct weight of smaller firms in Croatia's economy [20], which can amplify the reliance on leverage-based indicators. Furthermore, regulatory and market conditions, including access to credit and alternative bankruptcy proceedings, can influence which financial ratios are most predictive of distress. The marginal impact of operating performance ratios and interest coverage ratios in our Croatian model also aligns with Mihalovič [31] that these variables have limited use in this specific setting. Such contextual factors underscore the need to tailor bankruptcy prediction models to the country-specific institutional setting and firm characteristics to increase their practical relevance and accuracy. Previous research on Croatian firms has highlighted the relevance of macroeconomic and institutional factors in explaining FD dynamics [47, 2, 39]. Although the present study focuses exclusively on firm-level financial indicators, these findings suggest that incorporating macro-level determinants may further enhance distress prediction models and represents a promising direction for future research.

Together with these results, the outcomes of prior Slovak and Croatian studies employing ML approaches validate the increasing dominance of ML methods over traditional ones in FD prediction in Slovakia and Croatia. Previous work using MDA or logistic regression [4] showed acceptable yet comparatively weaker results. ANN and DT models [11, 12] have, however, surpassed these benchmarks, with several of them exceeding 90% accuracy. The developed DT models achieve competitive performance relative to these benchmarks, particularly in terms of overall accuracy and discriminatory power. This finding further supports the growing application of interpretable ML techniques in insolvency prediction.

4. Conclusions

This study develops and cross-compares FD prediction models for Slovak and Croatian firms, two CEE economies that remain relatively underrepresented in insolvency research. The findings confirm the dominant role of leverage indicators, particularly total indebtedness and debt-to-equity ratios, in both countries. At the same time, important country-specific differences emerge. In Slovakia, liquidity, especially the current ratio, provides substantial complementary predictive power alongside leverage, while in Croatia profitability ratios, particularly return on assets, assume relatively greater importance compared to liquidity measures. These results deepen the understanding of how institutional and structural differences shape the explanatory power of financial ratios and highlight the need for country-specific FD modelling frameworks.

From a practical perspective, the results provide relevant guidance for financial analysts, credit institutions, and policymakers by identifying key financial indicators for early-warning monitoring. The interpretability of DT models enhances their practical usability, as decision rules and threshold values can be transparently communicated and applied in risk assessment processes. The observed cross-country variation further suggests that predictive systems should reflect national economic structures and regulatory environments rather than rely on universal ratio hierarchies.

Several limitations should be acknowledged. The models rely exclusively on accounting-based financial data and therefore do not capture qualitative determinants such as management quality, governance structure, or market-specific shocks. The analysis is restricted to two countries, which limits broader generalization across the CEE region. Moreover, FD is operationalized through accounting-based criteria rather than observed bankruptcy events, and the pronounced class imbalance constrains recall despite high overall accuracy. The models were estimated using default hyperparameter settings and were not systematically benchmarked against alternative ML or regression-based approaches. Future research should therefore incorporate non-financial and macroeconomic determinants, explore cost-sensitive or optimized modelling strategies, and extend the analysis across additional countries and time horizons to strengthen robustness and external validity.

Acknowledgements

This research was financially supported by the Slovak Research and Development Agency Grant VEGA 1/0494/24: Metamorphoses and causalities of indebtedness, liquidity and solvency of companies in the context of the global environment.

References

[1] Adamko, P. and Švábová, L. (2016). Prediction of the risk of bankruptcy of Slovak companies. In Managing and Modelling of Financial Risks: 8th International Scientific Conference (pp. 15–20).

[2] Altman, E. I., Balzano, M., Giannozzi, A., Liguori, E. and Srhoj, S. (2025). Bouncing back to the surface: Factors determining SME recovery. Journal of small business management, 63(4), 1856-1883. doi: 10.1080/00472778.2024.2415302

[3] Amin, M. B., Asaduzzaman, M., Debnath, G. C., Rahaman, M. A. and Oláh, J. (2024). Effects of circular economy practices on sustainable firm performance of green garments. Oeconomia Copernicana, 15(2), 637-682. doi: 10.24136/oc.2795

[4] Arnerić, J. and Moćan, M. (2025). A Bayesian approach to logistic regression for bankruptcy prediction of trading companies under uncertainty conditions. Ekonomski pregled, 76(1), 15-34. doi: 10.32910/ep.76.1.2

[5] Asad, A. I., Popesko, B. and Damborský, M. (2024). The nexus between economic policy uncertainty and innovation performance in Visegrad group countries. Oeconomia Copernicana, 15(3), 1067-1100. doi: 10.24136/oc.2804

[6] Behr, A. and Weinblat, J. (2017). Default patterns in seven EU countries: A random forest approach. International Journal of the Economics of Business, 24(2), 181–222. doi: 10.1080/13571516.2016.1252532

[7] Berrada, I. R., Barramou, F. and Alami, O. B. (2024). Towards a Machine Learning-based Model for Corporate Loan Default Prediction. International Journal of Advanced Computer Science and Applications, 15(3), 565-573. doi: 10.14569/IJACSA.2024.0150357

[8] Boďa, M. and Úradníček, V. (2023). Definition of financial distress supported by data in Slovak corporate conditions. European Journal of International Management, 21(4), 582-604. doi: 10.1504/EJIM.2023.134646

[9] Bogdan, S., Šikić, L. and Bareša, S. (2021). Predicting bankruptcy based on the full population of Croatian companies. Ekonomski Pregled, 72(5), 643–669. doi: 10.32910/ep.72.5.1

[10] Ďurica, M., Frnda, J. and Švábová, L. (2019). Decision tree based model of business failure prediction for Polish companies. Oeconomia Copernicana, 10(3), 453–469. doi: 10.24136/oc.2019.022

[11] Ďurica, M., Frnda, J. and Švábová, L. (2023). Artificial neural network and decision tree-based modelling of non-prosperity of companies. Equilibrium. Quarterly Journal of Economics and Economic Policy, 18(4), 1105–1131. doi: 10.24136/eq.2023.035

[12] Ďuricová, L., Kovalová, E., Gazdíková, J. and Hamranová, M. (2025). Refining the Best-Performing V4 Financial Distress Prediction Models: Coefficient Re-Estimation for Crisis Periods. Applied Sciences, 15(6), 2956. doi: 10.3390/app15062956

[13] Dvorský, J. (2025). Impact of Artificial Intelligence on Enterprise Risk Management. A case study from the Slovak SME Segment. Journal of Business Sectors, 3(1), 96–103. doi: 10.62222/CAJA0666

[14] European Commission. (2025). Croatia – SME Country Fact Sheet. url: https://ec.europa.eu/docsroom/documents/60557

[15] Gadžo, A., Suljić, M., Jusufović, A., Filipović, S. and Suljić, E. (2025). Data mining approach in detecting inaccurate financial statements in government-owned enterprises. Croatian Operational Research Review, 16(1), 1–15. doi: 10.17535/crorr.2025.0001

[16] Gajdošíková, D., Valášková, K., and Pavić Kramarić, T. (2026). Supplementary material to: Country-specific financial distress prediction using decision trees: Empirical evidence from Slovakia and Croatia. Zenodo. doi: 10.5281/zenodo.18632753

[17] Galant, A. and Zenzerović, R. (2023). Can corporate social responsibility contribute to bankruptcy prediction? Evidence from Croatia. Organizacija, 56(3), 173-183. doi: 10.2478/orga-2023-0012

[18] Ghaffarian, S., Taghikhah, F.R. and Maier, H.R. (2023). Explainable artificial intelligence in disaster risk management: Achievements and prospective futures. International Journal of Disaster Risk Reduction, 98, 104123. doi: 10.1016/j.ijdrr.2023.104123
[19] Gregová, E., Valášková, K., Adamko, P., Tumpach, M. and Jaroš, J. (2020). Predicting financial distress of Slovak enterprises: Comparison of selected traditional and learning algorithms methods. Sustainability, 12(10), 3954. doi: 10.3390/su12103954
[20] Gričar, S., Šugar, V., and Bojnec, Š. (2019). Small and medium enterprises led-growth in two Adriatic countries: Granger causality approach. Economic research-Ekonomska istrazivanja, 32(1), 2161-2179. doi: 10.1080/1331677X.2019.1645711
[21] Horák, J., Krulický, T., Rowland, Z. and Machová, V. (2020). Creating a comprehensive method for the evaluation of a company. Sustainability, 12(21), 9114. doi: 10.3390/su12219114
[22] Horváthová, J., Mokrišová, M. and Petruška, I. (2021). Selected methods of predicting financial health of companies: Neural networks versus discriminant analysis. Information, 12(12), 505. doi: 10.3390/info12120505
[23] Horváthová, J., Mokrišová, M. and Schneider, A. (2024). The application of machine learning in diagnosing the financial health and performance of companies in the construction industry. Information, 15(6), 355. doi: 10.3390/info15060355
[24] Jenčová, S., Petruška, I. and Miškufová, M. (2024). Relationship between profitability and debt: The case of the Slovak Energy and Mining sector. Ekonomicko-manazerske spektrum, 18(2), 38-49. doi: 10.26552/ems.2024.2.38-49
[25] Kliatskova, T., and Savatier, L. B. (2019). Insolvency regimes and economic outcomes (No. 133). DIW Roundup: Politik im Fokus.
[26] Klieštik, T., Klieštiková, J., Kováčová, M., Švábová, L., Valášková, K., Vochozka, M., and Oláh, J. (2018). Prediction of financial health of business entities in transition economies. New York: Addleton Academic Publishers.
[27] Korol, T. (2019). Dynamic bankruptcy prediction models for European enterprises. Journal of Risk and Financial Management, 12(4), 185. doi: 10.3390/jrfm12040185
[28] Letkovský, S., Jenčová, S. and Vašaničová, P. (2024). Is artificial intelligence really more accurate in predicting bankruptcy? International Journal of Financial Studies, 12(1), 8. doi: 10.3390/ijfs12010008
[29] Liu, J., Wu, C. and Li, Y. (2019). Improving financial distress prediction using financial network-based information and GA-based gradient boosting method. Computational Economics, 53(2), 851-872. doi: 10.1007/s10614-017-9768-3
[30] Michalková, L., Krulický, T. and Kučera, J. (2024). Detection of earnings manipulations during the corporate life cycle in Central European countries. Equilibrium. Quarterly Journal of Economics and Economic Policy, 19(2), 623-660. doi: 10.24136/eq.3030
[31] Mihalovič, M. (2018). Applicability of scoring models in firms' default prediction: The case of Slovakia. Politicka Ekonomie, 66(6), 689–708. doi: 10.18267/j.polek.1226
[32] Mokrišová, M., Horváthová, J., Schneider, A. and Jusková, M. (2023). Prediction of corporate bankruptcy – Selecting an appropriate prediction method. In: 8th International Scientific Conference of the Economics, Management Business (pp. 787–797).
[33] Navarro-Galera, A., Lara-Rubio, J., Novoa-Hernández, P. and Cruz Corona, C. A. (2025). Using decision trees to predict insolvency in Spanish SMEs: Is early warning possible? Computational Economics, 65(1), 91–116. doi: 10.1007/s10614-024-10586-5
[34] Pervan, I., Bartulović, M. and Jozipović, Š. (2023). Are insolvency proceedings opened too late? The case of Germany, Croatia and Slovakia. Ekonomski vjesnik: Review of Contemporary Entrepreneurship, Business, and Economic Issues, 36(2), 387-398. doi: 10.51680/ev.36.2.13
[35] Pervan, I., Pervan, M. and Vukoja, B. (2011). Prediction of company bankruptcy using statistical techniques – Case of Croatia. Croatian Operational Research Review, 2(1), 158–167.
[36] Prusak, B. and Karas, M. (2024). Bankruptcy prediction in Visegrad group countries. Polish Journal of Management Studies, 30(1), 268-288. doi: 10.17512/pjms.2024.30.1.16
[37] Radovanovic, J. and Haas, C. (2023). The evaluation of bankruptcy prediction models based on socio-economic costs. Expert Systems with Applications, 227, 120275. doi: 10.1016/j.eswa.2023.120275
[38] Rech, F., Isaboke, C. and Xu, H. (2025). Surviving the Pandemic: Financial Distress Prediction

for Slovak SME Manufacturers. Journal of Business Sectors, 3(1), 41–51. doi: 10.62222/SNRN2189

[39] Šergo, Z., Gržinić, J. and Sučić Červa, M. (2023). Charting the Course: Total Factor Productivity Trends in Croatia Post-pre-bankruptcy Act. Lexonomica, 15(2), 189-212. doi: 10.18690/lexonomica.15.2.189-212.2023

[40] Švábová, L., Labošová, V., Borovcová, M. and Jarabicová, N. (2024). Customer satisfaction prediction: A case study for electro-bike customers. Ekonomicko-manazerske spektrum, 18(2), 85-98. doi: 10.26552/ems.2024.2.85-98

[41] Toudas, K., Archontakis, S. and Boufounou, P. (2024). Corporate Bankruptcy Prediction Models: A Comparative Study for the Construction Sector in Greece. Computation, 12(1), 9. doi: 10.3390/computation12010009

[42] Tudor, L., Popescu, M. E. and Andreica, M. (2015). A decision support system to predict financial distress: The case of Romania. Romanian Journal of Economic Forecasting, 18(4), 170–179.

[43] Valášková, K., Gajdošíková, D. and Belás, J. (2023). Bankruptcy prediction in the post-pandemic period: A case study of Visegrad Group countries. Oeconomia Copernicana, 14(1), 253-293. doi: 10.24136/oc.2023.007

[44] Weiss, E., Janoskova, M., Culkova, K., Zuzik, J. and Weiss, R. (2023). Prediction of Extraction Companies' Development. Montenegrin Journal of Economics, 19(4), 7-18. doi: 10.14254/1800-5845/2023.19-4.1

[45] Yao, G., Hu, X., Zhou, T., and Zhang, Y. (2022). Enterprise credit risk prediction using supply chain information: A decision tree ensemble model based on the differential sampling rate, Synthetic Minority Oversampling Technique and AdaBoost. Expert Systems, 39(6), e12953. doi: 10.1111/exsy.12953

[46] Zenzerović, R. (2009). Business Financial Problems Prediction-Croatian Experience. Economic research-Ekonomska istrazivanja, 22(4), 1-16. doi: 10.1080/1331677x.2009.11517387

[47] Žiković, I. (2018). Challenges in predicting financial distress in emerging economies: The case of Croatia. Eastern European Economics, 56(1), 1–27. doi: 10.1080/00128775.2017.1387059

[48] Žmuk, B. (2017). Speeding problem detection in business surveys: benefits of statistical outlier detection methods. Croatian operational research review, 8(1), 33-59. doi: 10.17535/crorr.2017.0003

Comparative analysis of modern machine learning models for retail sales forecasting

Luka Hobor[1,2,*], **Mario Brčić**[1,2], **Lidija Polutnik**[3] **and Ante Kapetanović**[4]

[1] *Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3 Zagreb, Croatia*
E-mail: ⟨{luka.hobor, mario.brcic}@fer.hr⟩

[2] *It From Bit d.o.o., Zagreb, Croatia*

[3] *Babson College, 231 Forest Street, Babson Park, MA 02457, Massachusetts, United States*
E-mail: ⟨polutnik@babson.edu⟩

[4] *mStart Plus d.o.o., Slavonska avenija 11a, 10000 Zagreb, Croatia*
E-mail: ⟨Ante.Kapetanovic@mstart.hr⟩

Abstract. Accurate demand forecasting is critical for brick-and-mortar retailers to optimize inventory management and minimize costs. This study evaluates statistical baselines, tree-based ensembles (XGBoost and LightGBM), and deep learning architectures (N-BEATS, N-HiTS, and the Temporal Fusion Transformer) on retail sales data characterized by intermittent demand, substantial missingness, and frequent product turnover. Models are compared across four configurations varying by aggregation level and imputation strategy, using evaluation protocols that reflect typical deployment patterns for each model class. Localized tree-based methods achieve superior performance, with XGBoost attaining the lowest RMSE of 4.833. While SAITS-based imputation improved neural network performance in aggregated settings, these models remained inferior to ensemble methods. The results suggest that, under the studied constraints, model selection should prioritize alignment with problem characteristics over architectural sophistication.

Keywords: gradient-boosted decision trees, neural networks, predictive analytics, retail sales forecasting, time-series analysis

Received: December 12, 2025; accepted: February 6, 2026; available online: May 13, 2026

DOI: 10.17535/crorr.2026.0021

Original scientific paper.

1. Introduction

Accurate sales forecasting is indispensable in retail, enabling better inventory planning, resource allocation, and the achievement of revenue and profit goals. The choice of the best possible forecasting model to use is essential as retailers aim to gain a competitive advantage in the context of irregular demand and limited historical sales data in various product categories. While classical approaches such as the Exponential Time Smoothing [30] or the Theta model [2] have historically formed the backbone of retail forecasting, they often fall short in capturing modern retail complexities and operational nuances. Recent advances in machine learning and deep learning present significant opportunities for retailers to improve their sales predictions and operational efficiencies with new alternatives, such as tree-based models like XGBoost [8] and LightGBM [16], as well as neural architectures such as N-BEATS [27], NHITS [5], TFT [20], and others [26, 25, 21].

*Corresponding author.

Recent comparative evaluations of forecasting models have shown mixed results across domains. Studies such as [23, 6, 3, 7] have explored the performance of neural networks versus gradient boosting models and highlighted that deep learning does not consistently outperform tree-based approaches, especially on tabular, sparse, or highly intermittent retail data. These findings align with results from the M5 forecasting competition [22], which used Walmart's highly granular retail sales data consisting of daily item-store records. While some series exhibited intermittent demand patterns, the dataset primarily featured dense, continuous observations with limited explicit missingness. Despite this, LightGBM-based ensembles outperformed deep learning models, including those developed by Amazon's forecasting team [15]. On the other hand, recent findings from Zalando suggest that transformer-based models exhibit scaling laws in retail forecasting: as the volume of training data increases, demand forecasting error decreases in a predictable manner [18]. These understandings motivate a closer investigation into the conditions under which different forecasting paradigms perform best.

Much of the recent deep learning success has been in large-scale online retail settings, where companies like Amazon [1] and Zalando [18] operate centralized warehouses and enjoy the benefits of aggregated demand and operational homogeneity. In contrast, brick-and-mortar (B&M) retail is much more fragmented: sales occur in thousands of physical stores, each with limited shelf space, store-level variability, and frequent changes in product assortment. These conditions lead to noisy, intermittent demand signals that are known to affect neural forecasting models [4, 17, 29].

In this study, we conducted a large-scale evaluation of forecasting models in the context of B&M retail, using real-world data from a major retailer in Southeast Europe (SEE). Our primary goal was to systematically compare the performance of diverse modeling approaches, including statistical baselines, tree-based ensembles, and state-of-the-art neural architectures, in providing daily sales predictions for a horizon of 26 weeks (182 days) into the future for products within the hygiene category. A key dimension of our comparison involves contrasting local modeling strategies, where separate models are trained for individual product groups, against global approaches that train a single model across all product groups jointly. This daily forecasting resolution was strategically chosen to capture swift changes in sales often driven by intra-week promotional activities and pricing changes, which occur within the retailer's planning cycle. For evaluation purposes, the daily predictions were subsequently aggregated to a weekly basis to assess forecast accuracy across the 26-week horizon.

The models deployed by the research team are benchmarked for their ability to handle operational realities of physical retail, including intermittent demand, missing values, and product censorship due to assortment shifts and other changes in the retail environment. Our work builds on lessons from academic literature and industrial benchmarks, including the Rossmann case study [24] and the M5 competition [22], and provides new insights into model performance evaluation under real-world retail constraints.

Our main contributions are threefold: we offer practical guidance for retail practitioners through an end-to-end modeling pipeline for long-horizon retail forecasting encompassing data preprocessing, feature engineering, and model training strategies; we provide comprehensive empirical comparisons of state-of-the-art forecasting models across multiple dimensions including model architecture (statistical, tree-based, neural), modeling scope (local vs. global), and data preprocessing strategies (imputed vs. nonimputed); and we demonstrate that localized tree-based models consistently outperform global neural architectures in brick-and-mortar retail settings while highlighting the limitations of neural models under operational constraints typical of physical retail.

2. Methodology and data

2.1. Data description

The dataset includes daily retail sales information with multiple dimensions, including time of transaction, store characteristics, product attributes, prices of sold products, promotional activities, and inventory changes. Each observation represents a daily record for a specific product-store combination, making the dataset well suited for longitudinal analysis across multiple hierarchical levels.

The area studied consists of geographies that have different levels of importance in terms of their contributions to revenue, market size, and overall positioning. The data provided includes the identification of the zones for the purpose of business strategy. Products themselves follow a three-level hierarchy: individual items are first organized into categories based on their main purpose, then into subcategories called groups and within each group, related products are further clustered into units-of-need (UoNs). Our dataset consisted of one category for hygiene products, which contained 27 groups and 79 UoNs in total over 446 stores.

Other product metadata includes unit pricing (with and without tax), cost of goods sold and promotions, which are extensively captured through binary and count features indicating tactical promotions and various loyalty initiatives. Although inventory data are fully tracked, including physical stock levels, incoming deliveries, and reserved or frozen quantities, we recognize that inventory record inaccuracy remains a well-documented issue in retail systems [28].

Metric	Value
Series Count	46841
Missingness	20.933%
Coverage Ratio	78.848%
Mean Series Length	726.923
Median Series Length	794
Minimum Series Length	2
Maximum Series Length	1124
Mean Zero-Sale Days within Series	491.801
Percent of Series with at least One Zero-Sale Day	99.981%

Table 1: Time series descriptive statistics.

Table 1 summarizes the dataset characteristics. The dataset comprises 46,841 daily SKU-level time series with a mean length of 727 observations and a median of 794. The missingness rate of 20.9% and coverage ratio of 78.8% reflect typical retail operations. Nearly all series (99.9%) contain at least one zero-sale day, with an average of 492 zero-sale days per series, indicating highly intermittent demand patterns. These characteristics pose challenges for traditional forecasting approaches designed for continuous demand patterns.

The dataset structure and challenges are comparable to those reported in major demand forecasting benchmarks such as the M5 Competition [22] or Rossmann study case [24]. However, unlike e-commerce datasets, which benefit from denser demand signals, B&M retail environments are more exposed to localized promotions and physical differences across stores.

2.2. Data preprocessing and feature selection

To ensure high-quality inputs for forecasting, a structured data preprocessing pipeline was implemented, including enrichment, imputation, transformation, and formatting steps appropriately adjusted for time series models.

Data loading and standardization. The preprocessing pipeline begins by loading raw retail data and applying systematic column renaming and standardization to ensure consistency. Invalid entries are cleaned, and leaky or unstable columns, particularly inventory bookkeeping fields, are removed to prevent data leakage. Time-based and categorical features are constructed, including calendar effects and indicators for holidays, stock availability, promotional activity, and pension-day flags, the latter being particularly relevant for understanding demand patterns in retail contexts. A critical step in this phase involves discarding products that were unavailable or out of stock during the observation period. This filtering eliminates out-of-stock bias, which is essential for sales forecasting problems [14, 9], as it ensures the model learns from genuine demand patterns rather than supply constraints. The target variable is clipped and sanitized to remove outliers and implausible values.

Competitive and macroeconomic environment. The point-of-sale, store-specific data were augmented with relevant competitive and macroeconomic indicator data. In order to control for the level of competition in the relevant geographic market, the count of competitors within a 1-kilometer radius and their prices for comparable items were added to the dataset. Further, macroeconomic environment variables from the National Statistical Office, such as Consumer Price Index (CPI), average salaries (national and regional), and population estimates within specific geographies, were included in order to control for the purchasing power and the size of the market.

Handling missing values. Our dataset is structured to include an entry for every SKU on every day when it is supposed to be in stores, regardless of whether a sale occurred. This means that a zero value explicitly indicates no sales, while missing values represent genuinely unavailable data for auxiliary variables (e.g., competitor prices), not the sales themselves. The pipeline addresses missing covariate values through a multi-stage approach: horizontal imputation across related features, forward and backward fill for time-continuous numeric variables, and group-level imputation based on store or item characteristics. These strategies ensure temporal consistency and are only applied to static features(e.g., item metadata) or ones that are rarely changed or occur in a predictable manner (e.g., close-date prices, costs, macro indicators).

To explore whether more sophisticated methods could enhance model performance, we conducted a separate experiment using a deep learning imputation model. Specifically, we trained and applied the SAITS model [13], implemented via the PyPOTS library [12], to impute missing values for the target and other variables in the training and validation sets.

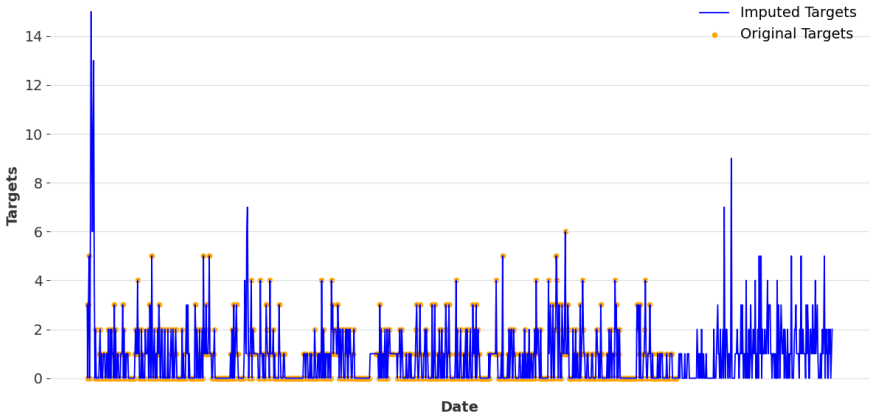


Figure 1: Sales Values Real and Imputed.

To rigorously assess imputation quality, we conducted a three-part analysis comparing SAITS with simpler alternatives. First, we examined distributional differences between original

and imputed values. SAITS-imputed values exhibit significantly different distributions from originals (all Kolmogorov-Smirnov p-values ≈ 0), with compressed variance and capped maximum values. For example, in product one group, imputed values show a standard deviation of 0.53 versus 5.42 for originals, with maximum values of 23 versus 353.

Second, we compared SAITS against simpler imputation methods including zero-fill, forward-fill, linear interpolation, and group-median imputation. Linear interpolation performs best for higher-volume groups, while zero-fill and group-median perform better for sparse groups. SAITS produces values weakly correlated with simple methods ($|r| < 0.15$), suggesting it captures different underlying patterns.

Finally, we conducted a LightGBM deep dive on a random group. Top-3 feature rankings remain stable regardless of imputed data inclusion.

Feature engineering. Several engineered features were introduced to enrich the temporal signal and provide the models with interpretable and context-aware inputs. The pipeline first trims observations after a specified cutoff date to define the modeling window. Lag and rolling-window statistics were computed for key variables such as sales, prices, and promotional activity using rolling, expanding, and exponentially weighted windows across multiple aggregation periods, specifically weekly, monthly, and yearly. The data are restructured by temporarily indexing on timestamps, and indicators are generated for whether a store was operating on the previous and next days to capture operational continuity patterns.

In-store relative positioning of products within the same Unit of Need (UoN) was captured through metrics like minimum and maximum price among comparable SKUs. All nominal prices were converted into real prices using CPI-adjusted values to control for inflation. Baseline prices for each item were established by extracting standard zone reference prices, filling them across stores, and adjusting for inflation. From these baselines, pricing-strategy features were derived, including promotion discounts, price multipliers, and baseline-relative indices. Relative pricing indicators were also constructed to measure the retailer's price advantage or disadvantage compared to competitors (out of store).

To prevent data leakage and ensure realistic forecasting conditions, all covariates were explicitly classified as either past-only or future-known variables. Past-only covariates include variables that are only observable up to the forecast origin and cannot be reliably predicted into the future: historical sales and derived features (lags, rolling statistics), competitor prices and counts (observable only historically), stock availability flags, etc. Future-known covariates include variables that are either constant, deterministic, or under direct control of the retailer over the forecast horizon: calendar features (day of week, month, holidays, pension days), store characteristics and metadata (location, size, zone), CPI projections (using forward-filled values from the most recent observation or exponential time smoothing for longer periods), planned promotional flags (directly controlled by the scenario simulation) and baseline and planned prices (inflation-adjusted, under the simulator's control).

For variables that exhibit both historical variability and forward-looking characteristics, such as competitor prices and certain inventory-related features, we applied controlled noise injection to smooth potential forward-looking signals that might leak into the training data. Specifically, competitor prices were treated as past-only during forecast generation, with no assumptions made about their future values. The pipeline enforces consistent item and store metadata throughout the temporal window. Revenue gaps identified in the historical period were repaired using median-based imputation, but no imputation was applied to future periods where values are genuinely unknown. This explicit separation ensures that both tree-based ensemble models and neural network architectures are trained and evaluated under identical, leakage-free conditions.

Feature selection. Feature selection was conducted in a single round early in the development process and remained fixed afterward. Core variables grounded in economic theory, such as prices, promotions, calendar effects, and competitor signals, were retained by default.

To reduce model complexity and enhance generalization, a filtering procedure was applied to additional engineered features based on their predictive utility. We used the Boruta algorithm [19], with LightGBM as the background estimator, to identify and keep only features with statistically verified predictive relevance. Features with excessive missingness, low variance, or multicollinearity were discarded to enhance efficiency and stability.

Train-calidation split. A time-based cutoff was used to divide the dataset into training and validation periods, and all series had observations in both periods.

Categorical encoding. Categorical variables were encoded using the CatBoost encoder [11] from the package `category_encoders`.

Final time-series transformation. For neural forecasting models, data were organized into time-series objects by grouping over product and store (SKU) identifiers. Target and numerical covariates were normalized, and to ensure series continuity, missing data points were simply interpolated and endpoints were replaced with forward and backward fills. Appropriate masks were created so that models are not trained on these points. For scalability sake, numerical fields were downcast to lighter data types to optimize memory usage during model training on large datasets.

2.3. Overview of forecasting models and implementation

This study evaluates the performance of a range of forecasting models commonly applied in time-series prediction. These models include tree-based ensembles, neural networks, and classical statistical approaches.

All neural network models were implemented using the `Darts` library, which supports modular time-series architectures and standardized training workflows. Each model was trained across four experimental setups: (1) training one model for each product group independently using nonimputed data; (2) training on each product group using imputed data; (3) training across all product groups jointly on nonimputed data; and (4) training jointly using the imputed dataset. This design allowed us to assess the impacts of SAITS-based imputation and group-level modeling on sales forecasting results.

To optimize model performance, a hyperparameter search was conducted using the HEBO algorithm [10], a scalable Bayesian optimization method. Due to computational constraints, hyperparameter tuning was conducted separately on a randomly sampled 10% subset for both nonimputed and imputed data configurations. This ensured that each experimental condition received appropriately optimized hyperparameters rather than applying configurations tuned on one data type to another.

