

Current Approaches in Applied Statistics - II

Editors: Assoc. Prof. Yalçın TAHTALI • Assoc. Prof. İbrahim DEMİR
Assist. Prof. Lütfi BAYYURT • Assoc. Prof. Samet Hasan ABACI



 OZGUR
PRESS

Current Approaches in Applied Statistics - II

Editors:

Assoc. Prof. Yalçın TAHTALI

Assoc. Prof. İbrahim DEMİR

Assist. Prof. Lütfi BAYYURT

Assoc. Prof. Samet Hasan ABACI



Published by

Özgür Yayın-Dağıtım Co. Ltd.

Certificate Number: 45503

📍 15 Temmuz Mah. 148136. Sk. No: 9 Şhitkamil/Gaziantep

☎ +90.850 260 09 97

📞 +90.532 289 82 15

🌐 www.ozguryayinlari.com

✉ info@ozguryayinlari.com

Current Approaches in Applied Statistics - II

Editors: Assoc. Prof. Yalçın TAHTALI • Assoc. Prof. İbrahim DEMİR
Assist. Prof. Lütfi BAYYURT • Assoc. Prof. Samet Hasan ABACI

Language: English

Publication Date: 2025

Cover design by Mehmet Çakır

Cover design and image licensed under CC BY-NC 4.0

Print and digital versions typeset by Çizgi Medya Co. Ltd.

ISBN (PDF): 978-625-5646-95-8

DOI: <https://doi.org/10.58830/ozgur.pub865>



This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc/4.0/>
This license allows for copying any part of the work for personal use, not commercial use, providing author attribution is clearly stated.

Suggested citation:

Tahtalı, Y. (ed), Demir, İ. (ed), Bayyurt, L. (ed), Abacı, S. H. (ed) (2025). *Current Approaches in Applied Statistics-II*. Özgür Publications. DOI: <https://doi.org/10.58830/ozgur.pub865>. License: CC-BY-NC 4.0

The full text of this book has been peer-reviewed to ensure high academic standards. For full review policies, see <https://www.ozguryayinlari.com/>



Preface

This book, titled *Current Approaches in Applied Statistics*, is a compilation of recent academic studies produced by researchers from different disciplines. The book covers not only theoretical contributions in the field of statistics, but also the innovative dimensions of methods used in a wide variety of application areas.

Today's rapidly increasing volume and diversity of data has led statistics to transcend being a science based solely on mathematical foundations and take on a critical role in many fields, from health sciences to engineering, social sciences to environmental research. Reflecting this broad sphere of influence, this book aims to present readers with both theoretical approaches and application examples from different disciplines.

The chapters in this book, prepared with contributions from international researchers, highlight the current importance of statistics, the methodological challenges encountered, and new solutions. Readers will encounter content that is useful both academically and practically in areas such as statistical modeling, data mining, machine learning, biostatistics, and social statistics.

We believe this work will provide researchers, graduate students, and practitioners with a comprehensive overview of current approaches to statistics. We thank all the authors and reviewers who contributed to this book and hope it will make a valuable contribution to the scientific community.

Contents

Preface	iii
---------	-----

Chapter 1

Assessment of FCM, PCM, and UFPC Algorithms Through Internal and Fuzzy Cluster Validity Indices on Multidisciplinary Benchmark Datasets	1
<i>Berna Özbaşaran</i>	
<i>Gözde Ulutağay</i>	

Chapter 2

Evaluation of the Performance of BRICST Countries in the Context of the Economic Freedom Index	19
<i>Bilal Saraç</i>	
<i>Çağlar Karamaşa</i>	

Chapter 3

Multi-Period Production and Inventory Planning in Textile Industry: A Case Study of Textile Company	43
<i>Çağdaş Yıldız</i>	
<i>Adem Tüzemen</i>	

Chapter 4

Optimal Product Mix And Resource Allocation In Furniture Manufacturing Using Linear Programming: A Case Of Furniture Company	55
<i>Çağdaş Yıldız</i>	
<i>Adem Tüzemen</i>	

Chapter 5

Latent Similarity Clustering of Video Games Based on Euclidean Distance and PCA 67

Diana Bratić

Chapter 6

Multivariate Anomaly Mapping in Video Games: A Mahalanobis Distance Approach 85

Diana Bratić

Chapter 7

Standardization in Elevator Systems: Call and Display Panel Design Compatible with CANopen Communication Protocol 101

Eyup Sayin

Murat Topuz

Muhammet Fatih Aslan

Akif Durdu

Chapter 8

Improving Production with Artificial Neural Networks and Integration into ERP Systems: An Approach within the Scope of Industry 4.0 117

Gizem Şara Onay

Mehmet Çakmakçı

Chapter 9

Using Statistical Moments in Hierarchical Machine Learning for Estimation of Birefringence in Mode-Locked Fiber Laser Systems 127

Hasan Arda Solak

Şeyma Koltuklu

Sueda Turgut

Mahmut Bağcı

Chapter 10

Examination of Supervised Machine Learning Algorithms in Employee Turnover Prediction 145

Melisa Dikici

Gökçe Sabriye Hörük

Deniz Efendioğlu

Chapter 11

A Macroeconomic Perspective on Türkiye's Climate Crisis Risk: Decision Tree Analysis 165

Mervenur Ünver

Şabika Gökmen

Chapter 12

Mplus Jamovi Amos and Spss Programs in Partial Mediation Model: A Simulation Study 175

Murat Yıldırım

Chapter 13

Investigation of the Structure and Function of Acid-Sensing Ion Channels 211

Ziya Çakır

Chapter 14

Assessment of Ankara University ERASMUS+ Programme Outgoing Students by Using Statistical Methodologies for the 2023-2024 Period 221

Cafer Yıldırım

Özlem Türkşen

Necdet Ünüvar

İlker Astarıcı

Parameter Estimation in Organic-Based Schottky Diodes Used in Solar Cell
Applications with Artificial Intelligence Optimization Algorithms 233

Murat Açıkgöz

Defne Akay

Özlem Türkşen

Assessment of FCM, PCM, and UFPC Algorithms Through Internal and Fuzzy Cluster Validity Indices on Multidisciplinary Benchmark Datasets¹

Berna Özbaşaran²

Gözde Ulutağay³

Abstract

In the contemporary context, due to the high volume of unlabelled data in fields such as medicine, agriculture, chemistry, and many more; unsupervised machine learning models have been topics of interest and of investment. One of the computationally inexpensive and fast models investigated in this paper will be the fuzzy form of K-Means known as Fuzzy C-Means (FCM). Since FCM like K-Means requires the cluster number beforehand it is also vital that the cluster validity indices be fuzzy. In this paper, the evolutionary steps of FCM will be compared by evaluating the models suggested to overcome the pitfalls of the FCM algorithm. As there are many other algorithms created for this purpose, the algorithms analyzed in this article will be Possibilistic C-Means (PCM) and Unsupervised Fuzzy Possibilistic C-Means (UFPC). The comparison of these models is crucial since the new parameters introduced affect the cluster number chosen as seen in the internal validity indices. For applying the algorithms 4 benchmark datasets will be studied in R that belong to fields from biology, chemistry, and demography. The researcher expects that the UFPC algorithm will surpass the others since, the algorithm uses parameters from both FCM and PCM, however, as real-life datasets are rather complex, it is significant that the analysis be compared to benchmark datasets as proposed in this article. The performance will be evaluated on 12 fuzzy clustering validity indices and 3 internal validity indices that being silhouette, gap, and WSS. Custom R libraries will be used to ease the process of applying the algorithms and validity indices.