2.3.1. Training and evaluation

The models were primarily trained using a Poisson likelihood objective to predict daily sales volumes, with the exception of the Temporal Fusion Transformer model, which was optimized using a Gaussian likelihood (a configuration determined through HEBO).

Forecasting strategy: Two evaluation protocols were employed, each reflecting standard deployment practices for its respective model class. For tree-based models, a one-step-ahead rolling protocol was used: for each day t, the model receives all information available up to t and predicts day $t+1$. This is consistent with their treatment of each observation independently and their widespread success in competitions [22, 15]. Neural models use a direct multi-horizon strategy, generating all 182 daily forecasts in a single forward pass, consistent with how architectures like N-BEATS, NHiTS, etc. are designed to learn sequential dependencies and cross-horizon patterns. To verify that this protocol difference did not disadvantage neural models, we also evaluated neural networks using rolling-origin prediction, matching the tree-based protocol. Results showed that neural networks with rolling-origin evaluation performed equally

or worse than fixed-origin multi-horizon prediction, confirming that the observed performance differences reflect genuine model capabilities rather than evaluation protocol asymmetry.

Forecast horizon and aggregation: Each model generated one sales forecast per SKU per day over a 26-week horizon. This daily resolution was intentionally chosen to accurately capture high-frequency sales volatility resulting from intra-week promotional and pricing changes that align with the retailer's operational planning cycle. The daily forecasts were subsequently aggregated to weekly totals, and performance metrics were computed at this weekly level to facilitate interpretation and align with business planning requirements.

In addition to standard forecasting evaluation metrics, we measured financial performance by computing demand error and bias [18]:

$$D_{T,h} = \frac{\sum_i \sum_{T=t+1}^{t+h} b_i(\hat{q}_{i,T} - q_{i,T})^2}{\sum_i \sum_{T=t+1}^{t+h} b_i q_{i,T}^2} \tag{1}$$

$$B_{T,h} = \frac{\sum_i \sum_{T=t+1}^{t+h} b_i(\hat{q}_{i,T} - q_{i,T})}{\sum_i \sum_{T=t+1}^{t+h} b_i q_{i,T}} \tag{2}$$

where b_i is the price of SKU and $\hat{q}_{i,T}, q_{i,T}$ are the predicted and actual sales values at a certain date.

For experiments involving imputed data, evaluation was performed exclusively on the original (nonimputed) validation data points to ensure a fair comparison across models.

All experiments were executed on a workstation equipped with an AMD Ryzen Threadripper PRO 7985WX 64-core CPU, 256 GB of RAM, and an NVIDIA RTX 4080 GPUs.

3. Results and discussion

3.1. Results

The results demonstrate clear superiority of gradient boosting methods (XGBoost and LightGBM) over deep learning approaches across most scenarios. In Case A (Individual Groups), XGBoost achieves the best performance with an RMSE of 4.833 and an MAE of 1.935, closely followed by LightGBM. The scale-independent metrics (RMSSE of 0.593 and MASE of 0.732) indicate significant error reduction compared to naive baselines. This pattern suggests that tree-based models' ability to capture complex feature interactions and handle heterogeneous data provides significant advantages for this retail forecasting problem.

Comparing Case A to Case B reveals minimal performance degradation for gradient boosting models, with XGBoost and LightGBM maintaining strong performance. However, deep learning models show deterioration, especially TFT, suggesting difficulty in capturing diverse patterns across the broader category scope. The introduction of SAITS-based imputation (Cases C and D) produces mixed results. In Case C, LightGBM experiences dramatic performance degradation (RMSE increases from 4.847 to 10.758), while XGBoost demonstrates greater resilience (RMSE of 6.027), possibly reflecting its superior regularization strategies. Case D shows improved performance over Case C for almost all models, with deep learning approaches becoming competitive, though XGBoost still maintains an edge.

The bias metrics reveal important systematic tendencies. Ensemble-based models exhibit slight negative bias (underestimation), while neural networks exhibit systematic overestimation. Case D achieves the most balanced predictions across models, with LGBM showing minimal bias (ME of -0.077, RB of -0.020). The imputation process significantly affects bias patterns, with Case C showing LightGBM shifting to severe positive bias (Demand Bias of 0.922), while Case D demonstrates that category-level imputation better preserves bias characteristics.

Model	RMSE	MAE	Demand Error	Demand Bias	RMSSE	MASE	ME	RB
Baseline								
Mean	12.022	4.947	1.211	**-0.028**	1.476	1.871	-0.500	-0.136
Theta	27.128	6.779	2.854	-0.356	3.331	2.563	-1.610	-0.440
ETS	26.027	6.355	2.585	-0.291	4.369	2.403	-1.331	-0.367
CSBA	34.972	7.989	1.072	0.084	1.795	3.648	0.419	0.314
Case A: Individual Groups								
LGBM	4.847	1.958	0.484	-0.040	0.595	0.740	-0.205	-0.056
NB	13.629	5.132	1.133	0.775	1.482	1.729	3.270	0.736
NH	12.764	4.575	1.254	0.948	1.537	1.712	3.252	0.849
TFT	7.773	3.049	0.753	0.220	0.945	1.130	0.673	0.175
XGB	**4.833**	**1.935**	**0.481**	-0.085	**0.593**	**0.732**	-0.356	-0.097
Case B: Whole Category								
LGBM	4.949	2.033	0.483	-0.105	0.601	0.754	-0.450	-0.117
NB	13.744	6.018	1.345	0.387	1.670	2.231	0.910	0.236
NH	13.744	6.018	1.345	0.387	1.670	2.231	0.910	0.236
TFT	14.060	6.087	1.357	0.398	1.708	2.257	1.130	0.294
XGB	5.045	2.041	0.493	-0.122	0.613	0.757	-0.506	-0.132
Case C: Individual Groups, Imputed Train Data								
LGBM	10.758	5.036	1.117	0.922	1.307	1.867	2.578	0.686
NB	15.362	7.704	0.628	0.065	1.867	2.856	1.210	0.085
NH	16.352	7.980	0.675	0.122	1.987	2.959	1.846	0.130
TFT	15.579	8.274	0.626	0.074	1.893	3.068	1.265	0.089
XGB	6.027	2.310	0.616	-0.224	0.732	0.856	-0.741	-0.197
Case D: Whole category, Imputed Train Data								
LGBM	7.932	3.450	0.769	-0.053	0.964	1.279	**-0.077**	**-0.020**
NB	7.853	2.928	0.773	0.067	0.954	1.086	0.263	0.069
NH	7.789	2.894	0.773	0.069	0.946	1.073	0.272	0.072
TFT	7.134	2.971	0.686	0.130	0.867	1.102	0.516	0.137
XGB	6.107	2.349	0.620	-0.307	0.742	0.871	-1.106	-0.294

Table 2: Summary statistics.

Abbreviations: ETS: Exponential Time Smoothing; CSBA: Croston SBA; TFT: Temporal Fusion Transformer; NH: NHiTS; NB: N-BEATS; LGBM: LightGBM; XGB: XGBoost; RB: Relative Bias; best results are bolded and second best underlined.

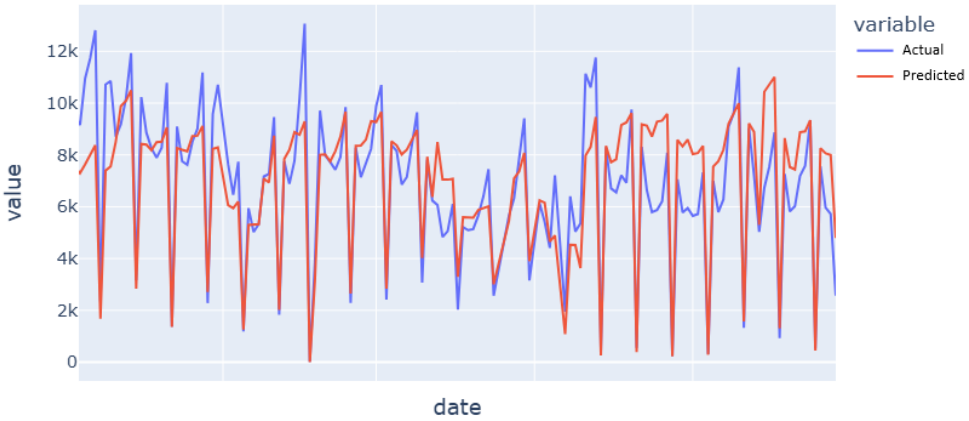
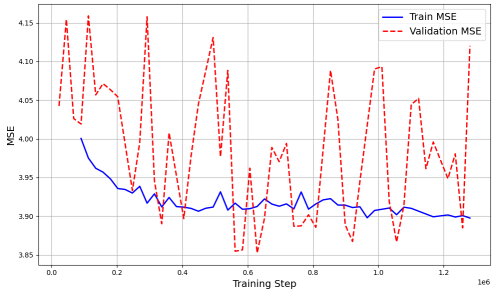
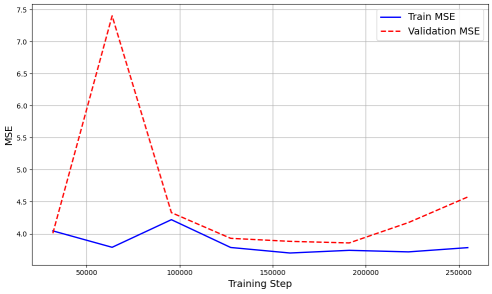


Figure 2: Example of aggregated predicted vs actual sales on a single group.

Beyond predictive accuracy, computational efficiency is another key factor used to evaluate forecasting models for practical use. Table 3 provides a comparison of training times in the



(a) N-BEATS (WC, imp): train vs. validation MSE.



(b) TFT (WC, raw): train vs. validation MSE.

Figure 3: Comparison of training and validation MSE for N-BEATS and TFT models.

	N-BEATS	NHiTS	TFT	LGBM	XGB
Training Time	146.715	145.928	4424.972	35.566	14.542

Table 3: Training time statistics (in minutes).
Note: per epoch for neural networks, total for LGBM and XGB.

whole-category setting for SAITS-imputed data. The results make it evident that ensemble approaches, especially XGBoost, are highly efficient: they require considerably less memory and train much faster than neural network approaches. For real-world applications where speed and accuracy matter, the lower computational footprint of the ensemble methods represents a clear advantage.

3.2. Discussion

The results offer several important insights into retail sales forecasting. Ensemble models such as LightGBM and XGBoost exhibit robust performance in both localized and aggregated configurations when they use nonimputed data, consistently achieving lower error rates while demonstrating significant training time advantages. This aligns with Borisov et al. [3], who note that gradient-boosted decision tree ensembles often remain state-of-the-art for heterogeneous tabular data due to their robustness to data irregularities.

A key takeaway from our study is the importance of tailoring the modeling approach to the characteristics of the dataset. When data are segmented into individual groups and they do not share too much information between them, localized modeling can capture unique patterns more effectively. In our case, training on the whole category does not introduce enough inter-group information to bring improvements, a phenomenon that resonates with the observations by McElfresh et al. [23], who note that differences in dataset properties, such as skewed feature distributions and irregularities, can diminish the relative advantage of complex neural network architectures over simpler, well-tuned tree-based methods.

Furthermore, experiments with imputed versus nonimputed data underscore the critical role of data quality. Our analysis of SAITS imputation quality revealed that imputed values exhibit compressed variance and substantially different distributions from originals. Deep neural networks, as discussed by Ramesh and Usman [6], tend to be more sensitive to inherent noise and missing values in tabular datasets. Our results extend this understanding by demonstrating that imputation-introduced artifacts also affect tree-based models, particularly the dramatic LightGBM performance collapse in Case C (RMSE increasing from 4.847 to 10.758). Detailed feature importance analysis revealed that while core predictors remain stable, the inclusion of imputed data shifts the relative importance of spatial features and introduces mild noise that degrades tree model performance.

Criterion	Statistical	Tree-Based	Neural
Best for	Stable, low-volume series	Heterogeneous, feature-rich data	Large-scale, dense data
Intermittent demand	Croston variants	Strong (handles zeros well)	Sensitive to sparsity
Feature handling	Limited	Excellent	Requires engineering
Training time	Minutes	Minutes–hours	Hours–days
Interpretability	High	Medium (SHAP)	Low
Data requirements	Low	Medium	High
B&M retail fit	Moderate	**Strong**	Weak–Moderate

Table 4: Practitioner guidance: Model selection by use case.

While recent work by Zalando [18] has demonstrated the presence of scaling laws for transformer-based models, where forecasting accuracy improves as training data size increases, such advantages were not observed in our experiments. This is likely due to the comparatively smaller scale of our dataset and the absence of centralized, high-density demand signals typical of e-commerce platforms. As a result, we did not observe a trade-off between computational cost and accuracy in favor of deep learning models. In our setting, ensemble methods remained both more efficient and more accurate, reinforcing their suitability for B&M retail forecasting tasks characterized by fragmentation, intermittency, and limited historical coverage.

4. Conclusion

This study demonstrates that ensemble methods, particularly XGBoost and LightGBM, substantially outperform neural network architectures in B&M retail forecasting when applied to localized, individual-group data. XGBoost achieved the best overall performance with an RMSE of 4.833, representing a significant improvement over naive baselines. Neural architectures exhibited sensitivity to intermittency and aggregation level; while SAITS-based imputation improved their performance in the whole-category setting, they still fell short of ensemble methods. Our imputation quality analysis revealed that SAITS introduces distributional compression and correlation shifts that particularly affect tree models in localized settings, explaining the dramatic LightGBM performance degradation in Case C.

For practitioners, these findings suggest that a localized modeling strategy leveraging tree-based ensembles ensures scalability, efficiency, and improved forecast accuracy, with direct applications in inventory management and demand planning. The superior computational efficiency of these methods makes them particularly suitable for real-time forecasting scenarios.

Several limitations warrant discussion. Aggregated error metrics may obscure performance variability across segments, and findings may not generalize to all retail environments. Future research should investigate hybrid approaches integrating tree-based efficiency with neural representation learning, and explore imputation methods that better preserve the underlying data structure. This work provides a systematic performance analysis under realistic conditions and offers actionable guidance for practitioners, reinforcing the importance of matching model complexity to problem characteristics.

Statement on the use of artificial intelligence

The Large Language Models were used in order to improve phrasing and clarity of the text. All ideas, arguments, analysis, and conclusions are the original work of the authors.

References

[1] Amazon Science (2021). *The history of Amazon's forecasting algorithm*. url: https://www.amazon.science/latest-news/the-history-of-amazons-forecasting-algorithm [Accessed 03/4/2025]

[2] Assimakopoulos, V. and Nikolopoulos, K. (2000). The theta model: a decomposition approach to forecasting. International Journal of Forecasting, 16(4), 521–530. doi: 10.1016/S0169-2070(00)00066-2

[3] Borisov, V., Leemann, T., Seßler, K., Haug, J., Pawelczyk, M., and Kasneci, G. (2024). Deep Neural Networks and Tabular Data: A Survey. IEEE Transactions on Neural Networks and Learning Systems, 35(6), 7499–7519. doi: 10.1109/TNNLS.2022.3229161

[4] Boylan, J. E. and Syntetos, A. A. (2021). *Intermittent demand forecasting: Context, methods and applications*. John Wiley & Sons. doi: 10.1002/9781119135289

[5] Challu, C., Olivares, K. G., Oreshkin, B. N., Garza, F., Mergenthaler-Canseco, M., and Dubrawski, A. (2022). N-HiTS: Neural Hierarchical Interpolation for Time Series Forecasting. doi: 10.1609/aaai.v37i6.25854

[6] Changa Ramesh, A. and Usman, M. (2024). *Empirical study of neural network and other machine learning models in time series forecasting using sales data*. Uppsala University, Master's Thesis.

[7] Chen, J., Yan, J., Chen, Q., Chen, D. Z., Wu, J., and Sun, J. (2024). ExcelFormer: A neural network surpassing GBDTs on tabular data. doi: 10.48550/arXiv.2301.02819

[8] Chen, T. and Guestrin, C. (Aug. 2016). XGBoost: A Scalable Tree Boosting System. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. ACM, 785–794. doi: 10.1145/2939672.2939785

[9] Chopra, S. and Meindl, P. (2019). Supply chain management: Strategy, planning & operation. In: Das Summa Summarum des Management: Die 25 wichtigsten Werke für Strategie, Führung und Veränderung (265-275). Springer. doi: 10.1108/02656710310461350

[10] Cowen-Rivers, A. I., Lyu, W., Tutunov, R., Wang, Z., Grosnit, A., Griffiths, R. R., Maraval, A. M., Jianye, H., Wang, J., Peters, J., and Ammar, H. B. (2022). HEBO Pushing The Limits of Sample-Efficient Hyperparameter Optimisation. doi: 10.48550/arXiv.2012.03826

[11] Dorogush, A. V., Gulin, A., Gusev, G., Kazeev, N., Prokhorenkova, L. O., and Vorobev, A. (2017). CatBoost: Unbiased Boosting with Categorical Features. doi: 10.48550/arXiv.1706.09516

[12] Du, W. (2023). PyPOTS: a Python toolbox for data mining on Partially-Observed Time Series. doi: 10.48550/arXiv.2305.18811

[13] Du, W., Côté, D., and Liu, Y. (June 2023). SAITS: Self-Attention-based Imputation for Time Series. Expert Systems with Applications, 219, 119619. doi: 10.1016/j.eswa.2023.119619

[14] Hyndman, R. J. and Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts. doi: 10.2307/1054108

[15] Januschowski, T., Wang, Y., Torkkola, K., Erkkilä, T., Hasson, H., and Gasthaus, J. (2022). Forecasting with trees. International Journal of Forecasting, 38(4), 1473–1481. doi: 10.1016/j.ijforecast.2021.10.004

[16] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In: *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*.

[17] Kourentzes, N. (2013). Intermittent demand forecasts with neural networks. International Journal of Production Economics, 143(1), 198–206. ISSN: 0925-5273. doi: 10.1016/j.ijpe.2013.01.009

[18] Kunz, M., Birr, S., Raslan, M., Ma, L., Li, Z., Gouttes, A., Koren, M., Naghibi, T., Stephan, J., Bulycheva, M., Grzeschik, M., Kekić, A., Narodovitch, M., Rasul, K., Sieber, J., and Januschowski, T. (2023). Deep Learning based Forecasting: a case study from the online fashion industry. doi: 10.1007/978-3-031-35879-1_11

[19] Kursa, M., Jankowski, A., and Rudnicki, W. (Jan. 2010). Boruta - A System for Feature Selection. Fundam. Inform., 101, 271–285. doi: 10.3233/FI-2010-288

[20] Lim, B., Arik, S. O., Loeff, N., and Pfister, T. (2020). Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting, 37(4). doi: 10.1016/j.ijforecast.2021.03.012

[21] Luo, D. and Wang, X. (2024). ModernTCN: A Modern Pure Convolution Structure for General Time Series Analysis. In: *The Twelfth International Conference on Learning Representations*. doi: 10.1007/978-3-540-73291-4

[22] Makridakis, S., Spiliotis, E., and Assimakopoulos, V. (2022). M5 accuracy competition: Results, findings, and conclusions. International Journal of Forecasting, 38(4), 1346–1364. doi: 10.1016/j.ijforecast.2021.11.013

[23] McElfresh, D., Khandagale, S., Valverde, J., C, V. P., Feuer, B., Hegde, C., Ramakrishnan, G., Goldblum, M., and White, C. (2024). When Do Neural Nets Outperform Boosted Trees on Tabular Data? doi: 10.1088/1742-5468/ac3a81

[24] Middel, C. and Davel, M. (2023). Comparing Transformer-based and GBDT models on tabular data: A Rossmann Store Sales case study. In: *Southern African Conference for Artificial Intelligence Research (SACAIR)*. doi: 10.1109/icemme51517.2020.00110

[25] Nie, T., Qin, G., Ma, W., Mei, Y., and Sun, J. (2024). ImputeFormer: Low Rankness-Induced Transformers for Generalizable Spatiotemporal Imputation. In: *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. KDD '24. Barcelona, Spain: Association for Computing Machinery, 2260–2271. ISBN: 9798400704901. doi: 10.1145/3637528.3671751

[26] Nie, Y., Nguyen, N. H., Sinthong, P., and Kalagnanam, J. (2023). A Time Series is Worth 64 Words: Long-term Forecasting with Transformers. In: *The Eleventh International Conference on Learning Representations*. doi: 10.48550/arXiv.2211.14730

[27] Oreshkin, B. N., Carpov, D., Chapados, N., and Bengio, Y. (2020). N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. In: *International Conference on Learning Representations*. doi: 10.48550/arXiv.1905.10437

[28] Rekik, Y., Oliva, R., Glock, C., and Syntetos, A. (2025). Inventory record inaccuracy in grocery retailing: Impact of promotions and product perishability, and targeted effect of audits. doi: 10.48550/arXiv.2506.05357

[29] Weerakody, P. B., Wong, K. W., Wang, G., and Ela, W. (2021). A review of irregular time series data handling with gated recurrent neural networks. Neurocomputing, 441, 161–178. doi: 10.1016/j.neucom.2021.02.046

[30] Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. Management science, 6(3), 324–342. doi: 10.1007/978-3-642-51565-1_116

Croatian Operational Research Review
CRORR **17:2**(2026), 297–308

Environmental fiscal policy and greenhouse gas emissions in the EU-27: Evidence from a dynamic panel GMM model

Irena Palić

[1] *Faculty of Economics and Business, University of Zagreb, Trg J.F. Kennedy 6, 10000 Zagreb, Croatia*
E-mail: ⟨ipalic@efzg.hr⟩

Abstract. This paper examines the effects of environmental fiscal policy variables on greenhouse gas (GHG) emissions in the EU-27 from 2009 to 2023. The Arellano and Bond difference GMM estimator is used to address potential endogeneity and persistence of GHG emissions. The panel model is estimated using the total GHG emissions as the dependent variable, while energy taxes, transport taxes, and government environmental protection expenditure serve as independent variables. The results indicate significant persistence in GHG emissions. Energy taxes and government environmental protection expenditure have a negative and statistically significant effect on emissions in both the short run and the long run, confirming the effectiveness of these fiscal instruments in reducing emissions. However, transport taxes show a positive and significant relationship with emissions, suggesting that their current structure, which is linked primarily to vehicle ownership rather than fuel consumption, is not effective in reducing emissions. Robustness checks using population as a control variable and CO_2 emissions as an alternative dependent variable confirm the stability of these results. The findings suggest that energy taxes and environmental expenditures are effective fiscal instruments for reducing emissions in the EU, while the design of transport taxes requires reconsideration to improve their environmental effectiveness.

Keywords: dynamic panel GMM, environmental expenditures, environmental taxation, european union, greenhouse gas emissions

Received: March 6, 2026; accepted: April 9, 2026; available online: May 13, 2026

DOI: 10.17535/crorr.2026.0022

Original scientific paper.

1. Introduction

Over the past few decades, increased living standards have improved well-being but have also placed great pressure on natural ecosystems, resulting in increased CO_2 emissions, biodiversity loss, and environmental imbalances. Balancing economic growth with environmental sustainability remains an economic policy concern [6]. Handoyo [22] points out that effective public governance plays a key role in shaping national environmental outcomes and determining how well governments respond to environmental challenges. According to Batini et al. [5], spending that helps reduce greenhouse gas emissions can significantly boost economic growth while supporting environmental sustainability.

While fiscal measures can promote employment and production, their environmental consequences differ due to differences in resource allocation and the incorporation of sustainability objectives. Consequently, governments often face the challenge of creating fiscal strategies that support economic growth without threatening environmental sustainability. Environmental outcomes largely depend on the extent to which fiscal frameworks support clean technologies, energy efficiency, and green innovation [7]. Consequently, integrating sustainable environmental

objectives into the global policy framework has become one of the priorities, together with economic development and social equity. According to the Kunming–Montreal Global Biodiversity Framework, governments should align fiscal and financial activities with global biodiversity and environmental sustainability goals [38].

Recently, in 2023, the EU's greenhouse gas footprint amounted to 9.0 tonnes of CO_2 equivalents per capita [12]. Regarding total environmental tax revenue in the EU, it was €341.5 billion in 2023, representing 2% of EU GDP and 5.1% of total EU government revenue from taxes and social contributions. Taxes on energy accounted for more than three-quarters of total revenues from environmental taxes (76% of the total), followed by taxes on transport (19%) and pollution and resources (5%) [13]. Moreover, the total general government expenditure on environmental protection in the EU was €142 billion and accounted for 0.8% of GDP in 2023 [14].

Previous empirical studies have examined the relationship between fiscal policy and environmental quality, focusing mainly on aggregate environmental taxes and CO_2 emissions [21, 37, 29]. This paper contributes to previous research in several ways. Firstly, it examines both the revenue side (environmental taxes) and expenditure side (government environmental protection expenditure) of fiscal policy, providing a more comprehensive view of fiscal instruments. Secondly, this research disaggregates tax revenues into energy, transport, and pollution/resource taxes. Thirdly, this research uses an overall measure, namely total greenhouse gas emissions, which includes CO_2, CH_4, N_2O, and F-gases, rather than focusing solely on CO_2 emissions, unlike the common literature [21, 37, 29]. While CO_2 from fossil fuels remains the dominant greenhouse gas, accounting for 73.7% of global GHG emissions in 2023, the remaining share is also important: methane contributes 18.9%, nitrous oxide 4.7%, and fluorinated gases 2.7% [25]. Importantly, non-CO_2 gases are increasing rapidly. Between 1990 and 2023, F-gas emissions increased by 294%, CH_4 by 28.2%, and N_2O by 32.4% [25]. Furthermore, the European Climate Law sets binding targets for reductions in total GHG emissions, not CO_2 alone [9]. To ensure comparability with the existing literature, CO_2 emissions are used as an alternative dependent variable in the robustness checks. Moreover, this research uses the most recent available data (2009–2023), capturing post-COVID recovery and recent EU climate policy developments under the European Green Deal. Finally, the dynamic panel model is estimated using the difference GMM estimator of Arellano and Bond [1] to address potential endogeneity and persistence in GHG emissions, in line with Okombi and Ndoum Babouama [33].

After the introduction, the literature review provides relevant empirical research on the role of fiscal policy in environmental sustainability, followed by the empirical part, providing data, methods, results, and discussion. Finally, the main conclusions are provided with research limitations and perspectives for future research.

2. Literature review

López et al. [30] examined how fiscal spending affects the environment in 38 selected countries and outlined the importance of the allocation of spending. When governments allocate more spending towards public goods like education, healthcare, and environmental protection, pollution tends to decrease. However, they outlined that simply increasing total government spending without changing allocation does not reduce pollution. Halkos and Paizanos [21] analysed fiscal policy impact on CO_2 emissions in the United States between 1973 and 2013. Using a Vector Autoregression model, they found that higher government spending can lower emissions from both production and consumption, while tax cuts financed by deficits increase consumption-related emissions. Their findings suggest that the environmental impact of fiscal policy depends on its design, with spending measures showing a more favourable impact than tax cuts. Savranlar et al. [37] analysed the influence of environmental taxes on CO_2 emissions in EU-27 countries using a panel vector autoregression (PVAR) approach, finding that

an increase in environmental taxes reduces CO_2 emissions by 0.14%. Leitão [29] investigated the impact of environmental taxes on CO_2 emissions in 38 OECD countries from 1995 to 2022, employing FMOLS, DOLS, quantile regression, and System GMM estimation.

Okombi and Ndoum Babouama [33] analyse the relationship between environmental taxation and inclusive green growth in developing countries from 2000 to 2021 using the System GMM estimator and outline that environmental taxes promote inclusive green growth. The study further demonstrates that institutional quality positively moderates this relationship. Kohlscheen et al. [28] analysed 121 countries using dynamic panel regressions and found that an increase in carbon taxes by \$10 per ton of CO_2 reduces CO_2 emissions per capita by 1.3% in the short run and 4.6% in the long run, providing robust evidence for the effectiveness of carbon pricing mechanisms.

Regarding single-country research, Katircioglu and Katircioglu [26] outlined that in Turkey, government spending and tax revenues are cointegrated with CO_2, and the increase in fiscal aggregates is related to lower emissions in the long-run. Moreover, Erdogan [8] empirically examines how green fiscal policy instruments, renewable energy use, and economic growth affect environmental degradation in Germany using the Autoregressive Distributed Lag (ARDL) approach. The study uses the load capacity factor as a comprehensive indicator of environmental sustainability. The findings indicate that economic growth and renewable energy consumption improve environmental quality, while environmental taxes do not significantly contribute to reducing environmental degradation. In contrast, environmental protection expenditures have a positive impact on sustainability, suggesting that expenditure-based fiscal instruments may be more effective than tax-based instruments in achieving environmental goals. Nguyen and Duong [31] analysed the effects of fiscal policy on environmental quality in Vietnam between 1990 and 2022 using a nonlinear ARDL framework and pointed to asymmetric long-run and short-run effects on CO_2 emissions. Specifically, increased government spending significantly reduces CO_2 emissions, while a contraction in public spending increases emissions. Nguyen and Duong [31] attribute the reduction in emissions to Vietnam's expansionary spending on sectors such as education, healthcare, and social welfare, which have lower carbon intensity compared to the manufacturing sector or consumption-driven service sector. In contrast, fiscal contractions are found to decrease investments in green infrastructure.

However, the effectiveness of environmental taxes may depend critically on their design. In the EU, transport taxes are ownership-based, linked to vehicle registration and possession, while taxes on transport fuels are classified under energy taxes [11]. Pigato and Hayde [34] noted that higher registration taxes may discourage new vehicle purchases, leading consumers to retain older, more polluting vehicles. Similarly, Aydin and Bozatli [3] found that transport taxes increase air pollution in countries with the highest transport tax revenues, and Ptak et al. [35] argued that the environmental impact of transport taxes can be relatively limited.

To summarise, the reviewed literature points to several gaps. Most empirical studies examine the effect of aggregate environmental taxes on CO_2 emissions, without distinguishing between tax categories that differ in design and incentive structure. The expenditure side of environmental fiscal policy remains largely underexplored, with only a few studies analysing government spending [30, 21, 8, 31]. Furthermore, under the Eurostat classification, transport taxes consist exclusively of ownership-based levies, while fuel taxes are classified under energy taxes [11]. However, this distinction has rarely been exploited in panel data studies to separately assess the environmental effectiveness of ownership-based and usage-based taxation. This research contributes to filling these gaps using disaggregated fiscal data and total GHG emissions for the EU-27.

3. Data and sample

The dependent variable is greenhouse gas (GHG) emissions, retrieved from Eurostat [15], measured in million tonnes of CO_2 equivalent. The indicator includes carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), and fluorinated gases (hydrofluorocarbons, perfluorocarbons, nitrogen trifluoride, and sulphur hexafluoride), excluding Land Use, Land-Use Change and Forestry (LULUCF), as well as emissions from international aviation and maritime transport. Using each gas's individual global warming potential (GWP), emissions are integrated into a single indicator expressed in CO_2 equivalents [15]. As a robustness check, CO_2 emissions without other emissions are also used as an alternative dependent variable [19].

The independent variables refer to both the revenue and expenditure sides of environmental fiscal policy. The revenue side is represented by transport taxes and energy taxes measured in million euros [16]. The pollution taxes were excluded from modelling due to data limitations. Three EU member states reported zero pollution tax revenues: Germany throughout the entire sample period (2009–2023), Greece from 2009–2017, and Luxembourg in 2009. Hence, including pollution taxes would exclude Germany from the analysis, introducing selection bias. Moreover, as previously mentioned, pollution taxes account for a relatively small share of total revenues from environmental taxes, which was 5% in 2023 [13]. Nevertheless, it should be acknowledged that the exclusion of pollution taxes may affect the generalisability of the findings if countries that levy higher pollution taxes differ systematically in their emission patterns. Transport taxes (TRATAX) include taxes related to the ownership and use of motor vehicles and also cover taxes on other transport equipment, such as aircraft and related transport services. Energy taxes (NRGTAX) refer to taxes on energy products used for both transport and stationary purposes. The most important energy products for transport purposes are petrol and diesel. Energy products for stationary use include fuel oils, natural gas, coal, and electricity. CO_2 taxes are included under energy taxes [16]. It should be noted that under the Eurostat classification of environmental taxes, taxes on transport fuels such as petrol and diesel are included under energy taxes, while transport taxes consist of ownership-based levies such as vehicle registration and annual vehicle taxes [11]. Therefore, the difference between NRGTAX and TRATAX in this paper relates to separation between usage-based and ownership-based taxation in the Eurostat methodology.

Concerning the expenditure side, the government expenditure on environmental protection (ENVEXP) is available at the Eurostat [17] database on general government expenditure by function. Environmental protection expenditure covers government spending on waste management, wastewater management, pollution abatement, protection of biodiversity and landscape, and other environmental protection activities [17].

Furthermore, GDP (2015=100), available at Eurostat [18], was first considered as a control variable, but due to high multicollinearity with the fiscal policy variables, it was not included in the model. The correlation coefficient between GDP and energy tax revenues is very high ($r = 0.98$), between GDP and environmental expenditure is $r = 0.96$, and for GDP and transport taxes, $r = 0.89$. For robustness check, population (POP) is included in alternative specification as control variable. Population is retrieved from Eurostat [20] and measured in thousands of persons.

All fiscal variables (NRGTAX, TRATAX, ENVEXP) are measured in million euros and deflated to real values using the implicit GDP deflator, 2015=100, available at Eurostat [18]. Table 1 provides variable description, as well as data sources with Eurostat codes.

The model estimation is conducted on a panel dataset of 27 European Union member states (EU-27) from 2009 to 2023. The descriptive statistics of data prior to logarithmic transformation is provided in Table 2.

Greenhouse gas emissions (GHG) show substantial variation across EU countries, with an average value of 138.37 million tonnes of CO_2 equivalent and a wide range from 1.90 to 934 Mt

Variable	Description	Source	Eurostat code
GHG	GHG emissions (million tonnes of CO_2 equivalent)	Eurostat [15]	sdg_13_10
CO_2	Carbon dioxide emissions (million tonnes)	Eurostat [19]	env_air_gge
NRGTAX	Energy taxes (million euro)	Eurostat [16]	env_ac_tax
TRATAX	Transport taxes (million euro)	Eurostat [16]	env_ac_tax
ENVEXP	General government expenditure on environmental protection (million euro)	Eurostat [17]	gov_10a_exp
POP	Population on 1 January (thousands of persons)	Eurostat [20]	demo_gind

Table 1: *Variable description and data sources.*

Variable	Mean	Std. Dev.	Min	Max	CV (%)
GHG	138.37	192.17	1.90	934.00	138.9
CO_2	109.85	155.22	1.40	762.00	141.3
NRGTAX	9,542.10	15,842.17	86.62	75,195	166.0
TRATAX	2,117.31	2,874.75	6.42	11,250	135.8
ENVEXP	3,858.94	6,038.64	-38.10	28,936	156.5
POP	16,400,000	21,700,000	410,926	83,200,000	132.3

Table 2: *Descriptive statistics of selected variables.*

CO_2e. This reflects differences in the size and industrial structure of EU economies. Energy tax revenues (NRGTAX) have the highest average among fiscal variables (€9.5 billion) and exhibit considerable variability across countries, while transport tax revenues (TRATAX) are substantially lower on average (€2.1 billion), reflecting the smaller share of transport taxes in overall environmental tax revenues in the EU. Environmental protection expenditure (ENVEXP) averages €3.9 billion, but also shows notable variation across countries. High coefficients of variation point to heterogeneity in emissions levels, environmental taxation and environmental spending, as well as population across EU member states. Moreover, the minimum value of ENVEXP (−38.10 million euros) which corresponds to Estonia in 2010, is negative. In the ESA 2010 accounting framework [10], negative values in government expenditure by function can occur due to netting in the transaction classifications.

For model estimation, all variables are transformed using natural logarithms, the first-differenced log variables represent approximate growth rates, and the estimated coefficients are interpreted as elasticities. Table 3 shows the results of the Im–Pesaran–Shin (IPS) panel unit root test. The IPS test examines the null hypothesis that all panel series contain a unit root [24].

Variable	IPS (level)	p-value	IPS (1st diff.)	p-value
LNGHG	5.368	1.000	-8.510^{***}	0.000
$LNCO_2$	−0.795	1.000	-8.636^{***}	0.000
LNNRGTAX	0.257	0.601	-8.215^{***}	0.000
LNTRATAX	−0.455	0.325	-7.203^{***}	0.000
LNENVEXP	−1.038	0.150	-7.927^{***}	0.000
LNPOP	−0.413	1.000	−0.767	0.221

Table 3: *Panel unit root tests (Im–Pesaran–Shin). Note:* *** $p < 0.01$.

All variables except LNPOP are non-stationary in levels but stationary in first differences,

i.e. integrated of order one, I(1). Population, which is not stationary in first differences, is included as an exogenous control variable for robustness check of the baseline model. The first-difference GMM estimator is used in model estimation, which is explained in the next section.

4. The empirical model

In line with the dynamic panel data literature [4], the following dynamic model is estimated as a baseline model:

$$y_{i,t} = \rho y_{i,t-1} + \mathbf{X}'_{i,t}\boldsymbol{\beta} + \alpha_i + \varepsilon_{i,t}, \quad |\rho| < 1, \tag{1}$$

where $y_{i,t} = \text{LNGHG}_{i,t}$ denotes the natural logarithm of greenhouse gas emissions for country i at time t, $\mathbf{X}_{i,t}$ is a vector of explanatory variables (LNNRGTAX, LNTRATAX, LNENVEXP), α_i represents time-invariant country-specific fixed effects, and $\varepsilon_{i,t}$ is the idiosyncratic error term with $E(\varepsilon_{i,t}) = 0$ and $E(\varepsilon_{i,t}\varepsilon_{j,s}) = 0$ for $i \neq j$ or $t \neq s$. The fiscal policy variables (energy taxes, transport taxes, and environmental protection expenditure) are treated as endogenous, as fiscal revenues and expenditures may simultaneously impact and be impacted by emission levels.

Model (1) is the baseline specification with total GHG emissions as the dependent variable. To assess the robustness of the baseline results, three alternative model specifications are estimated. Model (2) extends the baseline by including the total population as a control variable:

$$y_{i,t} = \rho y_{i,t-1} + \mathbf{X}'_{i,t}\boldsymbol{\beta} + \gamma \ln(\text{POP})_{i,t} + \alpha_i + \varepsilon_{i,t}. \tag{2}$$

Model (3) uses CO_2 emissions instead of total GHG emissions as the dependent variable, with the same set of explanatory variables as in Model (1), while Model (4) uses CO_2 emissions as the dependent variable and includes population as a control variable. Table 4 presents the estimation results for all four models.

The fiscal policy variables (LNNRGTAX, LNTRATAX, LNENVEXP) and the lagged dependent variable are treated as endogenous and instrumented using their own lagged levels from periods $t-2$ to $t-4$. Population, when included, is treated as strictly exogenous, meaning it is assumed to be uncorrelated with the idiosyncratic error term at all leads and lags: $E[\ln(\text{POP})_{i,s} \cdot \varepsilon_{i,t}] = 0$ for all s and t. Population dynamics are related to demographic processes that are not influenced by changes in greenhouse gas emissions. In the GMM estimation framework, strictly exogenous variables instrument themselves, appearing in both the regressor and instrument matrices, rather than being instrumented by their own lags as is the case for endogenous variables [36].

In the first-differenced equation estimated by the Arellano–Bond procedure, population enters in its differenced form, serving as its own instrument without requiring additional lags. This distinction is reflected in the instrument count: 12 instruments in Models (1) and (4), and 13 instruments in Models (2) and (3), where the additional instrument corresponds to the population variable.

The presence of the lagged dependent variable $y_{i,t-1}$ creates correlation with the fixed effects α_i, hence OLS and fixed effects estimators are biased and inconsistent [32]. To address this endogeneity problem, the first-difference GMM estimator proposed by Arellano and Bond [1] is used.

First-differencing baseline model equation (1) eliminates the fixed effects:

$$\Delta y_{i,t} = \rho \Delta y_{i,t-1} + \Delta \mathbf{X}'_{i,t}\boldsymbol{\beta} + \Delta \varepsilon_{i,t}. \tag{3}$$

However, the differenced lagged dependent variable $\Delta y_{i,t-1} = (y_{i,t-1} - y_{i,t-2})$ is correlated with the differenced error $\Delta \varepsilon_{i,t} = (\varepsilon_{i,t} - \varepsilon_{i,t-1})$ through the term $y_{i,t-1}$. Arellano and Bond [1]

propose using lagged levels of the dependent variable as instruments, exploiting the following moment conditions:

$$E[y_{i,t-s}\,\Delta\varepsilon_{i,t}] = 0, \quad s \geq 2;\ t = 3, \ldots, T. \tag{4}$$

Similarly, if the explanatory variables $\mathbf{X}_{i,t}$ are potentially endogenous, their lagged values can serve as instruments:

$$E[\mathbf{X}_{i,t-s}\,\Delta\varepsilon_{i,t}] = 0, \quad s \geq 2;\ t = 3, \ldots, T. \tag{5}$$

The problem that arises in GMM estimation is instrument proliferation. Following Roodman [36], the instrument matrix is collapsed, which reduces the number of instruments while maintaining consistency. By collapsing the instrument matrix, the number of instruments is below the number of cross-sectional units (27 countries), which is in line with the recommendations of Roodman [36] and Baltagi [4]. Without collapsing, the number of instruments would exceed the number of cross-sections, potentially leading to overfitting and weakening the Hansen test of overidentifying restrictions.

The two-step GMM estimator, which uses the residuals from a first-step consistent estimator to construct an optimal weighting matrix, is used. However, two-step standard errors are known to be downward-biased in finite samples. Hence, the Windmeijer [39] finite-sample correction is used to obtain reliable inference.

The validity of the GMM estimator is assessed through the Arellano–Bond test for serial correlation and the Hansen [23] J-test. The dynamic specification allows computation of both short-run and long-run effects. The short-run elasticity is given directly by the coefficient β, while the long-run elasticity is computed as [4]:

$$\text{Long-run elasticity} = \frac{\beta}{1-\rho}. \tag{6}$$

5. Estimation results and discussion

The dynamic panel dataset comprises 27 EU member states from 2009 to 2023. After first-differencing, which eliminates the first period, and the use of lagged levels from $t-2$ as instruments, which eliminates an additional period, the total number of observations is 349. Following Roodman [36], the instrument matrix is collapsed to limit the number of instruments. Lagged levels of greenhouse gas emissions, energy taxes, transport taxes, and environmental protection expenditure from periods $t-2$ to $t-4$ are used as instruments. To avoid biased standard errors, Windmeijer-corrected standard errors are used for the two-step GMM estimator [2, 27].

As previously mentioned, to assess the robustness of the baseline model, three alternative model specifications are estimated. In addition to Model (1), Model (2) includes total population as a control variable, while Model (3) uses CO_2 emissions as the dependent variable, and Model (4) uses CO_2 emissions as the dependent variable and population as control variable. Table 4 presents the estimation results for all four models.

In the baseline model, the coefficient on the lagged dependent variable ($\rho = 0.580$) is significant with a p-value of 0.000, and points to persistence in GHG emissions. In dynamic panel models, the inclusion of a lagged dependent variable captures such dynamic behaviour and allows the estimation of both short-run and long-run effects [4]. Current emission levels are strongly influenced by past emission levels, pointing to the persistence of emissions. Thus, changes in emissions adjust slowly over time. The estimated coefficient below one also points to partial adjustment dynamics. Long-run coefficients on fiscal variables are therefore calculated in line with equation (6).

In the baseline model, the energy taxes have a significant negative effect on emissions, with a short-run elasticity of -0.39 and a long-run elasticity of -0.93. This implies that a 1% increase in energy taxes reduces emissions by 0.39% in the short run and 0.93% in the long

Variable	(1) GHG	(2) GHG + POP	(3) CO_2 + POP	(4) CO_2
Lagged dependent variable	0.580***	0.367**	0.376**	0.586***
	(0.134)	(0.143)	(0.140)	(0.126)
ln(NRGTAX)	−0.392***	−0.586***	−0.605***	−0.433***
	(0.134)	(0.157)	(0.184)	(0.150)
ln(TRATAX)	0.411**	0.508***	0.618***	0.453**
	(0.156)	(0.153)	(0.171)	(0.175)
ln(ENVEXP)	−0.371***	−0.373***	−0.445***	−0.414***
	(0.108)	(0.129)	(0.139)	(0.128)
ln(POP)	–	−0.729*	−0.673	–
		(0.385)	(0.441)	

Table 4: *Two-step difference GMM estimation results. Source: Author's calculation, Stata 15.0.*

*Notes: standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Two-step difference GMM with Windmeijer (2005) corrected standard errors.*

run, which points to the effectiveness of energy taxation as an environmental policy instrument. As noted in Section 1, energy taxes account for approximately 76% of total environmental tax revenues in the EU [13]. This empirical result emphasises the important role of energy taxes for environmental sustainability. The negative significant impact is in line with the literature on environmental taxation impact [37, 28].

However, regarding transport taxes, the coefficient is also statistically significant, but positive, indicating that a 1% increase in transport taxes increases emissions by 0.41% in the short run and 0.98% in the long run. This result is unexpected and should be interpreted with caution. In the EU, transport taxes are linked to vehicle ownership and use, rather than directly to fuel consumption. As noted in Section 3, the Eurostat classification assigns fuel taxes to the energy tax category, while transport taxes are ownership-based [11]. The positive coefficient on TRATAX, therefore, specifically reflects the effect of ownership-based taxation, while the negative coefficient on NRGTAX captures the effect of usage-based energy taxation, including taxes on transport fuels. In 2023, households accounted for about 66.5% of total transport tax revenues, pointing to the connection between transport taxes and private vehicle ownership [13]. Transport taxes mainly include vehicle registration and similar charges, which may have only a limited effect on emissions. As outlined in Section 1, transport taxes account for only about 19% of total environmental tax revenues in the EU, compared with 76% for energy taxes [13]. Their relatively small revenue share, together with the structure discussed above, may explain the finding that they do not reduce emissions. National patterns also differ across countries. In Ireland, Malta, and Austria, most transport taxes are paid by households, while in Slovakia and Czechia, a larger share is paid by businesses, particularly in the services sector. This points to differences in transport habits, tax structures, and vehicle ownership across the EU [13]. Higher registration taxes may decrease new vehicle purchases but can also impact the decision to delay buying a new car or to purchase a used vehicle instead, increasing the share of older and more polluting vehicles [34]. Therefore, transport tax revenues may reflect the level of transport activity rather than reducing emissions. According to Ptak et al. [35], the environmental impact of fuel or transport taxes can be relatively limited. Similarly, Aydin and Bozatli [3] found that transport taxes increase air pollution for the whole panel of the ten countries with the highest transport tax revenues, concluding that transport taxes are not effective in reducing air pollution and that the structure of transport taxes should be changed. Erdogan [8] found that environmental taxes do not serve as an effective green fiscal policy tool in Germany.

Concerning government environmental expenditures, it has a significant negative effect, in-

dicating that 1% increase in environmental expenditures decreases emissions by 0.37% in the short run and 0.88% in the long run. This result is in line with López et al. [30], who stated that the allocation of government spending towards public goods such as environmental protection reduces pollution. Similarly, Halkos and Paizanos [21] found that higher government spending lowers emissions, with spending measures showing a more favourable environmental impact than tax-based instruments. Nguyen and Duong [31] confirmed that increased government spending significantly reduces CO_2 emissions in Vietnam. Furthermore, Batini et al. [5] concluded that green spending multipliers can significantly boost economic growth while supporting environmental sustainability.

The robustness check results confirm the stability of the baseline results in terms of signs and statistical significance of fiscal variables. Energy taxes and environmental protection expenditure consistently show statistically significant negative impacts on emissions, while transport taxes maintain their statistically significant positive impact on emissions. The population variable shows a marginally significant negative coefficient in Model (2) with GHG emissions (p-value = 0.069).

Long-run elasticities are presented in Table 5 and confirm the short-run results. Energy taxes and environmental protection expenditure reduce emissions in the long-run, while transport taxes remain positively associated with emissions.

Variable	**(1) GHG**	**(2) GHG + POP**	**(3) CO_2 + POP**	**(4) CO_2**
NRGTAX	−0.933	−0.926	−0.969	−1.045
TRATAX	0.979	0.802	0.991	1.093
ENVEXP	−0.883	−0.589	−0.713	−1.000

Table 5: *Long-run elasticities.*

The diagnostic tests are consistent across all models and confirm the adequacy of all model specifications, as reported in Table 6. The AR(1) test is significant, which is expected in first-differenced models, and the AR(2) test is insignificant, indicating that second-order autocorrelation is not present. Moreover, the Sargan test and Hansen test do not reject the null hypothesis for all models, which confirms instrument validity. The number of instruments remains relatively small (12 or 13, depending on the model) compared to the number of countries (27), preventing instrument proliferation.

Statistic	**(1) GHG**	**(2) GHG + POP**	**(3) CO_2 + POP**	**(4) CO_2**
Observations	349	349	349	349
Countries	27	27	27	27
Instruments	12	13	13	12
AR(1) p-value	0.018	0.051	0.059	0.015
AR(2) p-value	0.474	0.581	0.446	0.372
Sargan test p-value	0.208	0.329	0.373	0.346
Hansen test p-value	0.416	0.570	0.598	0.508

Table 6: *Model diagnostic tests' results.*

These results provide several important implications for the design of environmental fiscal policy in the European Union, particularly regarding the balance between tax-based and expenditure-based fiscal instruments. Both environmental taxation and protection expenditures contribute to emission reductions in the European Union. In particular, energy taxes and environmental expenditures are more effective in reducing emissions, while the current structure of transport taxes may limit their environmental effectiveness.

6. Conclusion

This paper assesses the impact of environmental fiscal policy instruments on greenhouse gas emissions in the EU-27 using difference GMM estimation from 2009 to 2023. The analysis provides an assessment of the impact of both the revenue and expenditure sides of environmental fiscal policy. The revenue side is represented by energy and transport taxes, while the expenditure side refers to government environmental protection expenditures.

The energy taxes have a significant negative effect on GHG emissions, with a 1% increase in energy taxes reducing emissions by 0.39% in the short run and 0.93% in the long run. Given that energy taxes account for approximately 76% of total environmental tax revenues in the EU, this finding underlines their importance in environmental policy. Government environmental protection expenditure also shows a significant negative effect on emissions, with short-run and long-run elasticities of -0.37 and -0.88, respectively. This suggests that expenditure-based fiscal instruments are effective in achieving environmental objectives. However, transport taxes exhibit a significant positive relationship with emissions, indicating that a 1% increase in transport taxes is associated with a 0.41% increase in emissions in the short run. This counterintuitive result can be explained by the structure of transport taxes in the EU, which are primarily related to vehicle ownership and registration rather than directly to fuel consumption or emissions.

These results are robust to the inclusion of population as a control variable and to the use of CO_2 emissions as an alternative dependent variable and have important policy implications for the design of environmental fiscal policy in the European Union. First, the negative and statistically significant impact of energy taxes on greenhouse gas emissions confirms that energy taxation is the most effective fiscal instrument on the revenue side for decreasing emissions. Given that energy taxes represent the largest share of environmental tax revenues in the EU, strengthening energy taxation could further support environmental sustainability. Second, the significant negative effect of government expenditure on environmental protection emphasizes the importance of public investment in environmental sustainability. This suggests that fiscal policy should not rely solely on taxation measures, but should also use expenditure-based instrument aimed at environmental protection. Finally, the positive relationship between transport taxes and emissions suggests that the current structure of transport taxation may not provide a sufficiently strong environmental incentive. Under the Eurostat methodology, transport taxes are exclusively ownership-based, while fuel taxation is captured by energy taxes [11]. The contrasting signs of the two tax variables confirm that usage-based taxation (energy taxes) effectively reduces emissions, whereas ownership-based taxation (transport taxes) does not. The positive coefficient on transport taxes therefore does not imply that all transport-related taxation is ineffective, but rather that the ownership-based component does not contribute to emission reductions. Policymakers may therefore consider restructuring ownership-based vehicle taxes to incorporate stronger emission-related criteria. Such reforms could better align transport taxation with environmental objectives and improve its effectiveness in reducing emissions.

The limitation of the study is that the model does not account for potential heterogeneity across member states in terms of economic development, energy mix, or institutional quality. Moreover, pollution taxes are not included in the analysis due to data constraints. Future research could extend this analysis by incorporating pollution taxes, as well as examining heterogeneous effects across different groups of EU member states. Additionally, future work could explore the inclusion of country-specific time trends, particularly as longer time series become available.

Acknowledgements

This research was funded by the European Union – NextGenerationEU under Call IIP-581-2025. The views and opinions expressed are solely those of the authors and do not necessarily reflect the official positions of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

References

[1] Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The Review of Economic Studies, 58(2), 277–297. doi: 10.2307/2297968

[2] Arnerić, J. and Šitum, A. (2022). PVAR model with collapsed instruments in the real exchange rates misalignment's analysis. Croatian Operational Research Review, 13(2), 203–215. doi: 10.17535/crorr.2022.0015

[3] Aydin, M. and Bozatli, O. (2022). Do transport taxes reduce air pollution in the top 10 countries with the highest transport tax revenues? A country-specific panel data analysis. Environmental Science and Pollution Research, 29(36), 54181–54192. doi: 10.1007/s11356-022-19651-8

[4] Baltagi, B. H. (2021). *Econometric Analysis of Panel Data* (6th ed.). Springer. doi: 10.1007/978-3-030-53953-5

[5] Batini, N., Di Serio, M., Fragetta, M., Melina, G. and Waldron, A. (2022). Building back better: How big are green spending multipliers? Ecological Economics, 193, 107305. doi: 10.1016/j.ecolecon.2021.107305

[6] Čeh Časni, A., Palić, I. and Žmuk, B. (2024). Impact of energy source composition on CO_2 emissions: An analysis using GDP and electricity production. In Družić, G. and Gelo, T. (Eds.), *Conference Proceedings of the 6th International Conference on the Economics of the Decoupling (ICED)* (pp. 195–210). Zagreb: Hrvatska akademija znanosti i umjetnosti; Ekonomski fakultet Sveučilišta u Zagrebu.

[7] Ehigiamusoe, K. U., Lee, C.-C. and Lean, H. H. (2025). Analysis of the economic and environmental imperatives of the service sector: The role of government in promoting sustainable development. Journal of Environmental Management, 376, 124470. doi: 10.1016/j.jenvman.2025.124470

[8] Erdogan, S. (2024). Linking green fiscal policy, energy, economic growth, population dynamics, and environmental degradation: Empirical evidence from Germany. Energy Policy, 189, 114110. doi: 10.1016/j.enpol.2024.114110

[9] EU. (2021). Regulation (EU) 2021/1119 establishing the framework for achieving climate neutrality (European Climate Law). Official Journal of the European Union, L 243, 1–17. url: https://eur-lex.europa.eu/eli/reg/2021/1119/oj

[10] Eurostat. (2013). European System of Accounts – ESA 2010. Publications Office of the European Union. doi: 10.2785/16644

[11] Eurostat. (2024). Environmental taxes – A statistical guide (2024 edition). Publications Office of the European Union. doi: 10.2785/730717

[12] Eurostat. (2025a). Greenhouse gas emission footprints. Eurostat Statistics Explained. url: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Greenhouse_gas_emission_footprints

[13] Eurostat. (2025b). Environmental tax statistics. Eurostat Statistics Explained. url: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Environmental_tax_statistics

[14] Eurostat. (2025c). Government expenditure on environmental protection. Eurostat Statistics Explained. url: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Government_expenditure_on_environmental_protection

[15] Eurostat. (2025d). Domestic net greenhouse gas emissions (sdg_13_10). url: https://ec.europa.eu/eurostat/databrowser/view/sdg_13_10/default/table

[16] Eurostat. (2025e). Environmental tax revenues. url: https://ec.europa.eu/eurostat/databrowser/view/env_ac_tax

[17] Eurostat. (2025f). General government expenditure by function (COFOG). url: https://ec.europa.eu/eurostat/databrowser/view/gov_10a_exp

[18] Eurostat. (2025g). Gross domestic product (GDP) and main components. url: https://ec.europa.eu/eurostat/databrowser/view/nama_10_gdp
[19] Eurostat. (2025h). Greenhouse gas emissions by source sector. url: https://ec.europa.eu/eurostat/databrowser/view/env_air_gge
[20] Eurostat. (2025i). Population on 1 January. url: https://ec.europa.eu/eurostat/databrowser/view/tps00001
[21] Halkos, G. E. and Paizanos, E. A. (2016). The effects of fiscal policy on CO_2 emissions: Evidence from the USA. Energy Policy, 88, 317–328. doi: 10.1016/j.enpol.2015.10.035
[22] Handoyo, S. (2024). Public governance and national environmental performance nexus: evidence from cross-country studies. Heliyon, 10(23), e40637. doi: 10.1016/j.heliyon.2024.e40637
[23] Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. Econometrica, 50(4), 1029–1054. doi: 10.2307/1912775
[24] Im, K. S., Pesaran, M. H. and Shin, Y. (2003). Testing for unit roots in heterogeneous panels. Journal of Econometrics, 115(1), 53–74. doi: 10.1016/S0304-4076(03)00092-7
[25] International Energy Agency. (2024). GHG emissions of all world countries. Publications Office of the European Union. doi: 10.2760/4002897
[26] Katircioglu, S. and Katircioglu, S. T. (2018). Testing the role of fiscal policy in the environmental degradation: The case of Turkey. Environmental Science and Pollution Research, 25(6), 5616–5630. doi: 10.1007/s11356-017-0906-1
[27] Kiviet, J., Pleus, M. and Poldermans, R. (2017). Accuracy and efficiency of various GMM inference techniques in dynamic micro panel data models. Econometrics, 5(14). doi: 10.3390/econometrics5010014
[28] Kohlscheen, E., Moessner, R. and Takats, E. (2025). Effects of carbon pricing and other climate policies on CO_2 emissions. National Institute Economic Review, 1–16. doi: 10.1017/nie.2025.10066
[29] Leitão, N. C. (2025). The impact of environmental taxes and renewable energy on carbon dioxide emissions in OECD countries. Energies, 18(10), 2543. doi: 10.3390/en18102543
[30] López, R., Galinato, G. I. and Islam, A. (2011). Fiscal spending and the environment: Theory and empirics. Journal of Environmental Economics and Management, 62(2), 180–198. doi: 10.1016/j.jeem.2011.03.001
[31] Nguyen, T. P. and Duong, T. T.-T. (2025). Asymmetric effects of fiscal policy and foreign direct investment inflows on CO_2 emissions – An application of nonlinear ARDL. Sustainability, 17(6), 2503. doi: 10.3390/su17062503
[32] Nickell, S. (1981). Biases in dynamic models with fixed effects. Econometrica, 49(6), 1417–1426. doi: 10.2307/1911408
[33] Okombi, I. F. and Ndoum Babouama, V. B. D. (2024). Environmental taxation and inclusive green growth in developing countries: Does the quality of institutions matter? Environmental Science and Pollution Research, 31(21), 30633–30662. doi: 10.1007/s11356-024-33245-6
[34] Pigato, M. and Hayde, E. (2021). Taxing vehicles. World Bank. url: https://documents1.worldbank.org/curated/en/863581636144031861/pdf/Taxing-Vehicles.pdf
[35] Ptak, M., Neneman, J. and Roszkowska, S. (2024). The impact of petrol and diesel oil taxes in EU member states on CO_2 emissions from passenger cars. Scientific Reports, 14(1), 52. doi: 10.1038/s41598-023-50456-y
[36] Roodman, D. (2009). How to do Xtabond2: An introduction to difference and system GMM in Stata. The Stata Journal, 9(1), 86–136. doi: 10.1177/1536867X0900900106
[37] Savranlar, B., Ertas, S. A. and Aslan, A. (2024). The role of environmental tax on the environmental quality in EU counties: evidence from panel vector autoregression approach. Environmental Science and Pollution Research, 31(24), 35769–35778. doi: 10.1007/s11356-024-33632-z
[38] United Nations Environment Programme. (2022). Kunming–Montreal Global Biodiversity Framework. url: https://www.cbd.int/doc/decisions/cop-15/cop-15-dec-04-en.pdf
[39] Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. Journal of Econometrics, 126(1), 25–51. doi: 10.1016/j.jeconom.2004.02.005

Study of novel normalization technique on weighting and ranking methodology in multi criteria decision making: Several cases in financial performance

Zaky Pratiwi[1], **Zahedi**[1,*] **and Badai Charamsar Nusantara**[2]

[1] *Faculty of Mathematics and Natural Science, Universitas Sumatera Utara, Medan 20155, Indonesia*
E-mail: ⟨zakypratiwi12@gmail.com, zahedi@usu.ac.id⟩

[2] *Faculty of Engineering, Universitas Sumatera Utara, Medan 20155, Indonesia*
E-mail: ⟨badainusantara9@gmail.com⟩

Abstract. The evaluation of corporate financial performance typically entails the use of multiple financial ratios characterized by heterogeneous properties, including the potential presence of negative values. Within this context, Multi-Criteria Decision-Making (MCDM) approaches have been extensively employed due to their capacity to systematically integrate diverse evaluation criteria into a coherent and objective ranking framework. However, most conventional normalization techniques in MCDM are not specifically designed to accommodate negative data, thereby potentially resulting in distorted weights and inconsistent ranking outcomes. This study proposes a novel normalization approach, termed the Root Distance Ratio (RDR), which is derived from the Weitendorf principle and formulated through a non-linear root transformation. The effectiveness of the RDR method is evaluated by integrating it into four MCDM techniques PSI, ARAS, MABAC, and EDAS across three corporate financial data scenarios, namely datasets with entirely positive values, negative values within benefit criteria, and negative values within cost criteria. The results demonstrate that, across the evaluated cases, the RDR method produces normalized values within the $[0, 1]$ interval, generates more proportionally distributed weights, and yields more consistent ranking outcomes compared to conventional normalization techniques. These findings indicate that RDR enhances both the stability and interpretability of financial data-driven multi-criteria decision-making processes, particularly in the presence of negative values.