- 1 This paper was presented at the VI. International Applied Statistic Congress (UYİK)
- 2 Ege University, Faculty of Science, Department of Statistics, Turkey
- 3 Ege University, Faculty of Science, Department of Statistics, Turkey

INTRODUCTION

As real-life datasets tend to be more complex in terms of their structure and volume, bettering and adapting an inexpensive model to be more robust seems as a viable option than using more complex and expensive models. The FCM (Dunn, J. C. 1973) algorithm (which is the fuzzy version of the K-Means) have been a topic of interest in this aspect, owing its simple but elegant nature. With its pitfalls came some suggestions from numerous academics as to better the algorithm and preserve its effectiveness. In this paper, two of those alternatives will be compared based on benchmark datasets, with reference to crisp and fuzzy indices. For this purpose, the algorithms evaluated will be FCM (Fuzzy C-Means), PCM (Possibilistic C-Means) (Krishnapuram, R., & Keller, J. M. 1993) and lastly UFPC (Unsupervised Fuzzy Possibilistic Clustering) (Yang, M., & Wu, K. 2005). The real-world benchmark datasets will be Glass, Seeds, German Credit and Wine Quality, respectively. Lastly the crisp indices considered will be the elbow method, silhouette score and gap statistics, alongside 12 fuzzy indices which will be named in the methods section. Hence, a comparison of evolutionary steps of an algorithm will be evaluated, with reference to both fuzzy and crisp cluster validity indices.

The investigation is significant in comprehending the evolution of an algorithm, since the latter algorithms have been suggested to better the previous approach via additional parameters. For instance, the pitfalls of the FCM (algorithm which is a fuzzy version of the K-Means) have been altered to overcome its vulnerability towards outliers and noise. In the PCM instance, two new parameters called typicality degree and typicality exponent were produced as it was expected to be a more flexible parameter than the membership degree which will be mentioned in the methods part.

As for the performance metrics, the R libraries *fcvalid* and *ppclust* will be utilized alongside others to extract the relevant cluster validity values. The datasets that have been chosen for the comparison are multidisciplinary samples that are known as benchmark datasets. These real-life datasets were chosen for two primary reasons, the first being their versatility in nature and that there is a ground truth meaning the classes are fixed in nature. This means that the datasets tend to have more noise and outliers than synthetic datasets, but also that the cluster number found could be compared with the ground truth to see if it is a match.

In this light, firstly the descriptive statistics will be shown to provide a sense of the data, and to investigate if there are peculiarities in terms of features. Secondly, the crisp validity indices will be shown to see if the

classical cluster validity indices are able to spot the ground truth for the designated dataset. Thirdly, the fuzzy indices will be analyzed for three of the algorithms which are FCM, PCM and UFPC, and the scatter plots will be colored according to cluster numbers for visualization. Finally, the comparison will be made in order to spot some sort of pattern between dataset characters, algorithms and indices. As an educated guess, it can be imagined that UFPC will perform better in rather complex datasets, since it combines the new parameters introduced by the FCM and PCM algorithms.

METHODS

Statistical Analysis

The main aim of this paper is to compare the three algorithms FCM, PCM and UFPC on the basis of their ability to predict the ground truth with help from crisp or fuzzy validity indices. The algorithms used for this purpose can be introduced as the following.

FCM (Fuzzy C-Means) (Dunn, J. C. 1973)

Fuzzy C-Means mathematically works by minimizing the overall weighted distance between data points and cluster centers, where the weights reflect how strongly each point belongs to each cluster. Its aim is to group similar data together while allowing partial membership, so that each point can belong to multiple clusters to varying degrees. The latter algorithms have been suggested since it has a vulnerability towards outliers and noise, as these values tend to receive higher membership degrees that affect the clustering result.

$$J = \sum_{i=1}^k \sum_{j=1}^n u_{ij}^m |x_j - \mu_i|^2 \quad (1)$$

where

k shows the cluster number

n shows the sample size

u_{ij}^m shows the membership degree

m is known as the fuzzifier parameter or the fuzziness exponent. The algorithm becomes more fuzzy as m increases

$|x_j - \mu_i|^2$ shows the Squared Euclidean distance

PCM (Possibilistic C-Means) (Krishnapuram, R., & Keller, J. M. 1993)

PCM (Possibilistic C-Means) works by minimizing a cost that measures how typical each data point is to clusters without forcing memberships to sum to one like in the membership degree parameter. Its goal is to group data while allowing points to have low typicality if they do not clearly belong to any cluster, making it robust to noise and outliers. A pitfall to this algorithm is that the initialization of the typicality degree is vital and that the algorithm tends to make coincidental clusters (Cebeci, Z. (2020)).

$$J = \sum_{i=1}^k \sum_{j=1}^n t_{ij}^m |x_j - \mu_i|^2 + \sum_{i=1}^k \eta_i \sum_{j=1}^n (1 - t_{ij})^m \quad (2)$$

where

t_{ij} shows the typicality degree. Unlike FCM's membership degree, the typicality degree is not constrained to sum to 1. This parameter reflects how typical or compatible the point is with cluster, independently of other clusters.