Keywords: criteria weight, financial performance, MCDM, normalization technique, root distance ratio

Received: May 5, 2025; accepted: March 30, 2026; available online: May 21, 2026

DOI: 10.17535/crorr.2026.0023

Original scientific paper.

1. Introduction

Financial performance data, typically represented by financial ratios including liquidity, solvency, activity, and profitability indicators [2, 16], often encompass both positive and negative values, reflecting varying degrees of financial health and risk exposure [6, 22]. These heterogeneous ratios constitute the empirical foundation of the present study's evaluation framework within the context of Multi-Criteria Decision Making (MCDM) context.

MCDM is widely employed to comprehensively evaluate corporate financial performance due to its capacity to integrate criteria characterized by differing scales, preference directions, and levels of importance [8]. MCDM process typically comprises three primary stages: normalization, weighting, and aggregation, which collectively produce structured and objective

*Corresponding author.

rankings that reflect the relative performance of alternatives [15]. Among these stages, normalization plays a pivotal role by transforming heterogeneous data into comparable scales [7]. However, most conventional normalization techniques are developed under the assumption of non-negative data [23], with outputs typically constrained within the $[0, 1]$ interval [10]. Empirical evidence further indicates that the choice of normalization method can substantially influence the resulting rankings. For example, [3], demonstrated that linear and vector normalization within the VIšekriterijumsko KOmpromisno Rangiranje (VIKOR) method produce divergent outcomes, thereby underscoring that normalization constitutes a critical determinant of overall MCDM validity rather than merely a preliminary processing step.

When datasets contain negative values, conventional normalization techniques often fail to preserve stability within the intended range, thereby generating distorted weights and irrational ranking outcomes. [24] reported that the application of normalization within the Preference Selection Index (PSI) method to datasets containing negative values produced outputs exceeding acceptable bounds, with a single criterion disproportionately dominating the weight distribution, thus biasing the resulting rankings. Such findings underscore the inherent vulnerability of MCDM frameworks when confronted with non-ideal or extreme data conditions.

Previous studies have proposed alternative normalization approaches to mitigate these effects. [23] identified Weitendorf normalization as the only technique capable maintaining stability under extreme data conditions, while [3], [9], and [21] confirmed that the choice of normalization method directly influences both weighting and ranking outcomes. Consistent with these findings, [19, 20] demonstrated that different normalization procedures can systematically alter preference structures and lead to inconsistent MCDM rankings, thereby reinforcing the critical importance of normalization stability. Furthermore, various hybrid or integrated MCDM approaches such as Best Worst Method (BWM)-Ranking Alternatives by Perimeter Similarity (RAPS) [1], entropy-fuzzy Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) [11], Stepwise Weight Assessment Ratio Analysis (SWARA)-Weighted Aggregated Sum Product Assessment (WASPAS) [25], VIKOR [12], and PSI [24] have enhanced financial performance evaluation; however, they predominantly focus on datasets with positive values, thereby leaving a methodological gap in handling negative data.

Despite the development of numerous normalization techniques within the MCDM framework, substantial limitations persist, particularly when decision matrices contain negative values. Linear and vector-based methods are highly sensitive to scale disparities and sign variations, often resulting in biased weights and distorted ranking outcomes. Nonlinear and hybrid approaches, including those examined by [19, 20], may alter preference distributions or generate unstable value ranges when criteria differ significantly in magnitude. Furthermore, two-step procedures such as Alternative Ranking Order Method Accounting for Two-Step Normalization (AROMAN) [4], although designed to enhance scaling robustness, rely on underlying linear transformations that fail to ensure stability in the presence of negative or extreme financial ratios.

Although MCDM methods permit hybridization to mitigate bias, through the integration of PSI with Weitendorf normalization such combinations do not inherently ensure robustness. Stability in the normalization stage does not necessarily translate into consistent weighting and ranking outcomes. Accordingly, a clear research gap remains: many existing normalization techniques exhibit limited capability in handling negative values within decision matrices, particularly in preserving stability and proportionality throughout subsequent weighting and ranking processes. The primary motivation of this study is therefore to address this limitation by developing a normalization method that remains bounded, distributionally balanced, and robust under conditions involving negative data. To this end, this study proposes the Root Distance Ratio (RDR), a normalization technique designed to maintain values strictly within the $[0, 1]$ interval while improving distribution balance, thereby enhancing the rationality and consistency of MCDM results.

2. Failure of conventional MCDM methods on negative data

Four MCDM methods were employed to capture methodological diversity in multi-criteria evaluation: PSI [14], Additive Ratio Assessment (ARAS) [26], Multi-Attributive Border Approximation Area Comparison (MABAC) [18], and Evaluation based on Distance from Average Solution (EDAS) [5]. Each method is characterized by distinct computational mechanisms, rendering them suitable for assessing sensitivity to negative data in financial performance analysis. Their combined application facilitates a comprehensive comparison of how data irregularities influence normalization, weighting, and ranking outcomes.

2.1. Case scenario design based on real company data

This study employs 2023 financial data from publicly listed companies on the New York Stock Exchange (NYSE) and the National Association of Securities Dealers Automated Quotations (NASDAQ), representing three industry sectors. Based on these data, three analytical scenarios were constructed to capture varying financial conditions: Case X, comprising exclusively positive values; Case Y, incorporating a negative value in a benefit criterion (Return on Assets, ROA); and Case Z, involving a negative value in a cost criterion (Debt-to-Equity Ratio, DER). Table 1 presents the corresponding decision matrices, with negative values explicitly highlighted to denote anomalies.

Cases	Alternative	Criteria			
		C1 (CR)	C2 (DER)	C3 (TATO)	C4 (ROA)
X	X1	1.3545	0.9912	0.0010	0.2452
	X2	2.4942	0.6791	0.7618	0.1868
	X3	2.0630	0.6527	0.8297	0.1368
	X4	1.1094	0.4136	0.4637	0.3490
	X5	3.3149	0.4116	0.3980	0.0810
Y	Y1	1.0232	3.0506	0.8737	0.0475
	Y2	4.2760	0.1518	0.3107	0.0879
	Y3	1.2518	2.9604	0.3701	0.0172
	Y4	6.8233	0.1314	1.2536	0.2031
	Y5	0.8548	4.6928	0.0635	0.0074
	Y6	7.0275	0.1846	1.0032	0.1018
	Y7	1.8837	2.1803	0.3401	**-0.1035**
	Y8	2.4938	0.3683	0.0518	**-0.1883**
Z	Z1	2.8555	0.7762	0.7787	0.1665
	Z2	0.7467	**-2.0189**	0.0513	0.0051
	Z3	3.9872	0.2929	0.4342	0.1592
	Z4	3.0798	40.3118	0.0001	0.1053
	Z5	5.6900	0.2736	0.3965	0.1195
	Z6	1.2073	0.8386	0.4779	0.0360
	Z7	9.4312	0.1363	0.1660	0.0226

Table 1: *Decision matrix.*

As presented in Table 1, Case X represents a fully positive and stable evaluation condition, Case Y includes two companies exhibiting negative ROA values indicative of financial losses, and Case Z contains negative DER values reflecting potential financial distress or heightened bankruptcy risk. These three purposively selected cases illustrate varying financial conditions and enable assessment of whether conventional MCDM methods can yield rational and stable results under non-ideal data. Subsequent analyses focus on how negative values affect normalization, weighting, and ranking integrity.

2.2. Normalization using PSI, ARAS, MABAC, and EDAS

Normalization across all cases was conducted in accordance with the embedded procedures of each method PSI, ARAS, MABAC, and EDAS. Table 2 summarizes the resulting normalization ranges. In Case X, all methods produced values within the expected $[0, 1]$ interval [10], indicating stable performance under conditions involving exclusively positive data. Notably, MABAC, which incorporates the Weitendorf normalization principle, maintained consistent and logically coherent results. However, in Cases Y and Z, methods relying on linear normalization such as PSI, ARAS, and EDAS exhibited notable distortions. In PSI, negative ROA values in Case Y produced normalized outputs below zero ($Y_{74} = -0.5095$, $Y_{84} = -0.9273$), the cost criterion (DER) generated extreme values, whereby an alternative expected to perform best (Z72 = 0.1363) was instead ranked as the worst (-14.807). ARAS demonstrated similar instability, as negative denominators led to disproportionately scaled outcomes, and EDAS likewise failed to adequately accommodate extreme negative values, resulting in biased preference scores. Overall, the presence of even a single negative value proved sufficient to distort the entire normalization process across these methods.

Cases	**Methods**			
	PSI	**ARAS**	**MABAC**	**EDAS**
X	$0 \leq \bar{X}_{ij} \leq 1$	$0 \leq \bar{X}_{ij} \leq 1$	$0 \leq \bar{X}_{ij} \leq 1$	$0 \leq \bar{X}_{ij} \leq 1$
Y	$-1 \leq \bar{X}_{ij} \leq 1$	$-1 \leq \bar{X}_{ij} \leq 1$	$0 \leq \bar{X}_{ij} \leq 1$	$0 \leq \bar{X}_{ij} \leq 2$
Z	$-14 \leq \bar{X}_{ij} \leq 1$	$-1 \leq \bar{X}_{ij} \leq 1$	$0 \leq \bar{X}_{ij} \leq 1$	$-1 \leq \bar{X}_{ij} \leq 2$

Table 2: *Normalization value range*

2.3. Impact of negative values on weight calculation

Criterion weights were primarily determined using the PSI method due to its simplicity, objectivity, and ability to generate weights directly from normalized data. Table 3 summarizes the resulting weight distributions. Case X produced proportional and logical weights, whereas Cases Y and Z showed several negative and extreme values (for example, $w_{C4} = 0.9984$ in Case Y and $w_{C2} = 1.0040$ in Case Z). These distortions indicate that negative data cause instability in the weighting process, leading certain criteria to dominate or even contribute negatively to final rankings. Additional analyses employing Entropy and the Method based on the Removal Effects of Criteria (MEREC) weighting approaches further corroborated this issue, as both methods rely on logarithmic transformations that are undefined for zero or negative values. Consequently, conventional weighting techniques demonstrate limited robustness when applied to datasets containing negative financial ratios.

Cases	**Weights** (w_j)			
	C1	**C2**	**C3**	**C4**
Case X	0.2895	0.2978	0.1482	0.2645
Case Y	-0.0553	0.1192	-0.0623	0.9984
Case Z	-0.0023	1.0040	-0.0014	-0.0003

Table 3: *Weight values using conventional normalization.*

2.4. Ranking evaluation

Ranking results obtained from PSI, ARAS, MABAC, and EDAS are summarized in Table 4. In Case X, all methods produced identical rankings, indicating that the normalization, weighting, and aggregation procedures operated consistently and proportionally.

Cases	Alternative	Rank MCDM				Spearman Rank Correlation					
		I	II	III	IV	I vs II	I vs III	I vs IV	II vs III	II vs IV	III vs IV
X	X1	5	5	5	5	1	0.9	1	0.9	1	0.9
	X2	3	3	3	3						
	X3	4	4	4	4						
	X4	1	1	2	1						
	X5	2	2	1	2						
Y	Y1	4	4	4	4	0.976	1	0.976	0.976	1	0.976
	Y2	2	3	2	3						
	Y3	5	5	5	5						
	Y4	1	1	1	1						
	Y5	6	6	6	6						
	Y6	3	2	3	2						
	Y7	7	7	7	7						
	Y8	8	8	8	8						
Z	Z1	4	4	5	5	1	-0.25	-0.25	-0.25	-0.25	1
	Z2	1	1	1	1						
	Z3	5	5	4	4						
	Z4	2	2	7	7						
	Z5	6	6	3	3						
	Z6	3	3	6	6						
	Z7	7	7	2	2						

Table 4: *Alternative spearman rank correlation (I) PSI (II) ARAS (III) MABAC (IV) EDAS.*

However, Cases Y and Z revealed clear ranking irrationalities. In Case Y, inconsistencies emerged among mid-ranked alternatives (e.g., Y_2 and Y_6) due to the extreme dominance of criterion C_4 (ROA), whose weight reached 0.9983, causing all methods to rely almost entirely on ROA values. In Case Z, alternative Z_2 was ranked highest across all methods despite its negative DER (C_2), a value typically associated with bankruptcy risk. This anomaly resulted from an inflated weight (> 1) assigned to C_2, demonstrating that conventional normalization leads to distorted weights and rankings when negative data are present.

Extending this analysis to two-step normalization methods reveals that procedures such as AROMAN [4] likewise fail to rectify the ranking distortions induced by negative entries. Although AROMAN integrates a vector-based transformation with a min–max normalization, the vector component remains negative whenever the original criterion values are negative. These negative values propagate to the aggregation and weighting stages, causing several elements of the weighted decision matrix to remain negative. This becomes problematic in the ranking phase, because the computation of the benefit and cost utility terms requires exponentiation of aggregated values; if these values are negative, the resulting expressions become undefined or non-real. Conceptually, this limitation implies that AROMAN can only yield valid and comparable rankings when all normalized inputs entering its final stages are strictly positive, a condition satisfied under the proposed RDR normalization but not under AROMAN's internal steps when negative financial ratios are present. In essence, AROMAN becomes mathematically infeasible when negative normalized inputs appear, because its final utility stage requires raising aggregated values to fractional exponents. Any negative value entering this step produces a non-real output, making the ranking computation undefined and preventing the method from generating a valid ordering of alternatives.

2.5. Root cause of irrational weights and rankings

As discussed in Sections 2.2 and 2.3, the irrational weights and ranking outcomes observed in Cases Y and Z stem from the inability of conventional normalization techniques to adequately accommodate negative data. In the PSI method, excessive preference variance led to negative deviation values, as evidenced in Case Y, where the ROA variance reached 2.531 (deviation −1.531) and in Case Z where the DER variance increased to 177.898 (deviation −176.898). Under theoretically consistent conditions, preference variance should remain within the $[0, 1]$ interval to ensure the derivation of logical and interpretable weights.

These inflated variances arise from substantial discrepancies between average performance values and normalized results when values fall outside the ideal range, leading to uncontrolled dispersion and distorted weight distributions. Consequently, such distortions propagate from mean scores to variances, deviations, and ultimately to the final rankings. Therefore, there is a clear need for a normalization method that preserves valid value ranges and ensures stability, even in the presence of negative data.

3. Evaluation of existing normalization techniques

3.1. Comparison of all normalization methods

A comparative analysis was conducted on nine conventional normalization techniques summarized by [23], namely linear (N_1), Weitendorf (N_2), vector (N_3), linear sum (N_4), logarithmic (N_5), maximum linear (N_6), minimum linear (N_7), Jüttler-Körth (N_8), Peldschus (N_9) and AROMAN - Based Double Normalization (N_{10}).

Figure 1 illustrates that in Case X all methods (N_1–N_9) produced normalized values within the valid range $[0, 1]$, thereby demonstrating conformity with fundamental normalization principles. In these figures, the Lower Bound (LB) is defined as 0 and the Upper Bound (UB) as 1, collectively establishing the permissible normalization interval. Figure 2 shows that in Case Y a single negative ROA value (Y_{74} and Y_{84}) leads to instability across several normalization methods. Extreme results are observed in N_7 (26.285) and N_8 (−0.9272), while N_5 fails to compute outputs due to logarithmic limitations. Similar effects are also observed in N_1, N_3, N_4, and N_6.

In method N10, although two normalization stages are applied, alternatives containing negative values in the decision matrix still produce negative final normalized results (e.g., $Y_{74} = -0.024$ and $Y_{84} = -0.143$). This phenomenon arises from the aggregation mechanism embedded in the double-normalization procedure, which combines vector normalization $\frac{x_{ij}}{\sqrt{\sum x_{ij}^2}}$ and Weitendorf normalization $\frac{x_{ij}-x_{\min}}{x_{\max}-x_{\min}}$. The vector component preserves the sign of negative values, since dividing a negative value by $\sqrt{\sum x_{ij}^2}$ does not eliminate its sign, while the Weitendorf component maps values into the non-negative range. As a result, the negative influence from the vector normalization propagates through the aggregation process.

Because the aggregation parameter β is set to 0.5 to balance the contributions of both normalization techniques, the negative influence originating from the vector normalization component continues to affect the final result. Consequently, whenever negative values are present in the decision matrix, the double-normalization procedure generally yields negative normalized outcomes. Theoretically, a strictly non-negative result would only be attainable if the contribution of the vector component were eliminated or made negligible (i.e., $\beta \to 0$), effectively reducing the procedure to pure Weitendorf normalization. With $\beta = 0.5$, it is therefore highly unlikely for the double-normalization method to produce strictly positive normalized values when alternatives contain negative entries.

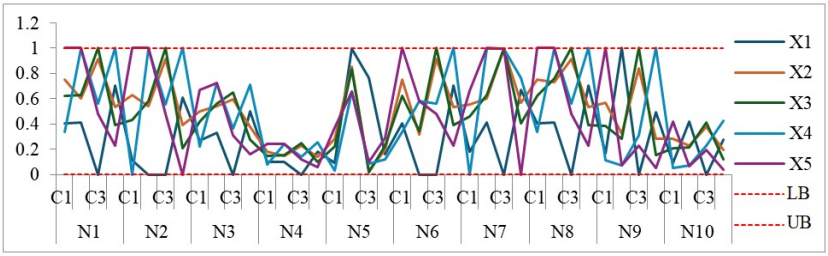


Figure 1: *The value of all normalization techniques for case X.*

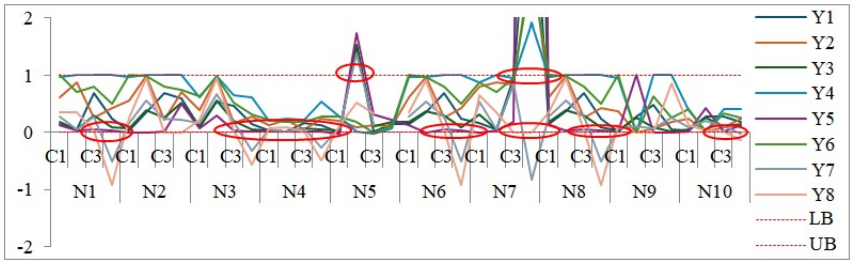


Figure 2: *The value of all normalization techniques for case Y.*

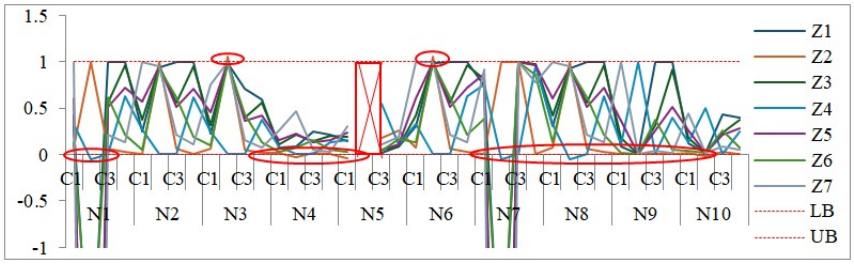


Figure 3: *The value of all normalization techniques for case Z.*

Figure 3 depicts similar instability in Case Z, where negative DER values affect multiple normalization methods. N_5 again fails to process the input values, while extreme values are observed in N_1 (-14.807 for Z_{72}) and N_6 (1.050 for Z_{22}). Other methods, such as N_3, N_4, and N_9, also exhibit distorted results. Using the same analytical reasoning as in Case Y, alternative Z_{22} also produced a negative value of -0.012 under the double-normalization method (N_{10}). This outcome occurs because the normalization technique does not explicitly distinguish between benefit and cost characteristics, thus reproducing the same distortion phenomenon observed previously. These distortions in normalized values subsequently give rise to irrational weighting outcomes and biased rankings. Among the evaluated techniques, only Weitendorf normalization (N_2) consistently maintained results within the $[0, 1]$ interval across Cases X, Y, and Z, thus confirming its stability and reliability for subsequent weighting and ranking analyses.

3.2. Application of Weitendorf as a partial solution

Weitendorf normalization was employed as a partial remedy to address the limitations of conventional linear normalization by integrating it into the PSI algorithm. The resulting criterion weights are presented in Table 5 and subsequently utilized for the ranking analysis.

Cases	Weights (w_j)			
	C1	**C2**	**C3**	**C4**
Case X	0.2425	0.2231	0.2518	0.2826
Case Y	-1.6402	-0.3046	0.1386	2.8061
Case Z	0.3933	0.2898	0.3330	-0.0161

Table 5: *Weight values using Weitendorf normalization.*

Despite all normalized values falling within the valid range, Table 5 indicates that Cases Y and Z still produced disproportionate weights, including negative and extreme values (e.g., -1.6402 and 2.8061). This anomaly arises from low average performance values that constrain deviation magnitudes, thereby impeding the generation of logical and proportionate weights. Thus, the issue lies not only in maintaining valid normalization ranges but also in ensuring a balanced value distribution. prevents mathematical inconsistencies, it remains insufficient in stabilizing weights or adequately improving value dispersion. Consequently, a more comprehensive approach is required one that simultaneously preserves valid normalization ranges and enhances distributional balance. To address this limitation, the present study introduces the RDR normalization technique as a more robust solution for generating consistent and interpretable weights, particularly under conditions involving negative data.

4. Proposed normalization technique: root distance ratio (RDR)

4.1. Motivation and theoretical basis

The irrationality of weight values in conventional normalization techniques, including Weitendorf, originates from low average performance values (N_j) produced during normalization. Alternatively, power transformations with positive integer exponents (x^p, $p > 1$) further reduce the magnitude of normalized values, resulting in $x^p < x$ for $0 < x < 1$, thereby exacerbating the problem of low average performance values. In contrast, root-based transformations ($x^{1/p}$) increase the magnitude of values within $(0, 1)$, making them more suitable for enhancing N_j. However, excessively high-order roots tend to compress the relative differences among alternatives, as all values tend to approach 1 (formally, $x^{1/p} \to 1$ as $p \to \infty$ for $0 < x \leq 1$), reducing the discriminative ability of the normalization.

Therefore, the square-root transformation ($p = 2$) is adopted as a balanced adjustment, as it increases normalized values while preserving sufficient differentiation among alternatives. Based on this reasoning, the Root Distance Ratio (RDR) normalization is defined as follows:

$$\bar{X}_{ij} = \begin{cases} \dfrac{\sqrt{x_{ij} - x_{\min}}}{\sqrt{x_{\max} - x_{\min}}}, & \text{for benefit criteria,} \\ \dfrac{\sqrt{x_{\max} - x_{ij}}}{\sqrt{x_{\max} - x_{\min}}}, & \text{for cost criteria.} \end{cases} \tag{1}$$

where x_{ij} represents the value of the i-th alternative on the j-th criterion, $x_{\max}$ denotes the maximum value for a given criterion, and $x_{\min}$ denotes the minimum value.

4.2. Distributional comparison: RDR vs Weitendorf

A comparative evaluation between the Weitendorf and RDR normalization techniques was conducted to examine RDR's adaptation of Weitendorf's principles and its capacity to enhance normalized values. Figures 4(a)-(c) show that Weitendorf results (red points) align closely with the linear reference line $y = x$, reflecting its linear transformation that maps negative values into the $[0, 1]$ range. In contrast, RDR results (green points) form a convex non-linear pattern above

the line, illustrating the effect of the square root transformation on distance ratios. These visual and quantitative observations confirm that while both techniques maintain valid normalization ranges, RDR achieves a more balanced and expanded value distribution, promoting stability and rationality in the subsequent weighting process.

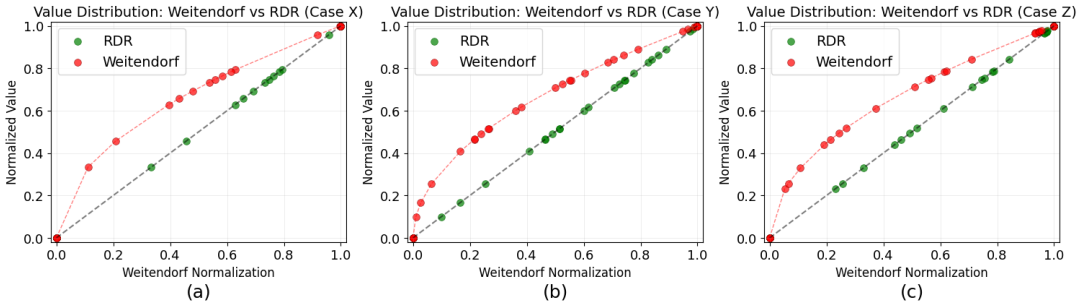


Figure 4: *Weitendorf vs RDR on (a) Case X, (b) Case Y, and (c) Case Z.*

4.3. PSI weighting integration using RDR

RDR normalization technique was implemented as a replacement for conventional normalization within the MCDM framework, applied to both weighting and ranking stages. In the PSI method, RDR substituted the original normalization component, producing the criterion weights shown in Table 6. As indicated in Table 6, all criteria across the three cases yielded proportional weights within the $[0, 1]$ interval. This improvement results from RDR's ability to elevate average performance values, thereby reducing deviation magnitudes and generating balanced, logically consistent weights.

Cases	**Weights** (w_j)			
	C1	**C2**	**C3**	**C4**
Case X	0.2544	0.2193	0.2377	0.2886
Case Y	0.0431	0.2747	0.1481	0.5341
Case Z	0.3634	0.1903	0.2924	0.1538

Table 6: *Weight values using RDR normalization.*

4.4. Substitution of RDR normalization in PSI, ARAS, MABAC, and EDAS rankings

The evaluation of ranking consistency across the conventional, Weitendorf, and RDR approaches reveals clear differences in pattern. Without normalization substitution, radar charts for Cases X, Y, and Z (Figure 5) displayed irregular polygon shapes and large fluctuations among alternatives, indicating that rankings were highly dependent on the specific MCDM method. The use of Weitendorf normalization improved distribution balance, though deviations persisted in cases with negative values. In contrast, applying RDR normalization produced markedly stronger consistency, as reflected in the more symmetrical radar chart shapes in Figure 5. Rankings across PSI, ARAS, MABAC, and EDAS became closely aligned, showing that RDR enhances both value distribution and decision stability.

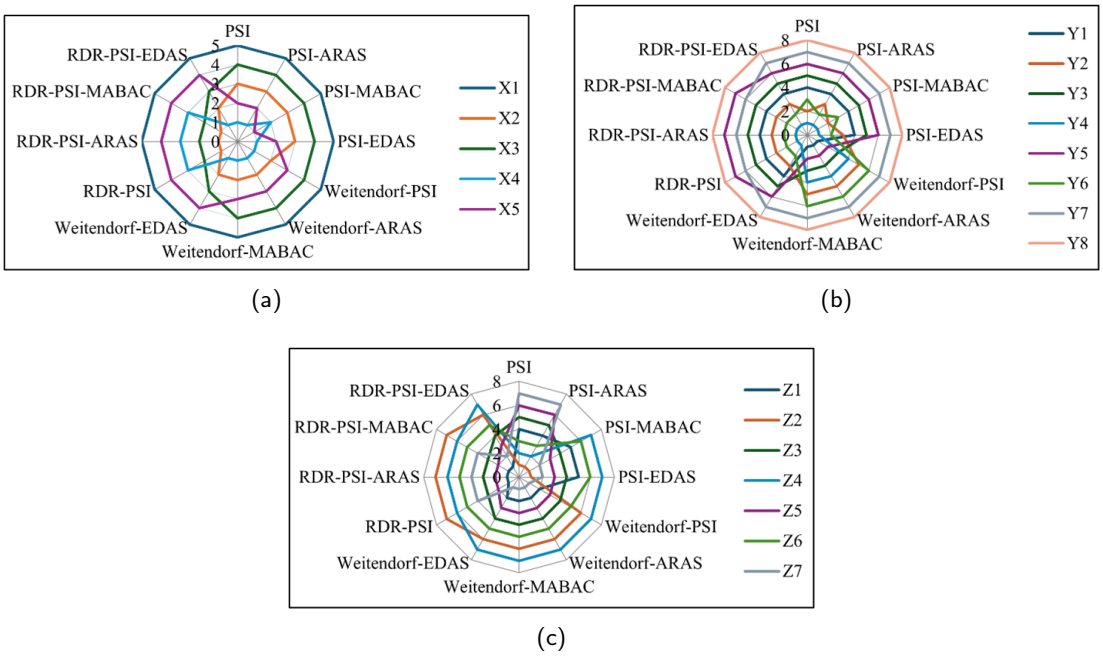


Figure 5: *Ranking results across methods for (a) Case X, (b) Case Y, and (c) Case Z.*

Cases	Spearman Rank Correlation					
	I vs II	**I vs III**	**I vs IV**	**II vs III**	**II vs IV**	**III vs IV**
X	1	1	0.7	1	0.7	0.7
Y	1	1	0.976	1	0.976	0.976
Z	1	1	0.857	1	0.857	0.857

Table 7: *Spearman rank correlation after the application of RDR (I) PSI (II) ARAS (III) MABAC (IV) EDAS.*

Spearman rank correlation results (Table 7) further confirm these findings. Correlations among methods increased substantially, indicating that RDR-based normalization yields more coherent and rational rankings than conventional approaches. The slight decrease observed in EDAS correlations does not suggest inconsistency but rather a more realistic alignment with the data's intrinsic structure, underscoring the robustness of the RDR normalization.

5. Evaluation using other datasets

5.1. Normalization simulation using equivalent interval ratios

In this section, a simulation is conducted using two datasets that possess identical interval ratios but differ in numerical ranges. Dataset A spans the interval 1–9 with a unit increment of 1 for nine alternatives (A_1–A_9), whereas Dataset B spans the interval 1–5 with an increment of 0.5 for the same nine alternatives (A_1–A_9), applied to both benefit ($*$) and cost ($**$) criteria.

This simulation evaluates the ability of normalization techniques to generate consistent outcomes when applied to datasets that have equivalent structural capacity yet different numerical scales. The analysis employs five normalization techniques: Root Distance Ratio (N_1), Linear (N_2), Vector (N_3), Sum-Linear (N_4), and AROMAN-based Double Normalization (N_5).

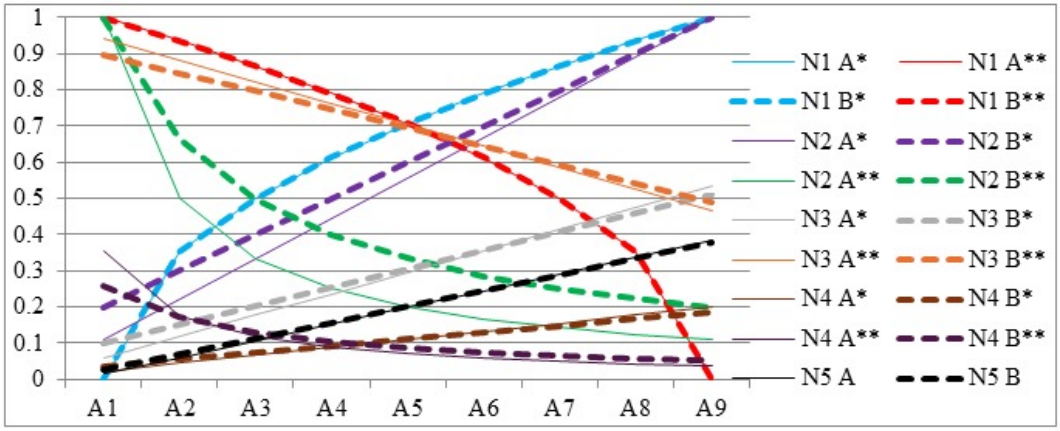


Figure 6: *Comparison of normalization results using equivalent interval ratios.*

Overall, all normalization methods employed in this simulation produced values within the $[0, 1]$ interval. This outcome is expected, as all raw data in both datasets are strictly positive. However, as illustrated in Figure 6, among the evaluated methods, N1 (RDR) demonstrates the closest overlap between Dataset A and Dataset B for both benefit and cost criteria.

This finding indicates that RDR is capable of producing highly consistent normalized values for datasets with equivalent interval structures, even when their numerical scales differ.

Although N_4 (Sum-Linear) and N_5 (AROMAN) yield curves that appear nearly overlapping, the resulting normalized values are not exactly identical, even though both datasets possess the same interval structure. Meanwhile, N_2 (Linear) and N_3 (Vector) exhibit substantially different normalization results, as reflected by the noticeable separation between the curves across the two datasets. Furthermore, the cost-type normalization under N_2 (Linear) fails to produce a truly linear pattern, revealing additional instability in the scale transformation process.

5.2. Weighting and ranking simulation based on case scenarios

Two simulated datasets, representing engineering and economic domains, were analyzed to compare the performance of linear and RDR normalization. Both datasets replicated real-world characteristics, including negative values in cost-type criteria (C_1 in Dataset 1 and C_4 in Dataset 2).

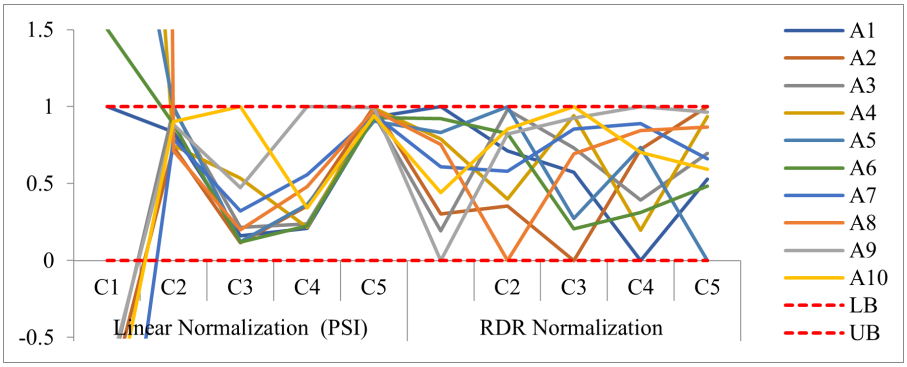


Figure 7: *Comparison of normalization value ranges in the first simulation.*

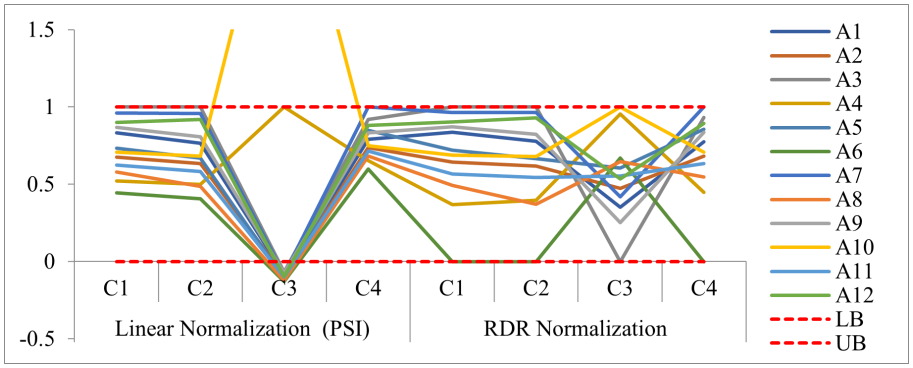
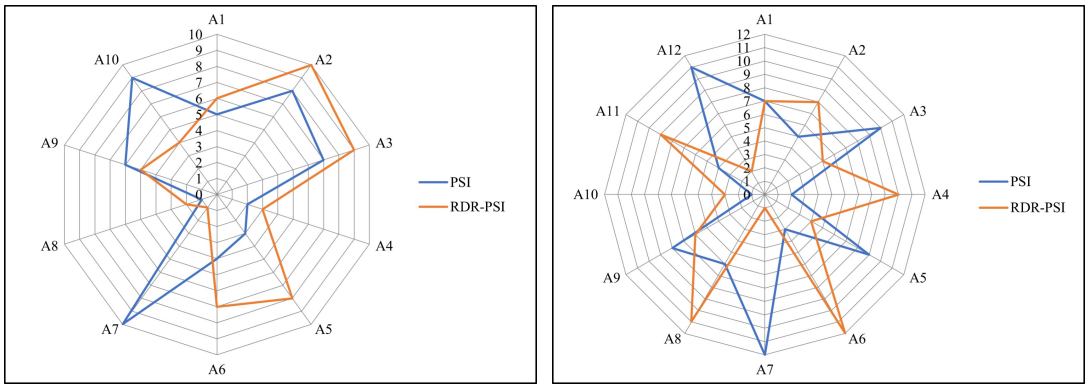


Figure 8: *Comparison of normalization value ranges in the second simulation.*



(a) (b)

Figure 9: *Ranking in the (a) first and (b) second simulation.*

Figures 7-9 illustrate the comparative outcomes across normalization, weighting, and ranking stages. Under linear normalization, several values exceeded the $[0, 1]$ range, producing distortions in both datasets. In Dataset 1, alternative A_7 under criterion C_1 yielded a negative normalized value despite representing the lowest-cost option, while in Dataset 2 the presence of negative values in criterion C_4 destabilized the entire normalization matrix. In contrast, the RDR approach maintained all values strictly within the $[0, 1]$ interval, thereby preserving proportional relationships among criteria and ensuring numerical stability.

The weighting results (Table 8) further underscore this improvement. Linear normalization produced inconsistent weights, including negative values and coefficients exceeding unity, thereby leading to unreliable aggregation. In contrast, RDR normalization generated balanced and theoretically consistent weights, all confined within the $[0, 1]$ interval, thus confirming greater accuracy and internal consistency within the PSI framework.

Data	Types of Normalization	C1	C2	C3	C4	C5
1	Linear Normalization (PSI)	1.00123	-0.00045	-0.00038	0.00041	-0.00049
	RDR Normalization	0.41079	0.18233	0.17075	0.23613	0.32588
2	Linear Normalization (PSI)	-0.12229	-0.11018	1.39137	-0.15891	
	RDR Normalization	0.18407	0.14974	0.30097	0.36523	

Table 8: *The criterion weight result.*

Ranking analyses (Figure 9) reinforce the same pattern. Linear PSI exhibited extreme preference disparities, such as the dominance of A_7 and the suppression of A_6 in Dataset 1, and the over-ranking of A_7 and A_{12} in Dataset 2. The RDR-PSI integration yielded smoother, more uniform radar profiles, where alternatives occupied logical positions with no unrealistic dominance.

Overall, distortions arising from normalization and weighting stages in the conventional PSI method were clearly propagated to the final ranking outcomes. The substituting of RDR normalization effectively mitigated these issues, resulting in coherent aggregate scores and enhanced decision reliability across both domains. Consequently, the RDR-PSI model demonstrates superior robustness and interpretability when applied datasets containing negative cost-type criteria.

6. Conclusion

This study investigated the limitations of conventional normalization techniques in Multi-Criteria Decision Making (MCDM) when applied to datasets containing negative values and proposed the Root Distance Ratio (RDR) as an alternative approach. The analysis focused on normalization stability, weight proportionality, and ranking consistency utilizing both real financial data (NYSE and NASDAQ) and simulated datasets.

The results indicate that linear normalization techniques are inadequate for handling negative values, leading to distorted weights and inconsistent ranking outcomes. In contrast, the RDR method maintained normalized values within the $[0, 1]$ interval, generated proportionate weight distributions, and produced generally consistent ranking patterns across the evaluated cases. These improvements were consistently observed across multiple MCDM methods (PSI, ARAS, MABAC, and EDAS) and further validated through rank correlation analysis and radar-based visualizations.

However, in specific scenarios involving negative cost-type criteria, such as the Debt-to-Equity Ratio (DER), minor variations in ranking positions may still occur, despite the stability of the normalization process. This suggests opportunities for further refinement, rather than indicating fundamental limitations of the proposed approach. Future research should therefore focus on developing an enhanced normalization framework capable of preserving both mathematical validity and semantic preference consistency, particularly under extreme or outlier conditions.

Acknowledgements

The authors would like to acknowledge support from Namira Rambe for her insightful discussions and analytical contributions, and from Nasya Nabila for her kind moral encouragement during the research and manuscript preparation.

References

[1] Alamoudi, M. H. and Bafail, O. A. (2022). BWM—RAPS Approach for Evaluating and Ranking Banking Sector Companies Based on Their Financial Indicators in the Saudi Stock Market. Journal of Risk and Financial Management, 15(10), 467. doi: 10.3390/jrfm15100467

[2] Alqam, M. A., Ali, H. Y. and Hamshari, Y. M. (2021). The Relative Importance of Financial Ratios in Making Investment and Credit Decisions in Jordan. International Journal of Financial Research, 12(2), 284. doi: 10.5430/ijfr.v12n2p284

[3] Budiman, E. and Hairah, U. (2021). Comparison of Linear and Vector Data Normalization Techniques in Decision Making for Learning Quota Assistance. Journal of Information Technology and Its Utilization, 4(1), 22–28. doi: 10.30818/jitu.4.1.3897

[4] Bošković, S., Svadlenka, L., Jovčić, S., Dobrodolac, M., Simic, V. and Bacanin, N. (2023). An Alternative Ranking Order Method Accounting for Two-Step Normalization (AROMAN) - A Case

Study of the Electric Vehicle Selection Problem. IEEE Access, 11, 39496-39507. doi: 10.1109/ACCESS.2023.3265818

[5] Keshavarz-Ghorabaee, M., Zavadskas, E. K., Olfat, L. and Turskis, Z. (2015). Multi-Criteria Inventory Classification Using a New Method of Evaluation Based on Distance from Average Solution (EDAS). INFORMATICA, 26(3), 435–451. doi: 10.15388/Informatica.2015.57

[6] Isayas, Y. N. (2021). Financial distress and its determinants: Evidence from insurance companies in Ethiopia. Cogent Business and Management, 8(1). doi: 10.1080/23311975.2021.1951110

[7] Jafaryeganeh, H., Ventura, M. and Guedes Soares, C. (2020). Effect of normalization techniques in multi-criteria decision making methods for the design of ship internal layout from a Pareto optimal set. Structural and Multidisciplinary Optimization, 62(4), 1849–1863. doi: 10.1007/s00158-020-02581-9

[8] Kaya, A., Pamucar, D., Gürler, H. E. and Ozcalici, M. (2024). Determining the financial performance of the firms in the Borsa Istanbul sustainability index: integrating multi criteria decision making methods with simulation. Financial Innovation, 10(1), 21. doi: 10.1186/s40854-023-00512-3

[9] Komatina, N. (2025). A Novel BWM-RADAR Approach for Multi-Attribute Selection of Equipment in the Automotive Industry. Spectrum of Mechanical Engineering and Operational Research, 2(1), 104–120. doi: 10.31181/smeor21202531

[10] Kumar, R., Prinshu, Mishra, A. K., Dutta, S. and Singh, A. K. (2023). Optimization and prediction of response characteristics of electrical discharge machining using AHP-MOORA and RSM. Materials Today: Proceedings, 80, 333–338. doi: 10.1016/j.matpr.2023.02.138

[11] Lam, W. H., Lam, W. S., Liew, K. F. and Fun, L. P. (2023). Decision Analysis on the Financial Performance of Companies Using Integrated Entropy-Fuzzy TOPSIS Model. Mathematics, 11(2), 397. doi: 10.3390/math11020397

[12] Lam, W. S., Lam, W. H., Jaaman, S. H. and Liew, K. F. (2021). Performance evaluation of construction companies using integrated entropy–fuzzy vikor model. Entropy, 23(3), 1–16. doi: 10.3390/e23030320

[13] Maity, R., Mishra, R., Pattnaik, P. K. and Pandey, A. (2023). Selection of sustainable material for the construction of UAV aerodynamic wing using MCDM technique. Materials Today: Proceedings. doi: 10.1016/j.matpr.2023.12.025

[14] Maniya, K. and Bhatt, M. G. (2010). A selection of material using a novel type decision-making method: Preference selection index method. Materials and Design, 31(4), 1785–1789. doi: 10.1016/j.matdes.2009.11.020

[15] Muthia, D., Zahedi and Nusantara, B. C. (2025). Ranking Quality of Life Index in Indonesian Provinces: A Multicriteria Approach for Sustainable Regional Development. Sustainable Development, 33(2), 3043–3061. doi: 10.1002/sd.3283

[16] Nukala, V. B. and Prasada Rao, S. S. (2021). Role of debt-to-equity ratio in project investment valuation, assessing risk and return in capital markets. Future Business Journal, 7(1), 1–23. doi: 10.1186/s43093-021-00058-9

[17] Osborne, J. W. (2005). Notes on the use of data transformations. Practical Assessment, Research and Evaluation, 8(1), 6. doi: 10.7275/4vng-5608

[18] Pamučar, D. and Ćirović, G. (2015). The selection of transport and handling resources in logistics centers using Multi-Attributive Border Approximation area Comparison (MABAC). Expert Systems with Applications, 42(6), 3016–3028. doi: 10.1016/j.eswa.2014.11.057

[19] Pavličić, D. M. (2000). Normalization of attribute values in MADM violates the conditions of consistent choice IV, DI and α. Yugoslav Journal of Operations Research, 10(1), 109–122. url: eudml.org/doc/261486 [Accessed 5/5/2025]

[20] Pavličić, D. M. (2001). Normalisation effects the results of MADM methods. Yugoslav Journal of Operations Research, 11(22), 251-265. url: eudml.org/doc/261661 [Accessed 5/5/2025]

[21] Petrović, N., Jovanovic, V., Petrović, M., Nikolic, B. and Mihajlović, J. (2025). Comparative Investigation of Normalization Techniques and Their Influence on MCDM Ranking – A Case Study. Spectrum of Mechanical Engineering and Operational Research, 2(1), 172–190. doi: 10.31181/smeor21202542

[22] Saragih, W. N. M., Zahedi and Nusantara, B. C. (2024). Comparative analysis of QUEST and CHAID decision tree methods in assessing the performance of conventional commercial banks.

Songklanakarin Journal of Science and Technology, 46(3), 279–285.

[23] Trung, D. D., Truong, N. X., Duc, D. V. and Bao, N. C. (2025). Data normalization in RAWEC method: Limitations and remedies. Yugoslav Journal of Operations Research, 35(3), 467–482. doi: 10.2298/YJOR240315020T

[24] Wardany, R. N. and Zahedi (2025). A study comparative of PSI, PSI-TOPSIS, and PSI-MABAC methods in analyzing the financial performance of state-owned enterprises companies listed on the Indonesia stock exchange. Yugoslav Journal of Operations Research, 35(2), 313–330. doi: 10.2298/YJOR240115017W

[25] Yazdi, A. K., Hanne, T. and Osorio Gómez, J. C. (2020). Evaluating the performance of colombian banks by hybrid multicriteria decision making methods. Journal of Business Economics and Management, 21(6), 1707–1730. doi: 10.3846/jbem.2020.11758

[26] Zavadskas, E. K. and Turskis, Z. (2010). A new additive ratio assessment (ARAS) method in multicriteria decision-making. Technological and Economic Development of Economy, 16(2), 159–172. doi: 10.3846/tede.2010.10

Croatian Operational Research Review
CRORR **17:2**(2026), 325–339

Mitigating the consequences of disruptions by locating temporary facilities and balancing the redistribution of users

Predrag Grozdanović[1,*] **and Miloš Nikolić**[1]

[1] *University of Belgrade, Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, Belgrade, Serbia*
E-mail: ⟨{p.grozdanovic, m.nikolic}@sf.bg.ac.rs⟩

Abstract. Many service systems are subject to disruptions that can reduce their operational capacity, forcing users to find alternative ways to meet their service needs. One way to mitigate such effects is to locate temporary facilities where users can receive services. Also, users can be served in facilities unaffected by the disruptions. This paper considers the problem of locating temporary facilities after a disruption and redistributing users to minimize the average travel distance. Compared with existing models, the proposed approach allows limited overcapacity and balances it across facilities to ensure a more even redistribution of users. A mathematical formulation of mixed integer programming is proposed and used to determine the locations of temporary facilities and the balanced redistribution plan. The final number of temporary facilities is determined using a TOPSIS method. The proposed approach integrates an exact approach with the TOPSIS method and enables the straightforward addition of new criteria without modifying the mathematical formulation. Testing on small-sized hypothetical examples demonstrates that, without capacity constraints, overcapacity in some facilities can reach 150%, whereas the proposed approach limits it to 45% and balances it so that differences among facilities do not exceed 10%. The results of testing illustrate that the proposed approach effectively solves the multi-objective location and redistribution problem while avoiding computationally intensive methods such as multi-criteria optimization or metaheuristics. The proposed approach provides a flexible and implementable framework that balances system management and user objectives under disruptions.

Keywords: allocation of users, capacity constraints, disruption events, facility location, temporary facilities

Received: November 24, 2025; accepted: April 11, 2026; available online: May 21, 2026

DOI: 10.17535/crorr.2026.0024

Original scientific paper.

1. Introduction

Many service systems are susceptible to disruptions that can reduce their ability to operate. As a result, many users have to find alternative ways to meet their service needs. One way to mitigate the negative effects of the disruption is to locate new temporary facilities where users can receive the desired services. Unfortunately, such events have become increasingly frequent in recent years. A disruption in the system can be acted on before or after disruptions (pre-disruption or post-disruption). Most of the papers in the literature propose approaches that ensure greater resistance of systems and their users to disruptions (pre-disruption) [7] and [1]. Fewer papers in the literature deal with the post-disruption period [3]. The response to a disruption after its occurrence ensures the immediate protection of people's health and

*Corresponding author.

lives by establishing temporary facilities and plans to return the system to the state before the disruption as soon as possible [8].

In case of various natural disasters, terrorist attacks, or other causes of disruption, one or more facilities in the system may be closed. One possible response is to develop a plan for redistributing users affected by the disruption to active facilities. However, such a way of responding to a disruption can only be applied in the case of disruptions of smaller dimensions. To overcome disruptions of larger dimensions, the introduction of temporary facilities must be considered. Creating a plan to redistribute users affected by the disruption to active permanent facilities (PF) and locate temporary facilities (TF) allows users to receive services in an acceptable time frame.

The importance of introducing temporary facilities after a disruption, when it is important to allocate resources quickly and efficiently, was highlighted in a review paper [2]. The authors provide a detailed review of the literature, including 65 papers that address location theory problems associated with the introduction of temporary facilities. One of the most famous location problem is the p-median problem [11]. The p-median model with capacity constraints is known as the capacitated p-median model. The p-median problem is much more often solved in the literature compared to the capacitated p-median problem. In the capacitated p-median problems [12] and [10], as well as in many p-median problems, redistribution of users from one node to one facility is allowed. This paper proposes a model based on the capacitated p-median problem, which enables the redistribution of users from one node to one or more facilities.

Unsatisfied demand is a problem that often occurs after disruptions in the system, especially in the case of capacity constraints. In the literature, there are papers dealing with the minimization of unsatisfied demand [4] and [6]. In the case of unsatisfied demand, fictitious facilities must be introduced, which is the case with classic capacity constraints [10]. Various forms of capacity constraints are used in the literature [12] and [4]. Common to those constraints is that they lead to the situation that if there is not enough capacity, not all users will receive the service. Capacity constraints designed in this paper allow all users to receive the requested service. This is achieved by introducing temporary facilities and allowing overcapacity to a certain limit, which enables normal functioning of the facility. These constraints solve the problem of unsatisfied demand while maintaining the efficiency of the system.

The use of hybrid approaches, which integrate the Exact Approach (EA) with multi-attribute decision-making (MADM) methods, is common in location problems under disruption conditions. MADM methods are most often used in the first phase, which includes the analysis of locations, to generate a set of potential locations of temporary facilities [5] and [14]. Unlike the papers in the literature, in this paper, the TOPSIS method was used after applying the optimization model (second phase) in order to select the best solution. By combining the EA for solving the capacitated p-median problem with the TOPSIS method (the EA-TOPSIS approach), several issues observed in the literature are addressed, such as the redistribution of users from one node to multiple facilities, unsatisfied demand, facility overloading, and non-balanced user redistribution plans. In related studies, similar problems are frequently addressed using robust, stochastic, or bi-level models, which require complex optimization methods for their solutions. Due to the conflicting objectives of minimizing user travel distance and limiting the number of temporary facilities, the considered problem has a multi-objective structure. Without the proposed hybrid framework, it would typically require multi-objective optimization or metaheuristic solution methods [9] and [1]. The developed EA-TOPSIS approach overcomes this by separating optimization and decision evaluation, thereby reducing methodological and computational complexity.

In this paper, a location model was developed that determines the location of TFs and creates a new plan for the redistribution of users allocated to the closed PFs. A temporary facility can be located at any demand node where there is no permanent facility. System adaptation to disruption is also reflected in the possibility of exceeding the capacity of PF and

TF to a certain limit. To evenly redistribute users across the functioning facilities, overcapacity balancing constraints have been implemented. The goal of the developed model is to minimize the average distance traveled by users affected by the disruption to obtain the requested service. It is in the interest of the system to overcome the consequences of the disruption with as few TFs located as possible. However, a smaller number of TFs affects the increase in the distance traveled by users and greater overcapacity in facilities. To create a balance between the system's goal and its users, the TOPSIS method is combined with a developed mathematical formulation to determine the number and location of TFs. In this way, the solution obtained by combining the EA and the TOPSIS method (EA-TOPSIS) represents a compromise between increasing the distance traveled by users and reducing the number of TF locations. We tested the EA-TOPSIS approach on a hypothetical example of 20 demand nodes and 5 PFs and obtained high-quality results. The obtained results confirm the possibility of direct application of the developed EA-TOPSIS approach in practice. The EA-TOPSIS approach can be applied to various systems, such as justice systems, police systems, healthcare systems, and other systems in which a single facility serves users from a particular region.

Several main contributions are highlighted in this paper. First, a mixed-integer location model is developed to simultaneously determine the locations of TFs and the redistribution plan for users affected by disruptions. Second, the model allows the redistribution of users from a single demand node to multiple facilities, enabling more flexible system adaptation under disruption conditions. Third, overcapacity balancing constraints are implemented, enabling controlled overcapacity in both PFs and TFs and ensuring an even redistribution of users across functioning facilities. Finally, an EA–TOPSIS hybrid approach is developed for determining the number and locations of TFs, and its practical applicability is demonstrated through computational testing.

This paper is organized as follows. Section one presents the introductory remarks and highlights the main contributions of the paper. The second section provides a description of the considered problem, its mathematical formulation, and the TOPSIS method that is adapted for application and combination with the previously given mathematical formulation. The developed EA-TOPSIS approach is then tested on a hypothetical example, and the test results are presented in the fourth section. Concluding remarks are given at the end of the paper.

2. Methodology

In this section, the problem and the system's response to disruptions are presented, along with the proposed EA–TOPSIS approach. The following subsections detail the mathematical formulation of the capacitated p-median model and the TOPSIS method.

2.1. Problem description and proposed approach

The problem considered in this paper is inspired by the functioning of systems in which a single facility provides services to all users in the region in which it is located. A region encompasses one or more municipalities in which users live. In such systems, all users of a region gravitate to a single facility and can only receive the requested service in that facility. Under normal conditions, users can only request service in the preassigned facility. They cannot receive service in any of the facilities from other regions. The most common examples of such systems are public buildings owned by the state, such as municipal buildings, courts, police, health institutions, etc.

A challenge in the operation of the described systems may arise if a facility in one region is affected by a disruption. In the event of an attack on one or more facilities, users preassigned to that facility or those facilities must be redistributed to active facilities (non-closed). In this case,

due to the reduction in the number of operational facilities (non-closed facilities), overcapacity of the facilities is permitted. The permitted overcapacity of facilities and the location of TFs will ensure that the consequences of disruption (attacks on one or more facilities) are as minimal as possible.

All municipalities in which there is no PF were taken as potential locations for TFs. The placement of TFs is considered with the following constraints. Only one TF can be established at each potential location. A TF cannot be established at the locations of the PF and, therefore, not at the site of the disruption. When addressing the described problem, it is necessary to determine which facilities are under attack and the maximum number of TFs that can be located. By resolving the stated issue, a solution is produced that identifies the number and location of TFs, as well as the redistribution of users affected by the disruption. The aim of solving this problem is to minimize the impact of the disruption on both users and the system. More specifically, the aim is to minimize the average distance traveled by users affected by the disruption to access the required service, while optimizing the number of located TFs, and balancing overcapacity in facilities.

Locating a larger number of TFs will reduce the consequences of disruptions, but will result in high construction costs. However, locating too few TFs may cause users to feel the effects of disruptions. The goal in selecting the final solution is to find a balance between the number of TFs located and the average distance that users affected by the disruption travel to receive service. Therefore, the paper proposes an EA-TOPSIS approach that combines the EA, used to solve the proposed mathematical formulation, with the TOPSIS method for multi-criteria evaluation of the generated solutions. The methodological framework of the proposed EA-TOPSIS approach is illustrated in Figure 1. The approach consists of two sequential phases. In Phase I (EA), the capacitated p-median model is solved iteratively for an increasing number of temporary facilities until a predefined maximum p_{max} is reached, generating a set of feasible alternatives. In Phase II (TOPSIS), the generated alternatives are evaluated and ranked using the TOPSIS method, with criteria weights determined using the direct rating (DR) method. This process leads to the selection of the most suitable recovery configuration.

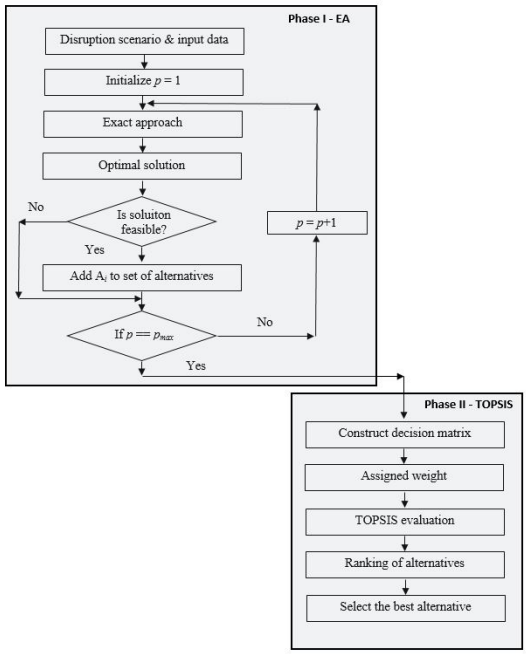


Figure 1: *Methodological framework of the proposed EA-TOPSIS approach.*

2.2. Mathematical formulation

This section presents the mathematical formulation of the considered problem. The notation used in the model is summarized in Table 1, listing all sets, parameters, and decision variables.

Sets	
P	set of locations of permanent facilities
T	set of potential locations of temporary facilities
F	set of all locations of facilities ($F = P \cup T$)
U	set of users represented by demand nodes (municipalities)
U_j	set of demand nodes preassigned to permanent facility j ($U_j \subseteq U, j \in P$) by territorial division
Parameters	
v_i	demand (number of users) at node i ($i \in U$)
p_{max}	the maximum number of temporary facilities that can be located
s_j	equals 1 if there is a permanent facility at location j, otherwise 0
r_j	equals 1 if a permanent facility at location j is closed, otherwise 0
d_{ij}	shortest distance between demand node i and facility j
Q_j	capacity of facility j
ρ	the maximum overcapacity ratio of the facility
β	the maximum tolerance of the difference in the overcapacity ratio between facilities
NAU	number of affected users in case of disruption
Decision variables	
x_j	$\begin{cases} 1, & \text{if the temporary facility is located at } j \ (j \in T) \\ 0, & \text{otherwise} \end{cases}$
y_j	$\begin{cases} 1, & \text{if the capacity of facility } j \text{ is exceeded} \\ 0, & \text{otherwise} \end{cases}$
z_j	overcapacity ratio of facility j ($j \in F$)

Table 1: *Notation used in the mathematical formulation.*

The mathematical formulation can be given in the following way:

$$\min Z = \frac{1}{NAU} \sum_{j \in F} \sum_{i \in U_j} d_{ij} q_{ij} \tag{1}$$

subject to:

$$\sum_{j \in T} x_j \leq p_{\max} \tag{2}$$

$$\sum_{j \in F} q_{ij} = v_i, \quad \forall i \in U \tag{3}$$

$$\sum_{\substack{l \in F \\ l \neq j}} q_{il} \leq v_i r_j, \quad \forall j \in P, \forall i \in U_j \tag{4}$$

$$q_{ij} \leq v_i (1 - r_j), \quad \forall i \in U, \forall j \in P \tag{5}$$

$$q_{ij} \leq v_i (s_j + x_j), \quad \forall i \in U, \forall j \in F \tag{6}$$

$$z_j \leq \rho y_j, \quad \forall j \in F \tag{7}$$

$$\sum_{i \in U} q_{ij} \geq Q_j(y_j + z_j), \quad \forall j \in F \tag{8}$$

$$\sum_{i \in U} q_{ij} \leq Q_j(1 + z_j), \quad \forall j \in F \tag{9}$$

$$z_j - z_l \leq \beta + \rho(2 - y_j - y_l), \quad \forall j, l \in F,\, j \neq l \tag{10}$$

$$x_j \in \{0, 1\}, \quad \forall j \in T \tag{11}$$

$$y_j \in \{0, 1\}, \quad \forall j \in F \tag{12}$$

$$q_{ij} \geq 0 \text{ and integer}, \quad \forall i \in U, \forall j \in F \tag{13}$$

$$z_j \geq 0, \quad \forall j \in F \tag{14}$$

The objective function (1) represents the average distance traveled by users affected by the disruption, and it should be minimized. This is achieved by redistributing the territorial assigned users from the closed facilities to the nearest non-closed facilities. Constraint (2) defines the maximum number of TFs that can be located. Constraints (3) to (6) apply to user flows. Constraints (3) ensure that all users receive the requested service. Constraints (4), (5), and (6) prevent the redistribution of users to a facility that is attacked, to a facility that does not exist, and to a temporary facility that is not located. Constraints (7) - (9) define the maximum allowed overcapacity ratio of the facility j. Constraints (10) also refer to the capacity of facilities, and they should achieve a balance between the overcapacity in the facilities, that is, they define the maximum allowed difference in the overcapacity ratio between facilities j and l. The decision variables in the model are defined over the following domains: binary variables x_j and y_j as specified in constraints (11) and (12), a non-negative integer variables q_{ij} as in constraints (13), and a non-negative continuous variables z_j as in constraint (14).