η_i shows the regularization parameter for a cluster. This parameter controls how wide or tight the cluster is and influences how quickly the typicality drops with distance from the center.

$\sum_{j=1}^n (1 - t_{ij})^m$ balances the influence of values on the clusters by penalizing nontypical values (like outliers). As the typicality degree is lower the equations result will be higher since it is subtracted and multiplied by the typicality parameter. Since the cost function ought to be minimal, nontypical values will have a higher cost function and be penalized.

UFPC (Unsupervised Fuzzy Possibilistic C-Means) (Yang, M., & Wu, K. 2005)

UFPC combines fuzzy memberships and possibilistic typicalities to cluster data by balancing soft assignments with typicality measures. It aims to improve clustering flexibility and robustness by integrating both approaches to overcome the pitfalls of FCM and PCM. PCA (Possibilistic Clustering Algorithm) is another algorithm that has been built on top of PCM (Krishnapuram, R., & Keller, J. M. 1996).

$$J_{U\text{FPC}}(X;U,V) = \sum_{j=1}^n \sum_{i=1}^c (a \cdot u_{ij,\text{FCM}}^n + b \cdot u_{ij,\text{PCA}}^n) d^2(x_j, v_i) + \frac{\beta}{n^2 \sqrt{c}} \sum_{j=1}^n \sum_{i=1}^c (u_{ij,\text{PCA}}^n \log u_{ij,\text{PCA}}^n - u_{ij,\text{PCA}}^n) \quad (3)$$

where

- A and b show the weighting coefficients balancing the influence for FCM and PCA (possibilistic clustering algorithm), respectively. When b is zero, the cost function turns to the FCM algorithm
- β shows the regularization parameter that balances the entropy term, helping control membership distribution

$\frac{\beta}{n^2 \sqrt{c}} \sum u_{ij,\text{PCA}}^n (\log u_{ij,\text{PCA}}^n - u_{ij,\text{PCA}}^n)$ shows the penalty term, inspired

by entropy. The equation encourages diversity, PCA (prevents all being 0 or 1) so as to promote meaningful typicality distributions.

While the first term combines FCM and PCA, the second term controls the shape of the possibilistic membership distribution and encourages a balanced spread of memberships.

The fuzzy cluster validity indices mentioned in Table 1 will be used with their abbreviations given below. The rationale behind every fuzzy index is intrinsic, hence, the selection of the lowest or highest value should be selected accordingly. For instance, while PC (Partition Coefficient) presents the best value with the maximum value, in PE (Partition Entropy) the minimum value must be chosen. One can analyze Table 1 to learn which value gives the best result from the different clustering values.

Table 1 Fuzzy Internal Validity Indices for FCM Algorithm for the Glass Dataset

Fuzzy Index	Full name	Optimum Cluster Value
PC	Partition Coefficient	Maximum
MPC	Modified Partition Coefficient	Maximum
PE	Partition Entropy	Minimum
XB	Xie-Beni Index	Minimum
K	Kwon Index	Minimum
TSS	Tang, Sun & Sun Index	Minimum
CL	Chen-Linkens Index	Maximum
FS	Fukuyama Sugeno Index	Minimum
PBMF	Pakhira-Bandyopadyang-Maulik	Maximum
FSIL	Fuzzy Silhouette Index	Minimum
FHV	Fuzzy Hyper Volume	Minimum
APD	Average Partition Density	Maximum

For this purpose, one ought to first pick the relevant sample, which in this case will be the real-world benchmark datasets Glass, Seeds, German Credit and Wine Quality.

The Glass Dataset (German, B., 1987).