The proposed model is designed for universal applicability across different systems and incorporates fundamental criteria that are common to all of them. When applied to a specific system, additional criteria may be introduced, particularly cost-related components such as fixed costs for locating temporary facilities and travel costs generated by user movements. These elements can be incorporated through an appropriate modification of the objective function.

The model also includes user splitting and explicit overcapacity control. User splitting ensures that no demand remains unsatisfied, while overcapacity control prevents excessive load concentration at individual facilities. By balancing the overcapacity ratio among facilities, the model protects infrastructure resources and enhances operational sustainability. Such load balancing contributes to system stability, especially under disruption conditions, enabling a more resilient and flexible network response.

From a computational perspective, the resulting formulation is a mixed-integer linear programming (MILP) model. Due to the presence of binary and integer decision variables, the problem belongs to the class of NP-hard optimization problems.

2.3. Method for determining the number of temporary facilities

When solving the MILP model based on the proposed mathematical formulation, the number of TFs to be located must be determined in advance. The number of TFs can be determined in advance, based on available resources, or by applying the TOPSIS method after testing several scenarios regarding the number of located TFs. It should be noted that when applying only the EA, the maximum number of TFs to be located must be defined in advance. However, if the TOPSIS method is used, it is necessary to generate several solutions with different numbers of located TFs, based on which the best solution will be selected using the TOPSIS method. The

EA-TOPSIS approach selects the best solution from a set of generated feasible solutions based on defined criteria.

In certain situations, the proposed approach may be easier to implement than a full multi-objective optimization framework. This is particularly evident when the analyst intends to incorporate additional evaluation criteria that are difficult to formalize within a strict mathematical structure. Such criteria may include location suitability, accessibility, or other context-specific and partially subjective factors. In these cases, separating the generation of feasible solutions from their final evaluation allows greater modeling flexibility. Additional criteria can be introduced at the evaluation stage without reformulating the entire optimization model, which simplifies practical application and enhances adaptability to real-world decision-making environments.

The TOPSIS method is an MADM method used to select the best alternative from a set of alternatives [13]. In this paper, the TOPSIS method is applied to select the best solution from a set of solutions that represent a plan for locating TFs and redistributing users to PFs and TFs, after disruptions. The set of alternatives depends on the available number of TF that can be located, but also on whether the solution obtained by the EA approach is unique. In the case where all solutions are unique, the number of alternatives is equal to the maximum number of TF that can be located, increased by one ($p_{max}+1$) (the alternative when the number of TF is zero). The number of alternatives can be even larger when the solution obtained by applying the EA approach is not unique. In this situation, more optimal solutions for the defined number of TF locations could be identified and inserted into the set of alternatives. For the selection of the best solution, 5 criteria were generated, shown in Table 2. The first criterion refers to the average distance traveled by users affected by the disruption (Z). The second criterion (Q_{max}) represents the highest value of the overcapacity ratio in facilities. The third criterion refers to the largest difference in the overcapacity ratio between facilities ($Q_{max}-Q_{min}$), where Q_{min} represents the smallest value of the overcapacity ratio in facilities. The fourth criterion shows the number of facilities in which capacity was exceeded (N^{Q^+}), and the fifth criterion refers to the number of located TFs. It should be noted that all criteria are minimization type. Based on these five criteria, the solution is evaluated, and then the ranking is established.

Type of criteria	**Criteria**				
	Z [km]	Q_{max} [%]	$Q_{max}-Q_{min}$ [%]	N^{Q^+}	Number of TFs
max/min	min	min	min	min	min

Table 2: *Criteria for choosing the best solution using the TOPSIS method.*

Steps of applying the TOPSIS method [13]:

1. Normalization of the initial matrix, where:
   J – set of maximization criteria
   J' – set of minimization criteria
   A – set of all alternatives
   m – number of alternatives
   i – index of alternative
   n – number of criteria
   j – criterion index
   x_{ij} – element of the initial matrix (value of alternative i for criterion j)
   r_{ij} – element of the normalized matrix (normalized value of alternative i for criterion j)

$$r_{ij}=\frac{x_{ij}}{\sqrt{\sum_{k=1}^{m}x_{kj}^2}},\quad \forall i=1,\dots,m,\ \forall j=1,\dots,n \tag{15}$$

2. Weighting of the normalized matrix by weights:
   v_{ij} – element of the weighted matrix

$$v_{ij} = w_j \cdot r_{ij}, \quad i = 1, \ldots, m,\ j = 1, \ldots, n \tag{16}$$

3. Determination of ideal (A^*) and anti-ideal (A^-) solutions:

$$A^* = \left\{ \left(\max_i v_{ij} \mid j \in J \right), \left(\min_i v_{ij} \mid j \in J' \right) \;\middle|\; i = 1, \ldots, m \right\} = \{v_1^*, v_2^*, \ldots, v_n^*\} \tag{17}$$

$$A^- = \left\{ \left(\min_i v_{ij} \mid j \in J \right), \left(\max_i v_{ij} \mid j \in J' \right) \;\middle|\; i = 1, \ldots, m \right\} = \{v_1^-, v_2^-, \ldots, v_n^-\} \tag{18}$$

4. Determining the distance of alternatives from the ideal (S^*) and anti-ideal (S^-) solutions:

$$S_i^* = \sqrt{\sum_{j=1}^{n} \left(v_{ij} - v_j^* \right)^2}, \quad i = 1, \ldots, m \tag{19}$$

$$S_i^- = \sqrt{\sum_{j=1}^{n} \left(v_{ij} - v_j^- \right)^2}, \quad i = 1, \ldots, m \tag{20}$$

5. Calculating the relative closeness of alternatives to the ideal solution (C_i^*):

$$C_i^* = \frac{S_i^-}{S_i^* + S_i^-}, \quad i = 1, \ldots, m \tag{21}$$

6. Rank the alternatives according to their C_i^* values. C_i^* values range from 0 to 1. An alternative that has a C_i^* value closer to 1 should be ranked higher. In other words, the alternatives should be ranked from the highest to the lowest C_i^* value.

The EA-TOPSIS approach is applied to analyze the obtained solutions and to identify the best one from both user and system perspectives.

3. Computational study

The main aim of this study is to develop a universal approach applicable to different real-world systems. The implementation of the proposed approach in a specific system requires the definition of additional criteria by domain experts, as these may vary depending on the system's characteristics and priorities. For this reason, the testing in this study is conducted on hypothetical instances and relies on a fundamental set of criteria relevant to any practical setting. This enables the evaluation of the methodological properties of the approach while preserving its general applicability.

The hypothetical example covers a territory with 20 demand nodes (municipalities), where 2,000 users live. Users are aggregated into demand nodes so that each node has 100 users. The distances used in the model are calculated according to the Euclidean distance based on the (x, y) coordinates of the demand nodes shown in Table 3. In the observed hypothetical example, there are 5 permanent facilities whose locations are shown in Figure 2 (nodes 1, 5, 9, 13, and 17). The locations of the PFs and the distribution of users they serve according to the territorial division are shown in Figure 2.

Under normal operating conditions of the system, users must respect the territorial division, i.e., each user is preassigned to exactly one PF. There are 20 demand nodes (municipalities) in

the system. The PF is located in 5 municipalities and serves users from the region to which it belongs. The capacity of PFs is defined based on the number of users who gravitate towards it, i.e., in this example, the capacity of each PF is 400. The remaining 15 municipalities in which there is no PF represent potential locations for locating TFs in the case of a disruption. TF can be located next to facilities of a similar purpose in the selected municipality, which facilitates the establishment of TF due to the proximity of trained personnel, equipment, and other necessary resources for the operation of TFs. In this example, the capacity of TFs is defined as 50 % of the capacity of PFs (TFs capacity is 200).

Demand node	(x, y)	Demand node	(x, y)	Demand node	(x, y)	Demand node	(x, y)
1	(20, 80)	6	(60, 70)	11	(98, 90)	16	(45, 20)
2	(10, 70)	7	(45, 55)	12	(106, 65)	17	(85, 18)
3	(12, 90)	8	(42, 80)	13	(25, 30)	18	(69, 8)
4	(22, 95)	9	(73, 73)	14	(15, 10)	19	(67, 30)
5	(60, 70)	10	(94, 65)	15	(10, 40)	20	(102, 25)

Table 3: *Coordinates (x, y) of the demand nodes.*

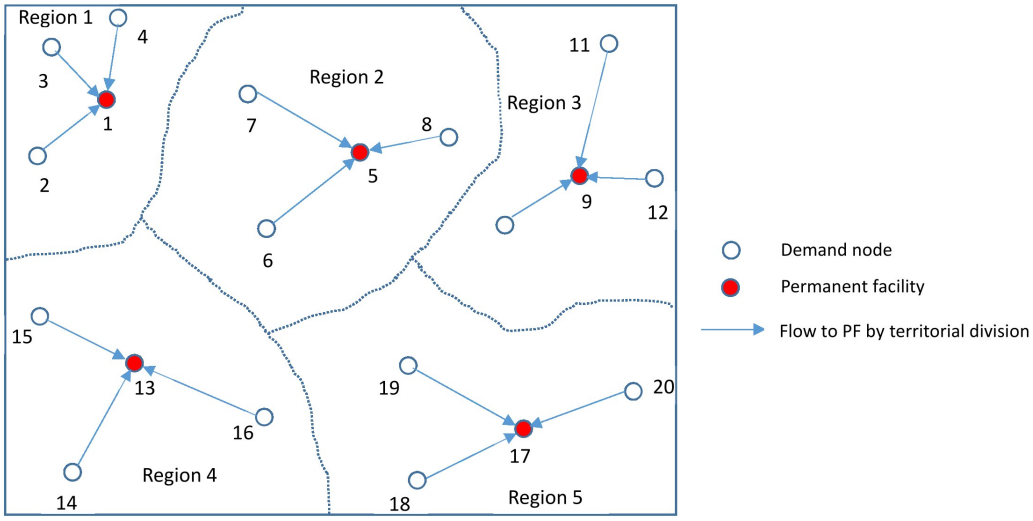


Figure 2: *The territorial division of the demand nodes by PFs and the potential location for TFs.*

To solve the proposed capacitated p-median model, we used ILOG CPLEX 12.1 on a 64-bit ACER computer with an Intel(R) Core (TM) i5 2.50 GHz processor and 8 GB of RAM. Python version 3.10.9 was used in Spyder environment version 5.4.5 to process data and search for solutions.

The model was tested for different variants of attacks on PF and TF locations. In all tested examples, a unique solution was obtained by applying the EA approach. Based on this, it can be concluded that the number of alternatives is equal to the available number of TF that can be located, increased by one (for the case when the temporary facilities are not located, i.e., $p = 0$). Due to the scope of the paper, an example of an attack on two PFs (1 and 5) is shown and explained in detail. The results of testing the developed approach in the case of an attack on PFs (1 and 5) are shown in Table 4 and illustrated in Figure 3. Table 4 shows the solutions obtained in 4 different situations, i.e., solutions obtained depending on the number of located

TFs (from 0 to 3). Column three shows the locations of the TFs. NAU denotes the number of users affected by the disruption in the case of an attack on PFs (1 and 5), and NAU is equal to 800 users in each scenario. Z denotes the value of the objective function, i.e., the average distance traveled by users affected by the disruption.

The total distance that users travel under normal conditions to obtain the requested service is 27,118 km, with the average distance traveled by users being 13.6 km. The disruption caused by the attack on two PFs (1 and 5) affected users in 8 demand nodes (1, 2, 3, 4, 5, 6, 7, and 8). In the event of an attack on the aforementioned two PFs, it is necessary to introduce 3 TFs so that users do not feel the consequences of the disruption in terms of traveled distance. In the scenario in which 3 TFs are introduced, the average distance traveled is 11.8 km, which is 1.8 km less than in normal conditions. Based on this, it can be concluded that in this example, by introducing 3 TFs, users would not feel the consequences of the disruption, from the perspective of the traveled distance. It can also be concluded that in this example, no more than 3 TFs should be located because locating more than 3 TFs would lead to unnecessary costs.

The highest value for $Z = 45.1$ km is obtained in the first scenario when zero TFs are located. In the case of locating only one TF, the value of Z decreases from 45.1 km to 30.8 km. In this scenario, it is best to locate the PF in Municipality 3. When locating two TFs, the objective function value is 17 km. The lowest value of the objective function is achieved by locating three TFs, with $Z = 11.8$ km.

Number of TFs	**Attacked PFs**	**Location of TFs**	***NAU***	***Z***
0		/		45.1
1	1 and 5	3	800	30.8
2		3 and 7		17.0
3		2, 4, and 6		11.8

Table 4: *Test results given by the EA approach in the case of the attack on the two PFs.*

Figure 3 shows the solutions in the case of attacking two PFs (1 and 5) and locating 0 TFs (Figure 3a), 1 TF (Figure 3b), 2 TFs (Figure 3c), and 3 TFs (Figure 3d). For each of the four TFs location variants, the redistribution of users from the 8 demand nodes affected by the attack on PFs 1 and 5 is shown. In the first example (Figure 3a), generating a redistribution plan without introducing temporary facilities is not possible due to capacity constraints (7)-(10). Figure 3a shows a user redistribution plan assuming that each user is served by the nearest active facility. According to the redistribution plan shown in Figure 3a, overcapacity in some facilities is 150%. This is proof that when generating a redistribution plan, it is necessary to take into account the facilities' capacities, which was done in the mathematical formulation proposed in this paper. Given that the attack on two PFs affected 800 users (45% of the total number of users in the system), PF and TF facilities must allow overcapacity in that ratio so that all users can be served ($\rho = 0.45$). Due to the constraints of the permitted overcapacity ratio of facilities to 45% and the difference in overcapacity ratio between facilities of 10% ($\beta = 0.1$), users from certain nodes have been redistributed to multiple facilities. The stated constraints are intended to ensure an even load on the active (non-closed) facilities. In Figures 3b, 3c, and 3d, the redistribution plan is defined respecting the capacity constraints. From the presented solutions, it can be seen that in the case of the introduction of 1 TF (Figure 3b), users from all regions feel the consequences of the disruption. In the case of the introduction of 2 TFs (Figure 3c), users from one region do not feel the consequences of the disruption, while in the case of the introduction of 3 TFs (Figure 3d), users from 2 out of 5 regions do not feel the consequences of the disruption. Of the four considered solutions, the solution of the introduction of three TFs (2, 4, and 6) provides minimal consequences for users after a disruption, but the highest cost for the system.

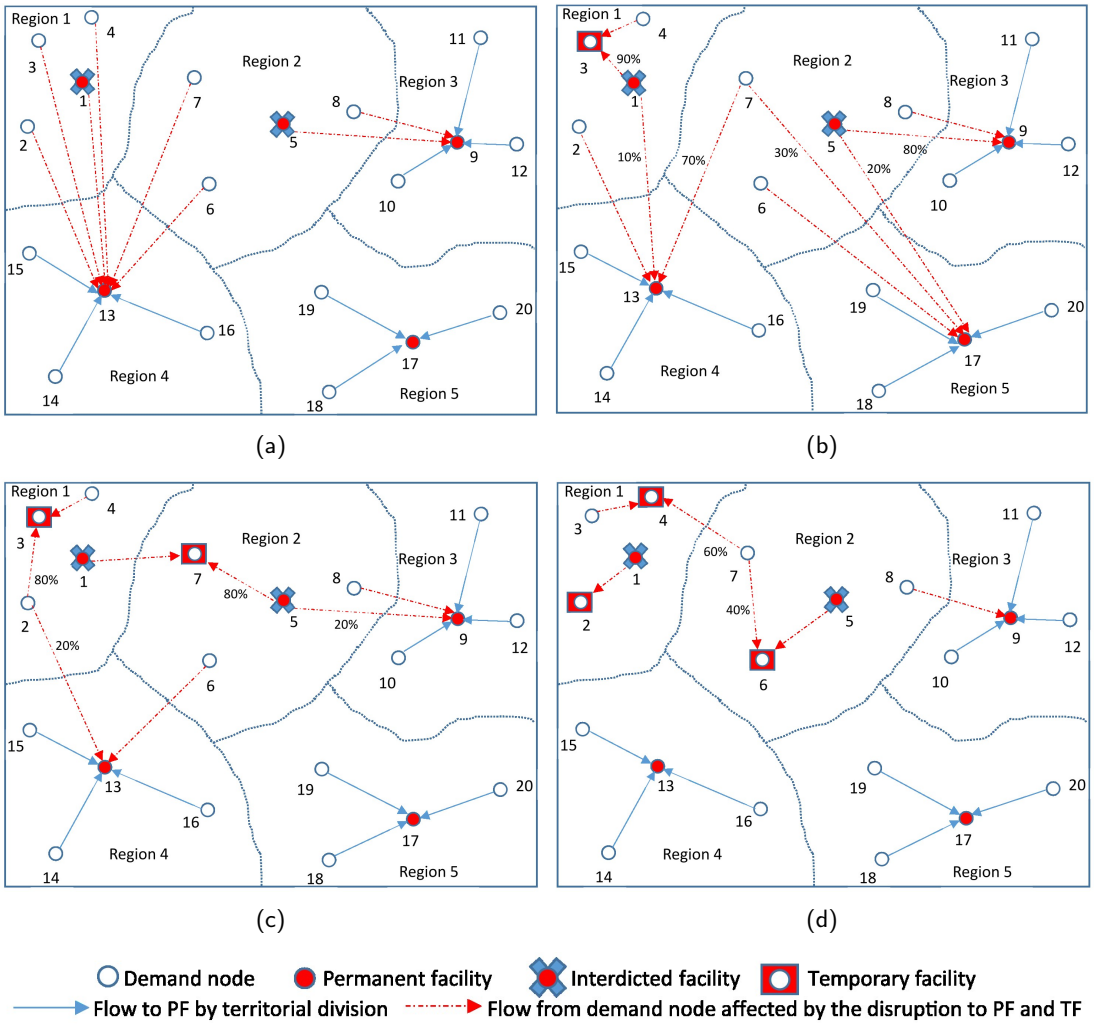


Figure 3: *Examples of attacks of two PFs and redistribution of users in case of location: a) zero TF; b) one TF; c) two TFs; d) three TFs*

To obtain the final solution of the considered problem, it is necessary to select the best scenario, i.e., the number of located TFs. The selection of the best solution can be performed either through expert analysis of the results or by applying the MADM method. In order to select the best solution, it is proposed to combine the exact approach for generating high-quality feasible solutions and the TOPSIS method for selecting the best solution. We call this approach EA-TOPSIS. The EA-TOPSIS method was developed in order to reduce subjectivity in selecting the best solution as much as possible. The input data required for applying the EA-TOPSIS approach to select the final solution in the case of an attack on the PF 1 and 5 are shown in Table 5. Table 5 shows the solutions obtained in 3 different situations (3 alternatives), that is, the solutions obtained depending on the number of located TFs (from 1 to 3). For the scenario where there are no temporary facilities (TF=0), the exact approach provides an infeasible solution and therefore cannot be included in the set of alternatives. The second to sixth columns in Table 5 show the criteria for choosing the best solution, which are explained in detail in Section 2.3. The last row of Table 5 shows the weighting coefficients that define the importance of each of the criteria.

The weighting coefficients were obtained by applying the DR method. When applying this method, the decision maker has a fixed number of points available, in this case, 100. Changing (correcting) the number of points assigned to one of the criteria is reflected in changing the number of points assigned to one or more of the others, i.e., their redistribution is carried out. The average distance (Z) traveled by users affected by the disruption is assigned the highest weight (40), indicating that it is the most important criterion. The number of located TFs is the second most important criterion. The criteria related to overcapacity ratio in facilities (the highest overcapacity ratio (Q_{max}) and the difference between the highest and lowest overcapacity values ($Q_{max} - Q_{min}$)) are in third place in importance. The least important criterion is the number of facilities in which capacity has been exceeded. The values assigned to the weight coefficients need to be normalized. The normalization of the weight coefficient values was performed using the following formula:

$$w_j = \frac{W_j}{\sum_{j=1}^{n} W_j}, \quad \forall j \in J \cup J' \tag{22}$$

W_j weighting coefficient of criterion $j \in J \cup J'$
w_j normalized weighting coefficient of criterion $j \in J \cup J'$

In Table 5, the last two columns show the output results obtained by the TOPSIS method. C_i^* represents the relative closeness of alternatives to the ideal solution, based on which the alternatives are ranked. The last column shows the obtained rank of alternatives, which indicates that the best solution is the location of 2 TFs in the Municipality 3 and 7. The EA-TOPSIS approach shows that locating TFs in these municipalities achieves the best balance between the number of located TFs and the average distance traveled by users affected by the disruption. The application of the EA-TOPSIS approach reduces the average traveled distance of the users by 28.1 km and the maximum overcapacity ratio in facilities from 150% to 40%, compared to the case without TFs location. Under normal operating conditions of the system, users travel on average only 3.4 km less to get the service. Respecting facility overcapacity limits and balancing in overcapacity prevents overloading of facilities in the system and ensures efficient and sustainable operation of the system. The solution generated by applying the developed EA-TOPSIS approach is directly applicable in practice. The developed method has the possibility of application in various systems in which there are disruptions in the operation of facilities (system of health facilities, system of police station facilities, system with several warehouses, system of facilities for water supply, etc.).

Input						**Output**	
Alternatives	**Criteria**					**Results**	
	$\boldsymbol{Z}$ **[km]**	$\boldsymbol{Q_{max}}$ **[%]**	$\boldsymbol{Q_{max} - Q_{min}}$ **[%]**	$\boldsymbol{N^{Q^+}}$	**Number of THs**	$\boldsymbol{C_i^*}$	**Rank**
1	30.8	45	7.5	4	1	0.39	3
2	**17.0**	**40**	**10**	**4**	**2**	**0.64**	**1**
3	11.8	30	10	3	3	0.61	2
max/min		min	min	min	min		
$\boldsymbol{W_i}$	40	15	15	5	25		

Table 5: *The values of solutions per criteria and the final rank.*

Sensitivity analysis of the weight coefficients was conducted to evaluate the robustness of the obtained ranking. For each criterion, the weight was independently increased and decreased within the interval of −20% to +20%. The weights were determined using the DR method, where the total sum equals 100. When a weight was changed, the remaining weights W_j ($j \neq k$)

were rescaled proportionally according to their original values to preserve the total sum. This was achieved using the following formula.

$$W_j' = (100 - W_k')\frac{W_j}{\sum_{\substack{j \in J \cup J' \\ j \neq k}} W_j}, \quad \forall j \in J \cup J',\ j \neq k \tag{23}$$

W_k' weight of criterion k after being changed by a specified percentage
W_j' $(j = 1, \ldots, n;\ j \neq k)$ weights of the remaining criteria adjusted proportionally to reflect the change in W_k

No rank reversal was observed for W_2, W_3, and W_4 throughout the perturbation interval. Rank changes occurred only for W_1 and W_5. As shown in Figure 4, the sensitivity analysis for these two criteria is illustrated separately. Changes in the relative closeness of alternatives C_i^* for different values of W_1 are presented in Figure 4a, and for W_5 in Figure 4b. For W_1, the ranking remained unchanged for variations from -20% to $+15\%$, and for W_5 from -15% to $+20\%$. Changes occurred only outside of these ranges. An increase of W_1 by 16% or more results in Alternative 3 becoming the top-ranked solution. Whereas a decrease of W_5 by 16% or more leads to a rank reversal between Alternatives 2 and 3. Since these changes occur only under relatively large deviations in weight, the adopted weight structure can be considered robust, and the resulting decision stable.

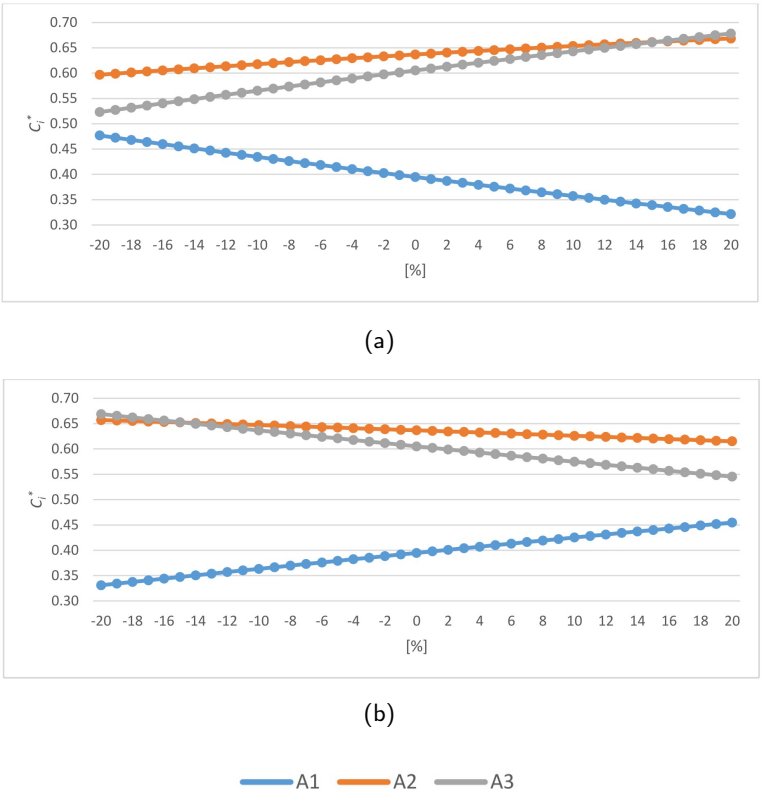


Figure 4: *Sensitivity analysis of the relative closeness of alternatives (C_i^*) with respect to weight variations (-20% to $+20\%$): a) W_1; b) W_5.*

4. Conclusion

In recent years, disruptions, such as the emergence of viruses, terrorist attacks, or earthquakes, have become increasingly frequent. Such events may interrupt the functioning of systems consisting of multiple facilities to which users are allocated, highlighting the need for models that ensure the system's operation in conditions of disruptions.

This paper discusses the problem of determining the location of temporary facilities and the redistribution of users to permanent and temporary facilities after disruptions. A mathematical formulation is proposed for this problem. In order to determine the final number of located TFs, the TOPSIS method was applied in combination with the exact approach. For the application of the TOPSIS method, 5 criteria were selected based on which the best solution from the set of solutions is selected. One criterion refers to the average traveled distance of users affected by the disruption, three criteria refer to the overcapacity ratio and balancing of the overcapacity ratio in PF and TF, and one criterion refers to the number of located TFs. The solution obtained by applying the EA-TOPSIS approach achieves a balance between the goals of system management and users.

The proposed EA-TOPSIS approach is tested on a hypothetical example of 20 demand nodes and 5 PFs. The model was validated by testing against different variants of PF attacks and by locating different numbers of TFs. The paper highlights an example of an attack on two PFs and locations from 0 to 3 TFs. Locating TF can achieve great savings in terms of the average distance traveled by users. In the analyzed example, it can be reduced by 31.7% by introducing one TF, or even 62.3% by introducing 3 TFs, compared to the case when no TFs are introduced. Also, the capacity constraints in the mathematical formulation ensure a more even redistribution of users. Without capacity constraints in facilities, the overcapacity ratio can reach up to 150%, while with capacity constraints, the overcapacity ratio is limited to 45%.

Overcapacity enables the service of affected users by allowing both TFs and PFs to adapt to increased demand, with capacity allowed to be exceeded up to a predefined ratio. Balancing the overcapacity ratio across facilities distributes users more evenly, preventing extreme overload at individual sites and supporting a more efficient and sustainable system operation under disruption conditions. The capacity constraints enable the direct application of the developed approach. The proposed approach is flexible and can be adapted to specific systems by incorporating additional criteria.

The study also has some limitations. The computational experiments were performed on small-sized instances rather than real-world cases. Furthermore, the approach involves a degree of subjectivity in the selection of the MADM method, the evaluation criteria, and their weights. Additional limitations may arise from the problem size that can be solved exactly using optimization software, such as CPLEX.

Overall, the study provides a theoretical contribution by integrating an exact approach with the TOPSIS method for multi-criteria decision-making under disruption. It further offers practical implications through a flexible framework that supports system managers in enhancing operational resilience.

There are several possible directions for future research, such as: testing the proposed solution approach on real examples (for example, in the case of disruptions in the healthcare system), including additional constraints and criteria (for example, available budget and the cost of locating temporary facilities), applying other MADM techniques, defining the number of temporary facilities that need to be located using fuzzy logic systems or some artificial intelligence methods.

References

[1] Aliakbarian, N., Dehghanian, F. and Salari, M. (2015). A bi-level programming model for protection of hierarchical facilities under imminent attacks. Computers and operations research, 64,

210-224. doi: 10.1016/j.cor.2015.05.016

[2] Carnero Quispe, M. F., Chambilla Mamani, L. D., Yoshizaki, H. T. Y. and Brito Junior, I. D. (2025). Temporary Facility Location Problem in Humanitarian Logistics: A Systematic Literature Review. Logistics, 9(1), 42. doi: 10.3390/logistics9010042

[3] Caunhye, A. M., Nie, X. and Pokharel, S. (2012). Optimization models in emergency logistics: A literature review. Socio-economic planning sciences, 46(1), 4-13. doi: 10.1016/j.seps.2011.04.004

[4] Cavdur, F., Kose-Kucuk, M. and Sebatli, A. (2016). Allocation of temporary disaster response facilities under demand uncertainty: An earthquake case study. International Journal of Disaster Risk Reduction, 19, 159-166. doi: 10.1016/j.ijdrr.2016.08.009

[5] Durmaz, Y. G. and Bilgen, B. (2020). Multi-objective optimization of sustainable biomass supply chain network design. Applied Energy, 272, 115259. doi: 10.1016/j.apenergy.2020.115259

[6] Gharib, Z., Tavakkoli-Moghaddam, R., Bozorgi-Amiri, A. and Yazdani, M. (2022). Post-disaster temporary shelters distribution after a large-scale disaster: An integrated model. Buildings, 12(4), 414. doi: 10.3390/buildings12040414

[7] Gul, M. and Guneri, A. F. (2015). Simulation modelling of a patient surge in an emergency department under disaster conditions. Croatian Operational Research Review, 6(2), 429-443. doi: 10.17535/crorr.2015.0033

[8] Holguín-Veras, J., Jaller, M., Van Wassenhove, L. N., Pérez, N. and Wachtendorf, T. (2012). On the unique features of post-disaster humanitarian logistics. Journal of Operations Management, 30(7-8), 494-506. doi: 10.1016/j.jom.2012.08.003

[9] Janáček, J. and Kvet, M. (2020). Discrete self-organizing migration algorithm and p-location problems. Croatian Operational Research Review, 11(2), 241-248. doi: 10.17535/crorr.2020.0019

[10] Liu, C. F. and Mostafavi, A. (2023). An equitable patient reallocation optimization and temporary facility placement model for maximizing critical care system resilience in disasters. Healthcare Analytics, 4, 100268. doi: 10.1016/j.health.2023.100268

[11] ReVelle, C.S. and Swain, R.W. (1970). Central facilities location. Geographical Analysis, 2(1), 30–42. doi: 10.1111/j.1538-4632.1970.tb00142.x

[12] Teixeira, J. C. and Antunes, A. P. (2008). A hierarchical location model for public facility planning. European Journal of Operational Research, 185(1), 92-104. doi: 10.1016/j.ejor.2006.12.027

[13] Teodorović, D. and Nikolić, M. (2021). Quantitative Methods in Transportation. CRC Press, Taylor and Francis Group. doi: 10.1201/9780429286919

[14] Yanıkoğlu, B. and Kizilay, D. (2024). A Mathematical Model Using AHP-Based Parameters for Solving Location-Allocation Problem. Journal of Optimization and Supply Chain Management, 1(2), 110-119. doi: 10.22034/ISS.2024.8212.1016

A discrete-time infinite-server batch arrival queue with application to serverless computing

Veena Goswami[1,*]

[1] *School of Computer Applications, Kalinga Institute of Industrial Technology, Bhubaneswar, India*
E-mail: ⟨veena_goswami@yahoo.com⟩

Abstract. Serverless computing systems behave like an infinite server queue, dynamically expanding the number of instances to handle incoming batch tasks. In serverless computing, to leverage its scalability, efficiency, and cost-effectiveness, it is essential to apply an infinite server queue with batch task arrival. The model is used to estimate resource usage and to optimize cost-performance trade-offs. This paper analyzes a discrete-time infinite-server batch/single-arrival queue with application to a serverless computing system. We obtain the probability-generating function of the system size, which characterizes a functional equation; we then solve it recursively to find state probabilities at various epochs and some performance measures. Unlike existing discrete-time infinite-server models, this paper explicitly compares the late-arrival system with delayed access and the early-arrival system policies in the context of batch arrivals, and establishes convergence to their continuous-time counterparts. The model's numerical results have been presented in graphs and tables to illustrate the impact of batch-size distributions and system parameters.

Keywords: bulk arrival, discrete-time queueing, infinite server, scalability, serverless

Received: January 4, 2026; accepted: March 12, 2026; available online: July 8, 2026

DOI: 10.17535/crorr.2026.0025

Original scientific paper.

1. Introduction

Serverless architecture offers numerous advantages, including lower operating costs, enhanced application performance, and multi-language support for developers. Due to these benefits, serverless computing is gaining popularity among companies seeking an affordable development and implementation model. Infinite-server bulk-arrival queueing models have intriguing applications in serverless computing, particularly for handling unpredictable or high-volume workloads. Serverless computing has emerged as a notable framework in cloud-native application development, primarily due to its underlying elasticity and cost efficiency [7, 14]. This paradigm simplifies runtime resource management, leading to its widespread adoption in Function-as-a-Service (FaaS) offerings, such as AWS Lambda [28]. [26] discussed the challenge of predicting the performance of serverless functions in a dynamic environment. [27] and [25] noted that the performance variability, which demonstrates differing end-to-end response latencies across identical function invocations, is an unconsidered problem within the research community.

Serverless function performance is crucial for both practitioners seeking to optimize their applications and cloud providers aiming to improve service quality [15]. It necessitates robust methodologies for performance evaluation, often involving empirical studies to determine the number of repetitions required for a serverless function with specific inputs, accounting for observed performance fluctuations [27]. Instead of coming in one at a time, several requests

*Corresponding author.

arrive simultaneously. There is never a waiting queue because each incoming request is promptly assigned to a server. Serverless platforms and other elastic resources are used in this model system. These models are ideal for examining systems with nearly infinite scalability, such as serverless operations or cloud-based microservices.

Bulk arrival models facilitate the simulation and forecasting of serverless systems' behavior, which often encounter unexpected spikes in requests. Analysis of latency, throughput, and resource usage without bottlenecks is possible with infinite server queues. Many researchers studied the continuous-time infinite server queues with many alterations like, overlap times [18], in a semi-Markovian environment [8], system's additional tasks and impatient customers [1], phase-type arrivals [20], batch arrivals [21, 23], batch arrivals and dependent service times [19]. The transient solution of the Markovian multi-server queue with balking and catastrophes has been discussed in [16].

The arrivals and departures may co-occur at a slot's border epoch in discrete-time queueing system. Arrival-first (AF) and departure-first (DF) ruling strategies, sometimes mentioned as the early arrival system (EAS) and late arrival system with delayed access (LAS-DA), respectively, can deal with rules in the case of simultaneity, see [12, 13]. In the literature, multiserver discrete-time models with various variations have been discussed: with late and early arrival system [9], optimal control of queue [2], baling behavior [10], batch arrival [22], renewal input [3], batch arrival with renewal input [5], batch service [11].

The discrete-time batch arrival infinite-server queueing model mapping enables the performance analysis of serverless systems to predict scalability, optimize resource allocation, and model bursty, event-driven workloads. The infinite-server assumption holds well below platform concurrency limits, making it ideal for analyzing short-term auto-scaling behavior. [17] analyzed the discrete-time infinite-server queueing system with batch arrivals and bulk service, formulated a functional equation for the system occupancy, and solved it by recursion rather than by the more common transform or matrix-analytic methods. [4] examined the discrete-time infinite-server bulk-arrival renewal input queue with geometrically distributed service times. They derived both the transient and steady-state distributions of the model's system size. [6] analyzed continuous-time infinite-server batch arrival queues with Poisson arrival epochs, handling both stationary and nonstationary arrival rates and general service distributions. This study extends [4] work by switching from discrete-time geometric service models to continuous-time general service distributions.

Serverless systems, by their features, are demand-driven and event-driven. Because system actions often occur in discrete steps or in response to discrete events, continuous-time models are less appropriate. Analysis of concurrency constraints, burst handling, and scaling dynamics is made easier by representing system evolution at slot boundaries. For the study of systems with batch arrivals, synchronization effects, and time-slotted control strategies, discrete-time queueing models offer manageable frameworks. While several studies address infinite-server queues with batch arrivals, fewer works consider discrete-time formulations under different arrival–departure conventions, which are particularly relevant for slot-based serverless platforms. In this article, we consider a discrete-time batch/single arrival infinite-server queueing system to enhance the effectiveness of serverless computing. The interarrival and service times are assumed to be geometrically distributed. We find the state probabilities at arbitrary and outside observer's observation epochs for both the late arrival system with a delayed access and the early arrival system by the recursion method. Particular cases are examined by assuming various probability distributions for the arrival batch sizes. Numerical experiments of the system have been examined in the form of tables and graphs. It is established that, in the limiting case, the results of this paper tend to the continuous-time counterpart.

This paper is structured as follows. Section 2 shows a serverless computing system architecture as a $Geo^X/Geo/\infty$ queueing system. Section 3 analyzes the $Geo^X/Geo/\infty$ queue by two methods for the late arrival system with a delayed access and the early arrival system policy.

In Section 4, we discuss exceptional cases by assuming various probability distributions for the arrival batch sizes and Section 5 presents the analysis of the $Geo/Geo/\infty$ queue. Section 6 illustrates the numerical results, and Section 7 concludes the paper.

2. System model

The cloud provider can dynamically allocate machine resources through a cloud architecture called "serverless computing". Due to their elasticity, serverless architectures can effectively manage batch arrivals and are inherently scalable. With a sufficient number of servers, serverless computing can handle batch arrivals as follows.

Automated Adjustment: Platforms without servers automatically allocate resources to meet incoming tasks. AWS Lambda system scales out in response to event volume and can manage thousands of concurrent executions. Google Cloud Functions automatically scales to handle the volume of incoming requests, ensuring each request is processed promptly. Azure Functions are suitable for batch processing; they scale out to handle multiple concurrent executions.

Event-driven invocation: Serverless functions are appropriate for batch job processing since a variety of events, such as scheduled tasks, message queues, and data processing pipelines, can trigger them.

Economy of cost: Serverless platforms bill according to the resources and real execution time, such as pay-per-use and no idle costs.

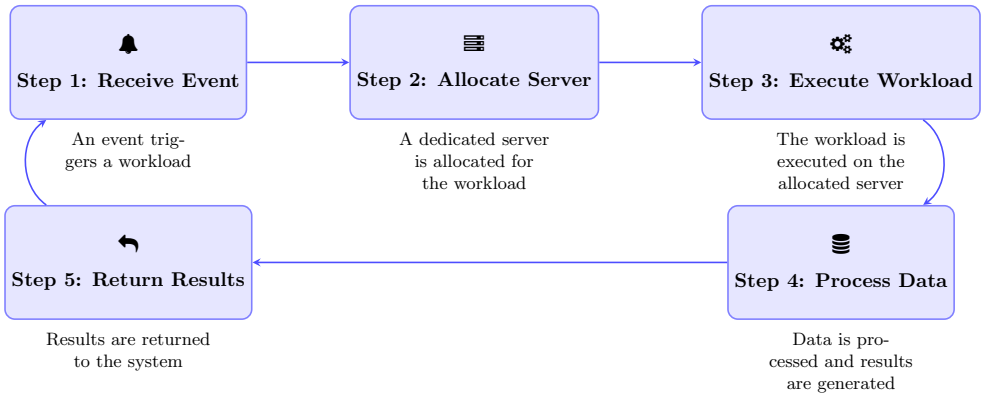


Figure 1: *Serverless architecture cycle.*

Serverless architecture processes bursty, event-driven workloads with automatic, near-infinite scale. The core principle is that every unit of work (job in a batch) gets its own dedicated execution environment (server) without waiting for previous jobs to finish. Fig. 1 depicts the serverless architecture cycle. Here is a visual representation of the core architecture.

Batch Arrival Generator (Event Sources): Here, batch arrivals originate and generate events that trigger the system, generally in bursts.

Queueing System (Trigger and Buffer Layer): After receiving the batch, this layer sends it to the compute layer. It decouples components and controls flow.

Infinite Server Farm (Compute Layer): The infinite server model consists of stateless, ephemeral compute services that scale on demand.

Shared State and Orchestration (Backing Services): The compute instances depend on managed services for durability and coordination because they are isolated and stateless.

Fig. 2 illustrates the flowchart of the serverless computing. Let us imagine a scenario where a large number of images need to be uploaded to a cloud storage bucket and then processed. Here, we consider an infinite server queueing model with packets arriving in batches. An

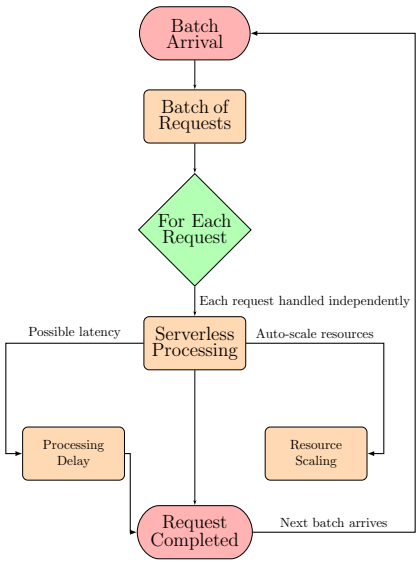


Figure 2: *Flowchart of serverless computing.*

infinite number of servers are ready to handle new assignments; a server is assigned to each task as soon as it arrives; every task is dealt with concurrently; and since there are always enough servers to manage the load, tasks never wait in a queue. This idea can be leveraged to handle such workloads in serverless computing by capitalizing on its scalability and event-driven architecture. By dynamically increasing or decreasing the number of instances to accommodate incoming batch tasks, serverless computing platforms can behave similarly to an infinite server queue. The discrete-time infinite-server batch-arrival model may not adequately explain systems that operate under capacity caps, burst-induced provisioning delays, or severe resource coupling, even while it is suitable for lightly loaded, unconstrained, and brief function executions.

3. Analytical model

We consider a discrete-time batch arrival queue wherein requests arrive in batches of random size Y with probability mass function (pmf) $g_j = P(Y = j),\ j \geq 1$, probability generating function (pgf) $G(z) = \sum_{j=1}^{\infty} g_j z^j$, $|z| \leq 1$ and mean batch size $\overline{g}$. The inter-arrival times A are independent and geometrically distributed with pmf $P(A = n) = \lambda\bar{\lambda}^{j-1}\ \ 0 < \lambda < 1,\ j \geq 1$. The batch sizes are independent of inter-arrival times. With pmf $P(S = n) = \mu\bar{\mu}^{j-1},\ 0 < \mu < 1,\ j \geq 1$ and mean service time $1/\mu$, the service time S of each of the infinite servers is independent and geometrically distributed. Additionally, given that there are k busy servers, the probability that ℓ servers finish service in the following interval is provided by

$$\Psi(\ell \mid k) = \begin{cases} \binom{k}{\ell}\mu^{\ell}\bar{\mu}^{k-\ell} & 0 \leq \ell \leq k, \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Since servers are infinite, stability does not need balancing arrival and service rates, unlike finite-server queues, because an infinite number of servers prevents queue formation. Rather, stability is ensured provided that both the mean service time and the mean arrival rate are finite. As it is assumed that the arrival batches and service times are independent and identically distributed throughout slots, the system dynamics are time-homogeneous. Thus, after

the system has run long enough, the process becomes a stationary stochastic process. Performance evaluation in real-world applications, such as serverless computing platforms, usually focuses on long-term average behavior rather than short-term startup impacts. The workload produced by large user populations can often be statistically consistent over sufficiently long observation periods. The steady-state study offers useful estimates of anticipated concurrency levels, resource usage, and system capacity requirements under these circumstances.

We study the system under both policies, the late arrival system with delayed access (LAS-DA) and the early arrival system (EAS). Assume that the time axis is denoted as $0, 1, 2, \ldots, m$, and that it is slotted into equal-length intervals with a slot length of unity. A thorough explanation of these ideas has previously been provided in several places [12, 13]. According to the LAS-DA policy, arrivals occur at the end of the slot and departures at the start. Service cannot be provided to newly arrived clients during the same time slot. In the EAS policy, arrivals happen at the start of the slot, and departures occur at its end. If the server is idle, arrivals may receive service immediately. Systematic delay discrepancies arise from differences in event ordering between LAS-DA and EAS policies. Queueing behavior is further influenced by batch-size variability, with greater variation leading to longer lineups and delays. The service can begin immediately in EAS policy, such as digital processors or packet-switched networks. But in the case of LAS-DA policy processing, decisions occur at slot boundaries, such as synchronized or gated systems.

3.1. $Geo^X/Geo/\infty$ queue with LAS-DA policy

In LAS-DA, prospective arrivals take place in the interval $(m-, m)$, whereas prospective departures occur in the period $(m, m+)$. More precisely, different periods of time during which events take place are shown in Fig. 3.

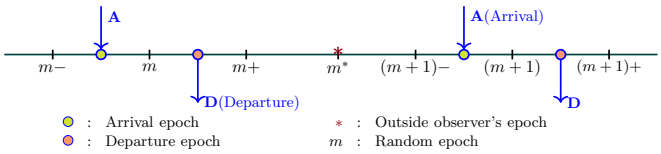


Figure 3: *Flowchart of serverless computing.*

Define the probability as $P_j(m-) = P\{j \text{ requests in the system at time } m-\}$, $j \geq 0$. Relating the system's states at two successive epochs $m-$ and $(m+1)-$, and using the notation $P_j(m)$ for $P_j(m-)$, we obtain

$$P_0(m+1) = \sum_{i=0}^{\infty} \bar{\lambda}\Psi(i|i)P_i(m), \tag{2}$$

$$P_j(m+1) = \sum_{i=j}^{\infty} \bar{\lambda}\Psi(i-j|i)P_i(m) + \sum_{k=0}^{\infty}\sum_{i=k}^{j+k-1} \lambda g_{j-i+k}\Psi(k|i)P_i(m), \quad j \geq 1. \tag{3}$$

Assuming that steady state exists, let $P_n = \lim_{m\to\infty} P_n(m)$, $n = 0, 1, \ldots$ and define pgf as $P(z) = \sum_{j=0}^{\infty} P_j z^j$. From (2) and (3), in steady-state we have

$$P_0 = \sum_{i=0}^{\infty} \bar{\lambda}\Psi(i|i)P_i, \tag{4}$$

$$P_j = \sum_{i=j}^{\infty} \bar{\lambda}\Psi(i-j|i)P_i + \sum_{k=0}^{\infty}\sum_{i=k}^{j+k-1} \lambda g_{j-i+k}\Psi(k|i)P_i, \quad j \geq 1. \tag{5}$$

Multiplying appropriate powers of z in Eqs. (4) and (5), and summing over j, after simplification, we have

$$P(z) = \left(\bar{\lambda} + \lambda G(z)\right) P\left(\mu + \bar{\mu} z\right), \tag{6}$$

which is a functional equation. From Eq. (6), using power series expansion around $z = 1$, we obtain

$$P(z) = \sum_{\ell=0}^{\infty} \frac{h_\ell}{\ell!} \, (z-1)^\ell, \tag{7}$$

where h_ℓ is given by the following recursive formula

$$h_0 = 1, \tag{8}$$

$$h_\ell = \frac{\lambda}{1-\bar{\mu}^\ell} \sum_{k=0}^{\ell-1} \binom{\ell}{k} \bar{\mu}^k \; h_k \; G^{(\ell-k)}(1), \quad \ell \geq 1, \tag{9}$$

$$= \frac{\lambda \ell!}{1-\bar{\mu}^\ell} \sum_{k=0}^{\ell-1} \bar{\mu}^k \frac{h_k}{k!} \sum_{j=\ell-k}^{\infty} \binom{j}{\ell-k} g_j, \quad \ell \geq 1, \tag{10}$$

and $G^{(n)}(1) = \frac{d^n}{dz^n} G(z)|_{z=1}$. Calculating the n-th ($n = 1, 2, \ldots$) derivatives on both sides of Eq. (6) and setting $z = 1$ yields the recursive formula mentioned above.From Eq. (7), we have

$$\sum_{\ell=0}^{\infty} P_\ell z^\ell = \sum_{\ell=0}^{\infty} z^\ell \sum_{n=\ell}^{\infty} (-1)^{n-\ell} \binom{n}{\ell} \frac{h_n}{n!}. \tag{11}$$

Thus,

$$P_\ell = \sum_{n=\ell}^{\infty} (-1)^{n-\ell} \binom{n}{\ell} \frac{h_n}{n!}, \quad \ell \geq 0. \tag{12}$$

Remark 1. *Taking inter-arrival and service times as exponentially distributed, with mean arrival and mean service rates γ and η, respectively. Divide the time axis into equal intervals $\Delta > 0$ such that $\lambda = \gamma\Delta$, $\mu = \eta\Delta$, and $\rho = \frac{\gamma\Delta}{\eta\Delta} = \rho_1$, where Δ is appropriately small. Setting $\lambda = \gamma\Delta$, $\mu = \eta\Delta$, and taking the limit as $\Delta \to 0$, we have*

$$h_\ell = \lim_{\Delta \to 0} \frac{\gamma\Delta \; \ell!}{1-(1-\eta\Delta)^\ell} \sum_{k=0}^{\ell-1} (1-\eta\Delta)^k \frac{h_k}{k!} \sum_{j=\ell-k}^{\infty} \binom{j}{\ell-k} g_j \tag{13}$$

$$= \frac{\gamma(\ell-1)!}{\eta} \sum_{k=0}^{\ell-1} \frac{h_k}{k!} \sum_{j=\ell-k}^{\infty} \binom{j}{\ell-k} g_j, \tag{14}$$

substituting h_ℓ in Eq. (12) *it tends to $M^X/M/\infty$ queue.*

Alternatively, we may solve Eq. (6) recursively, as shown below. Assume $H(z) = \left(\bar{\lambda} + \lambda G(z)\right)$. Thus,

$$P(z) = H(z) \cdot P\left(\mu + \bar{\mu} z\right), = H(z) \cdot H\left(\mu + \bar{\mu} z\right) \cdot P\left(\mu + \mu\bar{\mu} + \bar{\mu}^2 z\right) \tag{15}$$

After the nth recursion, we have

$$\begin{aligned} P(z) = & H(z) \cdot H\left(\mu + \bar{\mu} z\right) \ldots H\left(\mu + \mu\bar{\mu} + \mu\bar{\mu}^2 + \cdots + \mu\bar{\mu}^{n-1} + \bar{\mu}^n z\right) \\ & \cdot P\left(\mu + \mu\bar{\mu} + \mu\bar{\mu}^2 + \cdots + \mu\bar{\mu}^n + \bar{\mu}^{n+1} z\right). \end{aligned} \tag{16}$$

As $0 < \mu \leq 1$, so $0 \leq \bar{\mu} < 1$, when $n \to \infty$, in Eq. (16) the last factor is equal to one, because

$$\begin{aligned} &\lim_{n\to\infty} P\left(\mu + \mu\bar{\mu} + \mu\bar{\mu}^2 + \cdots + \mu\bar{\mu}^n + \bar{\mu}^{n+1}z\right) \\ &= \lim_{n\to\infty} P\left(\mu\left(1 + \bar{\mu} + \bar{\mu}^2 + \cdots + \bar{\mu}^n\right)\right) = P(1) = 1. \end{aligned} \tag{17}$$

Therefore, as $n \to \infty$, Eq. (16) reduces to

$$P(z) = H(z)\cdot H\left(\mu + \bar{\mu}z\right)\cdots H\left(\mu + \mu\bar{\mu} + \mu\bar{\mu}^2 + \cdots + \mu\bar{\mu}^{n-1} + \bar{\mu}^n z\right) \tag{18}$$

$$= H(z)\prod_{j=1}^{\infty} H\left(\sum_{i=1}^{j}\mu\bar{\mu}^{i-1} + \bar{\mu}^j z\right) = H(z)\prod_{j=1}^{\infty} H\left(1 - \bar{\mu}^j + \bar{\mu}^j z\right). \tag{19}$$

It is apparent from Eq. (19) that the service rate and the PGF of packet arrival batch size entirely describe the PGF of the system.

3.1.1. Outside observer's observation epoch

The outside observer's observation epoch is more likely to fall in the interval $(m+, (m+1)-)$. Since nothing happens (no arrival and no departure) in this interval, the outside observer's distribution P_n^o will be equal to the distribution of the number in the system noticed at $(m+1)-$ and hence $P_n^o = P_n$.

3.2. $Geo^X/Geo/\infty$ queue with EAS

As was previously mentioned, the order in which packets arrive and depart around a slot boundary varies between the two policies. In LAS-DA, prospective arrivals take place in the period $(m-, m)$, whereas prospective departures occur in the period $(m, m+)$. More precisely, different periods of time during which events take place are shown in Fig. 4.

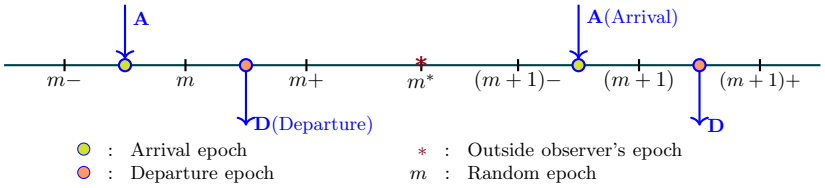


Figure 4: *Flowchart of serverless computing.*

$$Q_j = \bar{\lambda}\sum_{i=j}^{\infty} c(i-j|i)Q_i + \sum_{k=1}^{\infty}\sum_{i=j-k}^{\infty} \lambda g_k c(i-j+k|i+k)Q_i, \quad j \geq 0. \tag{20}$$

Multiplying Eq. (20) by appropriate powers of z and summing over j, after simplification, we have

$$Q(z) = \left(\bar{\lambda} + \lambda G(\mu + \bar{\mu}z)\right) Q\left(\mu + \bar{\mu}z\right), \tag{21}$$

which is a functional equation that defines the probability generating function $Q(z)$. Using power series expansion around $z = 1$ and following the same procedure as in LAS-DA policy, we obtain from Eq. (21),

$$Q(z) = \sum_{\ell=0}^{\infty} \frac{d_\ell}{\ell!}\ (z-1)^\ell, \tag{22}$$

where d_ℓ is given by the following recursive formula

$$d_0 = 1, \tag{23}$$

$$d_\ell = \frac{\lambda}{1-\bar{\mu}^\ell} \sum_{k=0}^{\ell-1} \binom{\ell}{k} \bar{\mu}^k \; d_k \; B^{(\ell-k)}(1), \quad \ell \geq 1, \tag{24}$$

where $B(z) = H(\mu + \bar{\mu}z)$. We obtain the alternative solution from Eq. (21) following the same procedure as in LAS-DA policy as

$$Q(z) = H(\mu + \bar{\mu}z)Q\left(\mu + \bar{\mu}z\right) = \prod_{j=1}^{\infty} H\left(1 - \bar{\mu}^j + \bar{\mu}^j z\right). \tag{25}$$

Taking derivative with respect to z and setting $z = 1$, we obtain $L_s = Q^{(1)}(z)|_{z=1} = \frac{\lambda\bar{\mu}\bar{g}}{\mu}$.

3.2.1. Outside observer's observation epoch

At the outside observer's observation epoch in EAS, let Q_n^o be the outside observer's distribution, that is, there are n requests in the system. In this instance, the observation epoch of the external observer is likely to occur between a possible arrival and a possible departure. Since we already know the system's number distribution at m, and since there is only one event—an arrival or no arrival—after m, the Q_n^o is determined by

$$Q_0^o = \bar{\lambda} Q_0, \tag{26}$$

$$Q_k^o = \bar{\lambda} Q_k + \lambda \sum_{j=1}^{k} g_j Q_{k-j}, \; k \geq 1. \tag{27}$$

The PGF at outside observer's observation epoch in EAS using Eq. (25) simplifies as

$$Q^o(z) = \left(\bar{\lambda} + \lambda G(z)\right) Q(z) = H(z) \prod_{j=1}^{\infty} H\left(1 - \bar{\mu}^j + \bar{\mu}^j z\right). \tag{28}$$

Note that $Q^o(z) = P^o(z) = P(z)$, which implies $Q_n^o = P_n$, meaning that the system size distribution at the observation epoch of the outside observer is equivalent in both LAS-DA and EAS. Taking derivative with respect to z and putting $z = 1$, we have

$$L_s^o = Q^{o(1)}(z)|_{z=1} = \frac{\lambda\bar{g}}{\mu}. \tag{29}$$

Remark 2. *Similar findings can be obtained from the observation epoch probabilities of outside observers in the EAS policy, even in the limiting case. This leads to the conclusion that both LAS-DA and EAS queue results tend to be the same in continuous time.*

4. Special cases

In this section, we assume various probability distributions for the arrival batch sizes and derive the system size distribution. The distribution of batch sizes significantly impacts queue variability and latency when arrivals occur in batches. Geometric, Shifted Poisson, Shifted Bernoulli, and Shifted Binomial probability distributions are discussed for arrival batch sizes.
Geometric batches. Let us assume the batch size is Geometrically distributed with pmf

$P(X=\ell)=\varphi(1-\varphi)^{\ell-1}$, where $0<\varphi<1$. Here, pgf of geometric distribution is $G(z)=\frac{\varphi z}{1-\bar{\varphi}z}$, and $H(z)=\bar{\lambda}+\frac{\lambda\varphi z}{1-\bar{\varphi}z}$. Applying this in Eq. (19), we have

$$P(z)=\left(\bar{\lambda}+\frac{\lambda\varphi z}{1-\bar{\varphi}z}\right)\prod_{j=1}^{\infty}\left(\bar{\lambda}+\frac{\lambda\varphi\left(1-\bar{\mu}^j+\bar{\mu}^j z\right)}{1-\bar{\varphi}\left(1-\bar{\mu}^j+\bar{\mu}^j z\right)}\right). \tag{30}$$

Shifted Poisson batches. We consider that the batch arrival is shifted Poisson distributed with pmf $P(X=i)=\frac{\varphi^{i-1}e^{-\varphi}}{(i-1)!}, i\geq 1$, where $\varphi>0$. Here, $G(z)=ze^{-\varphi(1-z)}$, and $H(z)=\bar{\lambda}+\lambda ze^{-\varphi(1-z)}$. Applying this in (19), we have

$$P(z)=\left(\bar{\lambda}+\lambda ze^{-\varphi(1-z)}\right)\prod_{j=1}^{\infty}\left[\bar{\lambda}+\lambda\left(1-\bar{\mu}^j+\bar{\mu}^j z\right)e^{-\varphi\bar{\mu}^j(1-z)}\right]. \tag{31}$$

Shifted Bernoulli batches. Let us assume the batch size is a shifted Bernoulli random variable that takes only two values: $P(X=1)=\bar{\varphi}$ and $P(X=2)=\varphi$. So, $G(z)=z(\bar{\varphi}+\varphi z)$, which gives $H(z)=\bar{\lambda}+\lambda z(\bar{\varphi}+\varphi z)$. In this case, the PGF is

$$P(z)=\left(\bar{\lambda}+\lambda z(\bar{\varphi}+\varphi z)\right)\prod_{j=1}^{\infty}\left[1-\lambda\bar{\mu}^j(1-z)\left(1+\varphi-\varphi\bar{\mu}^j(1-z)\right)\right]. \tag{32}$$

Shifted Binomial batches. Let us assume the batch size is shifted binomial random variable and its pmf is $P(X=i)=\binom{n}{i-1}\varphi^{i-1}\bar{\varphi}^{n-(i-1)}, i=1,2,\ldots,n+1$ and $G(z)=z(\bar{\varphi}+\varphi z)^n$, which gives $H(z)=\bar{\lambda}+\lambda z(\bar{\varphi}+\varphi z)^n$. In this case, the pgf is

$$P(z)=\left(\bar{\lambda}+\lambda z(\bar{\varphi}+\varphi z)^n\right)\prod_{j=1}^{\infty}\left[\bar{\lambda}+\lambda\left(1-\bar{\mu}^j(1-z)\right)\left(1-\varphi\bar{\mu}^j(1-z)\right)^n\right]. \tag{33}$$

Table 1 shows the impact of several probability distributions for the arrival batch sizes on queueing performance.