The dataset belongs to the field of chemistry and consists of 214 instances of glass samples described by 9 numerical features, including refractive index (RI) and elemental compositions like Na, Mg, Al, Si, K, Ca, Ba, and Fe. The ground truth for this dataset is considered as seven glass types, varying from materials used so as to build windows or vehicle headlamps.

The Seeds Dataset (Charytanowicz, M., et al. 2010)

This dataset which is from biology and agriculture contains 210 instances of wheat kernels described by 7 numerical features, including area, perimeter, compactness, kernel length, kernel width, asymmetry coefficient, and groove length. The ground truth for this dataset consist of three wheets which are the following: Kama, Rosa, and Canadian.

The German Credit Dataset (Hofmann, H. 1994).

The German Credit dataset is a dataset from the field of demography and economy that contains 1,000 instances of loan applicants described by 20 attributes, including credit amount, duration, age, employment status, and housing type. The labels that are mentioned in the liteature are binary, that is good loan chance and bad loan chance.

The Wine Quality Dataset (Cortez, P., et al. 2009)

The last dataset from the field of chemistry contains 1,143 red wine samples described by 11 physicochemical attributes, such as acidity, chlorides, and alcohol content; and the label. Wine quality is rated on a scale from 0 to 10 and is grouped into three classes as a benchmark dataset: low (≤ 4), medium (5–6), and high (≥ 7) quality.

For the descriptive analysis, the following numeric features for every variables was selected. In Figure 1, one can see the heatmap for the correlation matrix for all datasets. One can see that the diagonal is red for all the intercept with the variable itself, so this would mean the correlation is 1. The blues represent negative correlation. In the Glass dataset, we can see a strong negative correlation between the variable pairs: Ca, Mg; Al, MG; Ba, Mg; RI, Si; RI, Al, alongside strong positive correlation with RI and Ca. Thus, negative correlation has dominated in the Glass dataset case. For Seeds, one can see that except for the relationship with the assymetry coefficient, the table consist of predominantly strong positive correlation

which might suggest multicollinearity. In the German Credit dataset, one can see a stronger negative correlation between age and the other variables. Lastly, there does not seem to be quite as much correlation between the Wine Quality dataset, but one can spot a strong positive and negative correlation chunk between variables. This information will be useful as we are comparing the dataset in the summary section.