Distribution	Queueing performance
Geometric	High variability; bursty traffic patterns
Shifted Poisson	Analytical tractability; moderate randomness;
Shifted Bernoulli	Sparse arrivals; low variability
Shifted Binomial	Smoother traffic; moderate variability

Table 1: *Impact of batch-size distribution on the queueing performance.*

The geometric distribution typically exhibits the highest effective burstiness among these distributions, because it allows unbounded batch sizes, which can increase the expected number of jobs in the system and lead to larger fluctuations in instantaneous workload. In contrast, shifted Bernoulli and shifted binomial arrivals produce more regulated traffic patterns. They are frequently employed as baseline models for systems that are not heavily loaded or subject to regulations. The shifted Poisson distribution is used due to its memorylessness and mathematical tractability. As a result, the batch-size distribution selection should account for anticipated fluctuations in actual workloads, especially in burst-sensitive settings such as serverless computing systems.

5. *Geo/Geo/*∞ queue with LAS-DA policy

Taking $g_1=1$ and $g_i=0, \forall i\geq 2$, the above model reduce to $Geo/Geo/\infty$ queue with LAS-DA policy. Here, the functional equation is

$$P(z)=\left(\bar{\lambda}+\lambda z\right)P\left(\mu+\bar{\mu}z\right). \tag{34}$$

Taking derivative of Eq. (34) with respect to z and and putting $z = 1$, we obtain

$$P(z)^{(1)}|_{z=1} = \sum_{n=0}^{\infty} nP_n = \frac{\lambda}{\mu} = L_s. \tag{35}$$

Since we have as many servers as requests in the system, the average number of requests in the queue and the average waiting time in the queue is zero. The sojourn time is the mean service time, that is, $W_s = \frac{1}{\mu}$. Applying power series expansion about $z = 1$, from Eq. (34), we get Eq. (7), where

$$d_k = \frac{\ell\lambda\bar{\mu}^{k-1}}{1-\bar{\mu}^k} d_{k-1}, \quad d_0 = 1. \tag{36}$$

Using recursively the above, we get

$$d_k = k!\lambda^k \prod_{j=0}^{k-1} \frac{\bar{\mu}^j}{1-\bar{\mu}^{j+1}}, \quad k \geq 1, \quad \text{thus,} \quad \sum_{\ell=0}^{\infty} P_\ell z^\ell = \sum_{\ell=0}^{\infty} \frac{z^\ell}{\ell!} \sum_{k=\ell}^{\infty} \frac{(-1)^{k-\ell} d_k}{(k-\ell)!}. \tag{37}$$

Remark 3. *Taking inter-arrival and service times as exponentially distributed, with mean arrival and mean service rates γ and η, respectively. Divide the time axis into equal intervals $\Delta > 0$ such that $\lambda = \gamma\Delta$, $\mu = \eta\Delta$, and $\rho = \frac{\gamma\Delta}{\eta\Delta} = \rho_1$, where Δ is appropriately small. Setting $\lambda = \gamma\Delta$, $\mu = \eta\Delta$, and taking the limit as $\Delta \to 0$, we have*

$$\lim_{\Delta\to 0} d_k = k! \lim_{\Delta\to 0} (\gamma\Delta)^k \prod_{j=0}^{k-1} \frac{(1-\eta\Delta)^j}{1-(1-\eta\Delta)^{j+1}} = \left(\frac{\gamma}{\eta}\right)^k = \rho_1^k. \tag{38}$$

So,

$$\sum_{\ell=0}^{\infty} P_\ell z^\ell = \sum_{\ell=0}^{\infty} \frac{\rho_1^\ell z^\ell}{\ell!} \sum_{k=\ell}^{\infty} \frac{(-1)^{k-\ell}\rho_1^{k-\ell}}{(k-\ell)!} = e^{-\rho_1} \sum_{\ell=0}^{\infty} \frac{(\rho_1 z)^\ell}{\ell!}, \tag{39}$$

which matches with the results of Shortle et al. [24].

5.1. *Geo/Geo/∞* queue with EAS policy

Using the same procedure as discussed in previous subsection, we get the following expressions

$$Q(z) = \left(\bar{\lambda} + \lambda(\mu + \bar{\mu}z)\right) Q\left(\mu + \bar{\mu}z\right), \tag{40}$$

which is a functional equation. We obtain the probabilities after simplification

$$Q_j = \sum_{k=j}^{\infty} \binom{k}{j} (-1)^{k-j} \lambda^k \prod_{i=1}^{k} \frac{\bar{\mu}^i}{1-\bar{\mu}^i}, \quad j \geq 0. \tag{41}$$

The PGF of outside observer's observation epoch is $Q^o(z) = \left(\bar{\lambda} + \lambda z\right) Q(z)$, and

$$Q_j^o = \bar{\lambda} \sum_{k=j}^{\infty} \binom{k}{j} (-1)^{k-j} \lambda^k \prod_{i=1}^{k} \frac{\bar{\mu}^i}{1-\bar{\mu}^i} + \sum_{k=j-1}^{\infty} \binom{k}{j-1} (-1)^{k-j+1} \lambda^{k+1} \prod_{i=1}^{k} \frac{\bar{\mu}^i}{1-\bar{\mu}^i}. \tag{42}$$

The mean number of tasks in the system at outside observer's epoch is $L_s^o = \frac{\lambda}{\mu}$, which is same as of LAS-DA policy.

6. Numerical illustrations

In this section, we examine how model parameters affect several probability distributions and various performance metrics. The numerical experiments demonstrate some of the qualitative aspects of the system under study and validate the analytical results obtained in the sections above. The numerical calculations were performed on a PC with an Intel(R) Core(TM) i7-1165G7 CPU @ 2.80 GHz and 16GB RAM, using Maple 22 software.

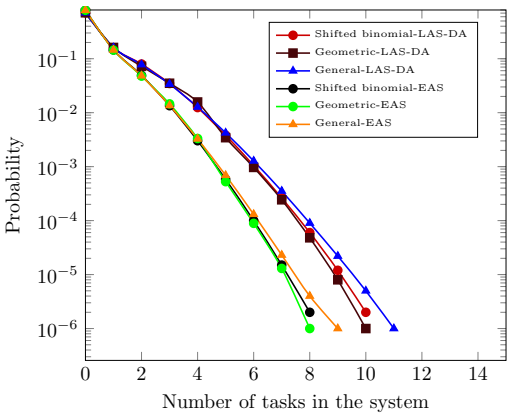


Figure 5: *Number of tasks vs probability.*

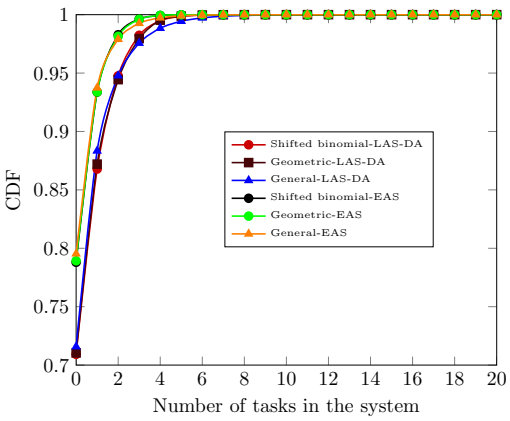


Figure 6: *Number of tasks vs CDF.*

Fig. 5 illustrates the probabilities of several requests for various probability distributions of arrival batch sizes for EAS and LAS-DA policies. As the number of tasks increases, the probability decreases. For a shifted binomial batch, the probability is lower than for geometric- and general-distributed batch sizes, as the batch size is small or bounded, resulting in more compactly focused arrival forms. In contrast, distributions with higher variability, such as the arbitrary and geometric batch size distributions, exhibit noticeably larger fluctuations. Namely, large batches can push many jobs into the system simultaneously, increasing the likelihood of transient workload spikes. Fig. 6 depicts the cumulative distribution function (CDF) of the number of tasks for various distributions of fixed mean batch size. The CDF increases with the number of tasks in the system, and eventually attains its maximum value. As expected, the number of tasks is maximum when the probability distributions of arrival batch sizes are arbitrary and minimum when the distribution is shifted binomial. Fig. 7 shows the probability distribution of the number of tasks in the system for both EAS and LAS-DA policies. When the system is empty, the probability in EAS is greater than in LAS-DA, because arrivals are admitted immediately before departures. When the system is not idle, EAS provides immediate service and reduces delays, but it can create sharper peaks in resource demand. Contrary, LAS-DA imposes a one-slot delay, which can be advantageous for systems that require batching or synchronization, as it reduces the risk of overload at slot boundaries.

For serverless systems, this difference translates into cost and performance trade-offs. Therefore, whether the objective is to minimize delay or to stabilize resource consumption under bursty arrival patterns, managers must align the system selection with operational requirements. Fig. 8 presents the number of tasks in the system when the arrival of tasks is single for various ρ. Here, we consider LAS-DA policy. For fixed ρ, as the number of tasks in the system increases, the probability decreases. For a fixed number of tasks in the system, the probability decreases as ρ increases. In addition to the qualitative trends observed in the figures, several quantitative insights emerge regarding the impact of the batch-size distribution on system performance. Specifically, the findings emphasize the importance of arrival variability. For distributions with the same mean batch size, the variance significantly affects the expected

number of jobs in the system.

Table 2 depicts steady-state probabilities of the $Geo^X/Geo/\infty$ queue for various probability distributions for the arrival batch sizes. The parameters are $\lambda = 0.1, \mu = 0.4$ and mean batch size $\bar{g} = 2$. In all cases, the probability decreases as the number of tasks increases. Probabilities in the case of an idle server are higher when the arrival batch size follows a Poisson distribution. The structure of arrival bursts influences system performance in addition to the mean arrival rate. Large numbers of function instances may be present in serverless computing systems, even when the average workload remains constant; distributions with greater variance can significantly raise the peak number of concurrent executions when launched. This observation underscores the importance of accounting for arrival variability when designing resource provisioning techniques, scaling policies, and concurrency constraints.

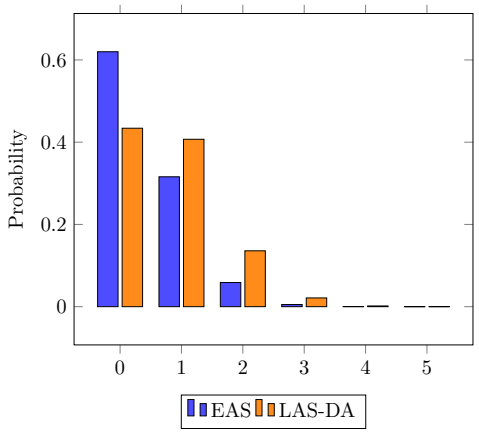


Figure 7: *Distribution of the number of tasks in the system.*

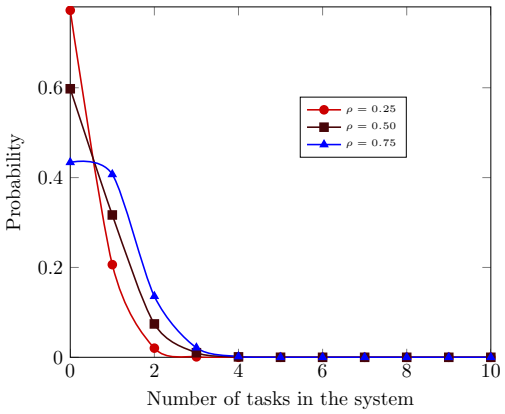


Figure 8: *Distribution of the number of tasks in the system with single server.*

	$\lambda = 0.1, \mu = 0.4, \bar{g} = 2$					
	Geometric		Shifted Poisson		Shifted Binomial	
n	P_n	Q_n	P_n	Q_n	P_n	Q_n
0	0.715786	0.795318	0.709933	0.788815	0.709210	0.788012
1	0.167613	0.142052	0.159749	0.145255	0.158627	0.145723
2	0.064377	0.041546	0.077744	0.048202	0.079778	0.049075
3	0.027881	0.013679	0.033875	0.013610	0.034745	0.013471
4	0.012738	0.004743	0.012725	0.003269	0.012435	0.003020
5	0.005994	0.001691	0.004229	0.000691	0.003817	0.000582
10	0.000166	0.000011	0.000005	0.000000	0.000002	0.000000
15	0.000005	0.000000	0.000000	0.000000	0.000000	0.000000
20	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Sum	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

Table 2: *Steady-state probabilities of the $Geo^X/Geo/\infty$ queue.*

An effective model for comprehending serverless dynamics is the discrete-time batch-arrival infinite-server queue. Managers should use it as a forecasting tool to ensure governance, budget for volatility, and anticipate bursts. This model yields several practical qualitative insights for managers operating serverless policies. In serverless computing, the infinite server queue enables instant allocation of compute resources without waiting. Managers must anticipate cost surges when large batches trigger simultaneous function executions. Due to discrete time modelling, time-slotted arrivals mirror real-world scheduling and help managers forecast peak-demand windows and optimize resource allocation. Pay-per-use billing can cause budget shocks

when large batches arrive in short succession, leading to volatile costs. Infinite servers eliminate queue delays but complicate system observability, and managers must track thousands of concurrent executions. Large batch arrivals may involve sensitive data; managers must ensure regulatory compliance during high-volume processing. The infinite-server batch-arrival queue model captures serverless elasticity; managers should exploit it for bursty workloads. Use predictive analytics to forecast batch arrival patterns and set budget alerts. Establish monitoring dashboards for batch arrivals and execution concurrency. Diversify across providers to reduce lock-in and compliance exposure. Infinite servers can scale beyond budget; managers must set usage thresholds. Batch arrivals may mask inefficiencies; invest in real-time analytics. Engage compliance teams early when handling sensitive batch workloads.

7. Conclusions

This paper examines the effectiveness of serverless computing in a discrete-time infinite-server batch/single-arrival queue, where customers arrive in batches rather than individually. Each customer is immediately served by one of an infinite number of servers, thereby eliminating waiting times due to unlimited processing capacity. We obtain state probabilities at various epochs, along with their corresponding performance measures. We illustrate numerical illustrations in the form of graphs and tables. The batch-arrival infinite-server queue paradigm in a serverless architecture is ideal for managing concurrent workloads on cloud platform. The discrete-time model offers valuable insights for serverless computing architectures, as it can capture key aspects of event-driven, auto-scaling systems. Thus, using discrete-time queueing models to analyze serverless system performance helps predict scalability, optimize resource allocation, and forecast scalability. In serverless computing platforms, where each incoming request is executed in a separate execution environment and, in theory, may be served instantly without waiting in line, this paradigm is particularly pertinent. Naturally, traffic spikes caused by external factors, such as IoT signals, user surges, and planned jobs, are modelled as batch arrivals. However, the infinite-server assumption imposes significant restrictions. Infrastructure limitations may result in implicit resource coupling, and providers typically set concurrency limits at the account or region level. It can be further extended for greater realism and practical relevance, such as assuming an infinite number of servers, a massive but finite number, leading to discrete-time $GI^X/G/c$ models, or state-dependent service distributions to model initialisation delays. Generalising discrete-time batch-arrival models along these lines would serve as a bridge between analytically tractable infinite-server queues and the operating conditions of contemporary cloud-native and serverless systems.

References

[1] Altman, E. and Yechiali, U. (2008). Infinite-server queues with system's additional tasks and impatient customers. Probability in the Engineering and Informational Sciences, 22(4), 477–493. doi: 10.1017/S0269964808000296

[2] Artalejo, J. R. and Hernández-Lerma, O. (2003). Performance analysis and optimal control of the Geo/Geo/c queue. Performance Evaluation, 52(1), 15–39. doi: 10.1016/S0166-5316(02)00161-X

[3] Chan, W. C. and Maa, D. Y. (1978). The GI/Geom/N queue in discrete time. INFOR: Information Systems and Operational Research, 16(3), 232–252. doi: 10.1080/03155986.1978.11731705

[4] Chaudhry, M. L. and Kim, J. D. (2004). Complete analytic and computational analyses of the discrete-time bulk-arrival infinite-server system: $GI^X/Geom/\infty$. Computers & Operations Research, 31(13), 2119–2135. doi: 10.1016/S0305-0548(03)00167-9

[5] Chaudhry, M. L., Gupta, U. C. and Goswami, V. (2001). Modeling and analysis of discrete-time multiserver queues with batch arrivals: $GI^X/Geom/m$. INFORMS Journal on Computing, 13(3), 172–180. doi: 10.1287/ijoc.13.3.172.12627

[6] Daw, A. and JPender, J. (2019). On the distributions of infinite server queues with batch arrivals. Queueing Systems, 91(3), 367–401. doi: 10.1007/s11134-019-09603-4

[7] Finol, G., París, G., García-López, P. and Sánchez-Artigas, M. (2024). Exploiting inherent elasticity of serverless in algorithms with unbalanced and irregular workloads. Journal of Parallel and Distributed Computing, 190, 104891. doi: 10.1016/j.jpdc.2024.104891

[8] Brian H Fralix, B. H. and Adan, I. J. (2009). An infinite-server queue influenced by a semi-markovian environment. Queueing Systems, 61(1), 65–84. doi: 10.1007/s11134-008-9100-y

[9] Goswami, V. and Gupta, U. C. (1998). Analyzing the discrete-time multiserver queue Geom/Geom/m queue with late and early arrivals. Information and Management Sciences, 9(2), 55–66. url: eudml.org/doc/249750 [Accessed 4/1/2025]

[10] Goswami, V. and Mund, G. B. (2017). Computational analysis of multi-server discrete-time queueing system with balking, reneging and synchronous vacations. RAIRO-Operations Research, 51(2), 343–358. doi: 10.1051/ro/2016025

[11] Goswami, V., Gupta, U. C. and Samanta, S. K. (2006). Analyzing discrete-time bulk-service $Geo/Geo^b/m$ queue. RAIRO-Operations Research, 40(3), 267–284. doi: 10.1051/ro:2006021

[12] Gravey, A. and Hébuterne, G. (1992). Simultaneity in discrete-time single server queues with Bernoulli inputs. Performance Evaluation, 14(2), 123–131. doi: 10.1016/0166-5316(92)90014-8

[13] Hunter, J. J. (1983). *Mathematical Techniques of Applied Probability, Volume II, Discrete Time Models: Techniques and Applications*. New York: Academic Press.

[14] Kaffes, K., Yadwadkar, N. J. and Kozyrakis, C. (2021). Practical scheduling for real-world serverless computing. arXiv preprint arXiv:2111.07226. doi: 10.48550/arXiv.2111.07226

[15] Kelly, D., Glavin, F. and Barrett, E. (2020). Serverless computing: Behind the scenes of major platforms. In: *2020 IEEE 13th International Conference on Cloud Computing(CLOUD)*, 304–312. doi: 10.1109/CLOUD49709.2020.00050

[16] Kumar, R. (2017). A transient solution to the M/M/c queuing model equation with balking and catastrophes. Croatian Operational Research Review, 8(2), 577–591. doi: 10.17535/crorr.2017.0037

[17] Nassar, H. (2003). A novel approach to analyzing the occupancy of the $Geo^X/Geo^Y/\infty$ queueing system. AEU-International Journal of Electronics and Communications, 57(6), 391–394. doi: 10.1078/1434-8411-54100190

[18] Palomo, S. and Pender, J. (2024). Overlap times in the infinite server queue. Probability in the Engineering and Informational Sciences, 38(1), 21–27. doi: 10.1017/S0269964822000456

[19] Pang, G. and Whitt, W. (2012). Infinite-server queues with batch arrivals and dependent service times. Probability in the Engineering and Informational Sciences, 26(2), 197–220. doi: 10.1017/S0269964811000337

[20] Ramaswami, V. and Neuts, M. F. (1980). Some explicit formulas and computational methods for infinite-server queues with phase-type arrivals. Journal of Applied Probability, 17(2), 498–514. doi: 10.2307/3213039

[21] Reynolds, J. F. (1968). Some results for the bulk-arrival infinite-server Poisson queue. Operations Research, 16(1), 186–189. doi: 10.1287/opre.16.1.186

[22] Rubin, I. and Zhang, Z. (2002). Message delay and queue-size analysis for circuit-switched TDMA systems. IEEE Transactions on Communications, 39(6), 905–914. doi: 10.1109/26.87180

[23] Shanbhag, D. N. (1966). On infinite server queues with batch arrivals. Journal of Applied Probability, 3(1), 274–279. doi: 10.2307/3212053

[24] Shortle, J. F., Thompson, J. M., Gross, D. and Harris, C. M. (2018). Fundamentals of queueing theory. New York: John Wiley & Sons, Inc. doi: 10.1002/9781119453765

[25] Sinha, P., Kaffes, K. and Yadwadkar, N. J. (2024). Shabari: Delayed decision-making for faster and efficient serverless functions. arXiv preprint arXiv:2401.08859. doi: 10.48550/arXiv.2401.08859

[26] Wen, J., Chen, Z., Sarro, F. and Liu, X. (2023). Revisiting the performance of serverless computing: An analysis of variance. arXiv preprint arXiv:2305.04309. doi: 10.48550/arXiv.2305.04309

[27] Wen, J., Chen, Z., Sarro, F. and Wang, S. (2025). Unveiling overlooked performance variance in serverless computing. Empirical Software Engineering, 30(2), 30–59. doi: 10.1007/s10664-025-10615-3

[28] Zhu, L., Giotis, G., Tountopoulos, V. and Casale, G. (2021). RDOF: deployment optimization for function as a service. In: *2021 IEEE 14th International Conference on Cloud Computing (CLOUD)*, 508–514. doi: 10.1109/CLOUD53861.2021.00066

Croatian Operational Research Review

Information for authors

Journal homepage: http://www.hdoi.hr/crorr-journal

Aims and Scope: The aim of the Croatian Operational Research Review journal is to provide high quality scientific papers covering the theory and application of operational research and related areas, mainly quantitative methods and machine learning. The scope of the journal is focused, but not limited to the following areas: combinatorial and discrete optimization, integer programming, linear and nonlinear programming, multiobjective and multicriteria programming, statistics and econometrics, macroeconomics, economic theory, games control theory, stochastic models and optimization, banking, finance, insurance, simulations, information and decision support systems, data envelopment analysis, neural networks and fuzzy systems, and practical OR and application.

Papers blindly reviewed and accepted by two independent reviewers are published in this journal.

Author Guidelines: There is no submission fee for publishing papers in this journal. All manuscripts should be submitted via our online manuscript submission system at the journal homepage. A user ID and password can be obtained on the first visit. The initial version of the manuscript should be written using LaTex program or Word program, and submitted as a single PDF file. After the reviewing process, if the paper is accepted for publication in this journal, the corresponding author will be asked to submit proofreading confirmation along with signed copyright form. CRORR uses the Turnitin Similarity Check plagiarism software to detect instances of overlapping and similar text in submitted manuscripts. Any copying of text, figures, data or results of other authors without giving credit is defined as plagiarism and is a breach of professional ethics. Such papers will be rejected and other penalties may be accessed.

Electronic Presentations: The journal is also presented electronically on the journal homepage. All issues are also available at the Hrcak database – Portal of science journals od the Republic of Croatia, with full-text search of individual, at

http://hrcak.srce.hr/crorr

Offprints: The corresponding author of the published paper will receive a single print copy of the journal. All published papers are freely available online.

Publication Ethics and Malpractice Statement

Croatian Operational Research Review journal is committed to ensuring ethics in publication and quality of articles. Conformance to standards of ethical behaviour is therefore expected of all parties involved: Authors, Editors, Reviewers, and the Publisher. Our ethic statements are based on COPE's Best Practice Guidelines for Journals Editors.

Publication decisions. The editor of the journal is responsible for deciding which of the articles submitted to the journal should be published. The editor may be guided by the policies of the journal's editorial board and constrained by such legal requirements as shall then be in force regarding libel, copyright infringement and plagiarism. The editor may confer with other editors or reviewers in making this decision.

Fair play. An editor will at any time evaluate manuscripts for their intellectual content without regard to race, gender, sexual orientation, religious belief, ethnic origin, citizenship, or political philosophy of the authors.

Confidentiality. The editor and any editorial staff must not disclose any information about a submitted manuscript to anyone other than the corresponding author, reviewers, potential reviewers, other editorial advisers, and the publisher, as appropriate.

Disclosure and conflicts of interest. Unpublished materials disclosed in submitted manuscript must be used in an editor's own research without the express written consent of the author.

Duties of Reviewers

Contribution to Editorial Decisions. Peer review assists the editor in making editorial decisions and through the editorial communications with the author may also assist the author in improving the paper.

Promptness. Any selected referee who feels unqualified to review the research reported in a manuscript or knows that its prompt review will be impossible should notify the editor and excuse himself from the review process.

Confidentiality. Any manuscripts received for review must be treated as confidential documents. They must not be shown to or discussed with others except at authorized by the editor.

Standards of Objectivity. Reviews should be conducted objectively. Personal criticism of the author is inappropriate. Referees should express their views clearly with supporting arguments.

Acknowledgement of Sources. Reviewers should identify relevant published work that has not been cited by the authors. Any statement that an observation, derivation, or argument had been previously reported should be accompanied by the relevant citation. A reviewer should also call to the editor's attention any substantial similarity or overlap between the manuscript under consideration and any other published paper of which they have personal knowledge.

Disclosure and Conflict of Interest. Privileged information or ideas obtained through peer review must be kept confidential and not used for personal advantage. Reviewers should not consider manuscripts in which they have conflicts of interest resulting from competitive, collaborative, or other relationships or connections with any of the authors, companies, or institutions connected to the papers.

Duties of Authors

Reporting Standards. Authors' manuscripts of original research should present an accurate of the work performed as well as an objective discussion of its significance. Underlying data should be represented accurately in the paper. A paper should contain sufficient detail and references to permit others to replicate the work. Fraudulent or knowingly inaccurate statements constitute unethical behaviour are unacceptable.

Data Access and Retention. When submitting their article for publication in this journal, the authors confirm the statement that all data in article are real and authentic. Authors should be prepared to provide the raw data in connection with a paper for editorial review, and should be prepared to provide public access to such data (consistent with the ALPSP-STM Statement on Data and Databases), if practicable, and should in any event be prepared to retain such data for a reasonable time after publication. In addition, all authors are obliged to provide retractions or corrections of mistakes if any.

Originality and Plagiarism. The authors should ensure that they have written entirely original works, and if the authors have used the work and/or words of others that has been appropriately cited or quoted. Multiple, Redundant or Concurrent Publication An author should not in general publish manuscripts describing essentially the same research in more than one journal or primary publication. Submitting the same manuscript to more than one journal concurrently constitutes unethical publishing behaviour and is unacceptable.

Acknowledgement of Sources. Proper acknowledgement of the work of others must always be given. Authors should cite publications that have been influential in determining the nature of the reported work.

Authorship of the Paper. Authorship should be limited to those who have made a significant contribution to the concept, design, execution, or interpretation of the reported study. All those who have made significant contributions should be listed as co-authors. Where there are others who have participated in certain substantive aspects of the research project, they should be acknowledged or listed as contributors. The corresponding author should ensure that all appropriate co-authors and no inappropriate co-authors are included on the paper, and that all co-authors have seen and approved the final version of the paper and have agreed to its submission for publication.

Disclosure and Conflicts of Interest. All authors should disclose in their manuscript any financial or other substantive conflict of interest that might be construed to influence the results or interpretation of their manuscript. All sources of financial support for the project should disclosed. When an author discovers a significant error or inaccuracy in his/her own published work, it is the author's obligation to promptly notify the journal editor or publisher and cooperate with editor to retract or correct the paper.

Operational Research

Operational Research (OR) is an interdisciplinary scientific branch of applied mathematics that uses methods such as mathematical modelling, statistics, and algorithms to achieve optimal or near optimal solutions to complex problems. It is also known as Operations Research, and is related to Management Science, Industrial Engineering, Decision Making, Statistics, Simulations, Business Analytics, and other areas.

Operational Research is typically concerned with optimizing some objective function, or finding non-optimal but satisfactory solutions. It generally helps management to achieve its goals using scientific methods. Some of the primary tools used by operations researchers are optimization, statistics, probability theory, queuing theory, game theory, graph theory, decision analysis, and simulation.

Areas/Fields in which Operational Research may be applied can be summarized as follows:

- Assignment problem,
- Business analytics,
- Decision analytics,
- Econometrics,
- Linear and nonlinear programming,
- Inventory theory,
- Optimal maintenance,
- Scheduling,
- Stochastic process, and
- System analysis.

Operational Research has improved processes in business and all around us. It is used in everyday situations, such as design of waiting lines, supply chain optimization, scheduling of buses or airlines, telecommunications, or global resources planning decisions.

Operational Research is an important area of research because it enables the best use of available resources. New methods and models in operational research are to be developed continuously in order to provide adequate solutions to process enormous information growth resulted from rapid technology development in today's new economy. By improving processes it enables high-quality products and services, better customer satisfaction, and better decision making. There is no doubt that operational research contributes to the quality of life and economic prosperity on microeconomic, macroeconomic, and global level.

Croatian Operational Research Society

Croatian Operational Research Society (CRORS) was established in 1992. as the only scientific association in Croatia specialized in operational research. The Society has more than 150 members and its main mission is to promote operational research in Croatia and worldwide for the benefit of science and society. This mission is realized through several goals:

- to encourage collaboration of scientists and researchers in the area of operational research in Croatia and worldwide through seminars, conferences, lectures and similar ways of collaborations,
- to organize the International Conference on Operational Research (KOI),
- to publish a scientific journal Croatian Operational Research Review (CRORR).

Since 1994, CRORS is a member of The International Federation of Operational Research Societies (IFORS), an umbrella organization comprising the national Operations Research societies of over forty-five countries from four geographical regions: Asia, Europe, North America and South America. CRORS is also a member of the Association of European Operational Research Societies (EURO) and actively participates in international promotion of operational research.

If you are a researcher, academic teacher, student or practitioner interested in developing and applying operations research methods, you are welcome to become a member of the CRORS and join our OR community.

Former Presidents:	Prof. Luka Neralić, PhD, University of Zagreb (1992-1996) Prof. Tihomir Hunjak, PhD, University of Zagreb (1996-2000) Prof. Kristina Šorić, PhD, University of Zagreb (2000-2004) Prof. Valter Boljunčić, PhD, University of Pula (2004-2008) Prof. Zoran Babić, PhD, University of Split (2008-2012) Prof. Marijana Zekić-Sušac, PhD, University of Osijek (2012-2016) Prof. Zrinka Lukač, PhD, University of Zagreb (2016-2020) Prof. Mario Jadrić, PhD, University of Split (2020-2022)
President:	associate professor Tea Šestanović, PhD, University of Split (since 2022)
Vice President:	Prof. Snježana Pivac, PhD, University of Split
Secretary:	assistant professor Tea Kalinić Milićević, PhD, University of Split
Treasurer:	assistant professor Marija Vuković, PhD, University of Split

If you want to become a member of the CRORS, please fill in the application form available at

http://www.hdoi.hr