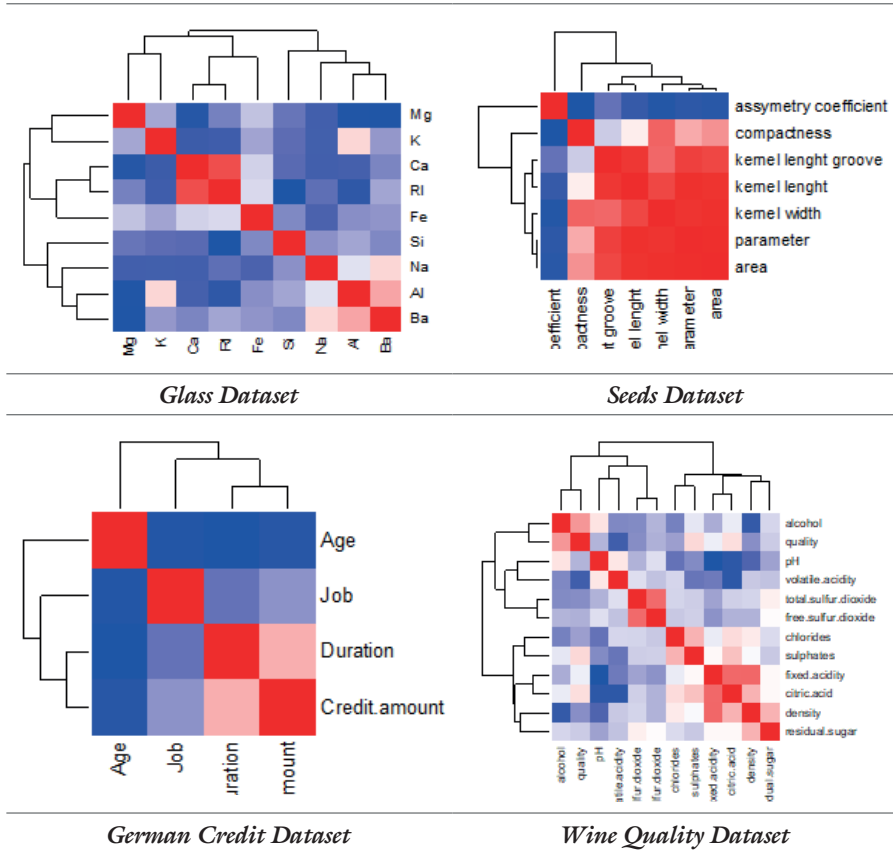


Figure 1 the correlation is represented as positive with red subtones, and negative with blue subtones. The stronger the color, the more strenght in correlation

RESULTS

The analysis for this paper has been done in R, with libraries such as pastecs, e1071, gridextra, factorextra, ppclust and lastly fcvallid. In Figure 2, one can view the ground truth and the crisp/ classical clustering validity indices. While all internal validity indices have failed in finding the ground truth in the datasets Glass and Wine, in the Seeds dataset, only Gap statistics

could find the actual number of clusters. However, in the German Credit dataset, Gap statistics was the only index which failed in finding the accurate number of clusters according to the label variable.

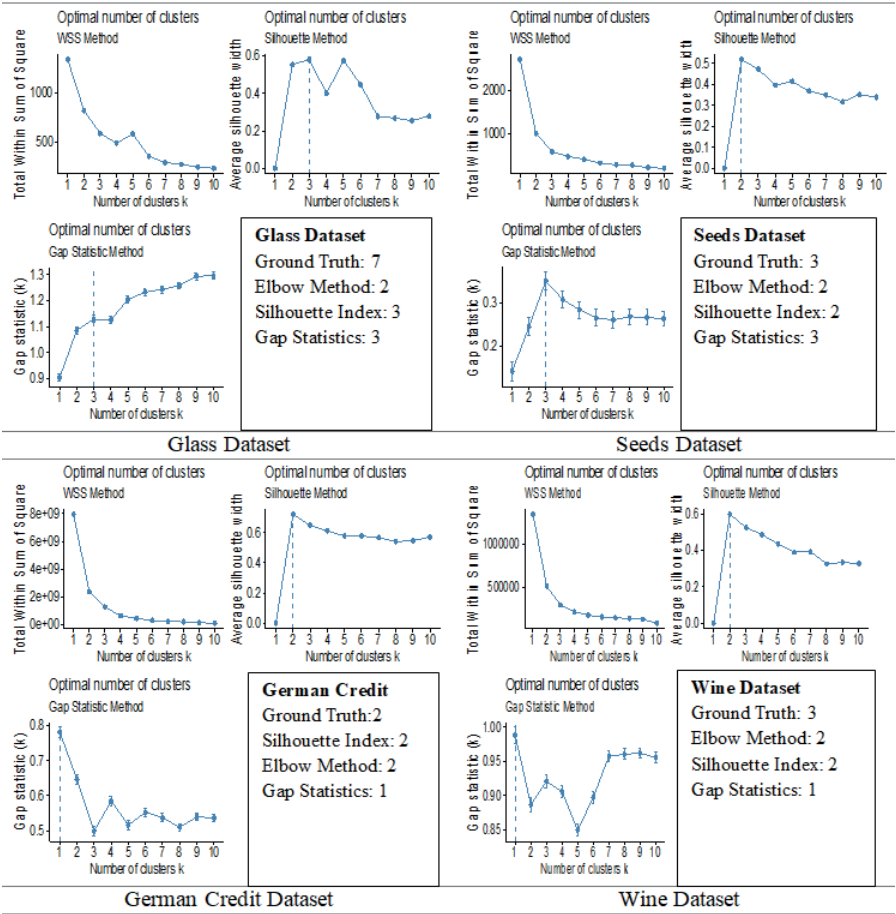


Figure 2 Results for Crisp Internal Validity Methods

As seen in Table 2, 3, 4 and 5 the relevant value according to the previous table (Table 1) was extracted for each fuzzy index for the FCM algorithm. In the Glass dataset, as the ground truth is 7, one can see that TSS (Tang, Sun & Sun Index) which searches the minimal value has matched the ground truth while neither the crisp indices or the other fuzzy indices could. In the seeds dataset it seems that PBMF and APD could find the ground truth 3. Looking at the German Credit dataset, one can see that all except for TSS, PBMF, FSIL and APD could find the ground truth, however, in the later

algorithms we see that these indices also manage to find the accurate cluster number. Lastly, in the Wine Quality dataset FHV and APD have managed to find the actual cluster number which was 3.

Table 2 Fuzzy Internal Validity Indices for FCM Algorithm for the Glass Dataset

FCM	c = 3	c = 4	c = 5	c = 6	c = 7
PC	0.756	0.634	0.499	0.493	0.499
MPC	0.635	0.513	0.374	0.392	0.415
PE	0.461	0.690	0.948	0.996	1.008
XB	0.129	0.591	2.989	2.358	1.972
K	27.920	129.532	659.322	531.306	466.117
TSS	25.907	111.814	11.988	8.628	1.377
CL	0.708	0.564	0.390	0.395	0.408
FS	-714772.886	-599838.578	-472174.360	-466378.017	-471225.827
PBMF	23.935	33.632	9.300	13.757	23.556
FSIL	0.818	0.622	0.414	0.401	0.463
FHV	0.000	0.000	0.000	0.000	0.000
APD	854420626.495	661626128.419	323505171.749	148062214.934	103475658.018

Table 3 Fuzzy Internal Validity Indexes for FCM Algorithm for Seeds Dataset

F C M	c = 2	c = 3	c = 4	c = 5
PC	0.805	0.726	0.639	0.575
MPC	0.610	0.589	0.519	0.469
PE	0.322	0.500	0.691	0.841
XB	0.102	0.151	0.164	0.265
K	21.694	32.615	35.570	58.111
TSS	21.159	31.005	11.396	29.573
CL	0.735	0.668	0.577	0.508
FS	-31937.055	-29588.269	-25970.056	-23820.749
PBMF	47.012	109.262	45.939	95.275
FSIL	0.791	0.744	0.675	0.623
FHV	0.00000171	0.00000177	0.00000262	0.0000028
APD	111,363,680.158	130,082,900.134	81,572,202.790	89,941,575.179

Table 4 Fuzzy Internal Validity Indexes for FCM Algorithm for German Credit

FCM	c=2	c=3	c=4	c=5
PC	0.900	0.829	0.813	0.773
MPC	0.800	0.743	0.750	0.716
PE	0.177	0.311	0.354	0.444
XB	0.053	0.083	0.109	0.173
K	52.718	83.807	112.806	181.378
TSS	52.468	26.446	3.711	3.029
CL	0.871	0.787	0.776	0.732
FS	-9888988248.183	-11621096150.883	-12951305710.689	-12078601698.260
PBMF	44863314.899	108280617.789	134662779.782	208025590.231
FSIL	0.899	0.856	0.842	0.816
FHV	382229.742	544645.189	644889.918	696087.279
APD	0.002	0.003	0.006	0.001

Table 5 Fuzzy Internal Validity Indexes for FCM Algorithm for The Wine Dataset				
FCM	c = 2	c = 3	c = 4	c = 5
PC	0.843	0.771	0.713	0.666
MPC	0.687	0.656	0.617	0.582
PE	0.262	0.418	0.547	0.655
XB	0.088	0.131	0.148	0.186
K	100.987	150.567	171.123	216.905
TSS	100.724	66.412	20.299	32.581
CL	0.787	0.719	0.661	0.609
FS	-1930172.422	-2066430.093	-2015191.039	-1974403.848
PBMF	4309.187	9469.070	14165.947	13527.127
FSIL	0.834	0.789	0.751	0.715
FHV	0.000008	0.000008	0.000009	0.000010
APD	165506620.897	275607986.305	75829566.987	288738803.942

One can view the different clustering results as the cluster number changes. The x and y values have been chosen as to visualize the sepeation in the best way. One can remind themselves that the ground truth for the datasets were 7, 3, 2, 3 for Glass, Seeds, German Credit and Seeds respectively.

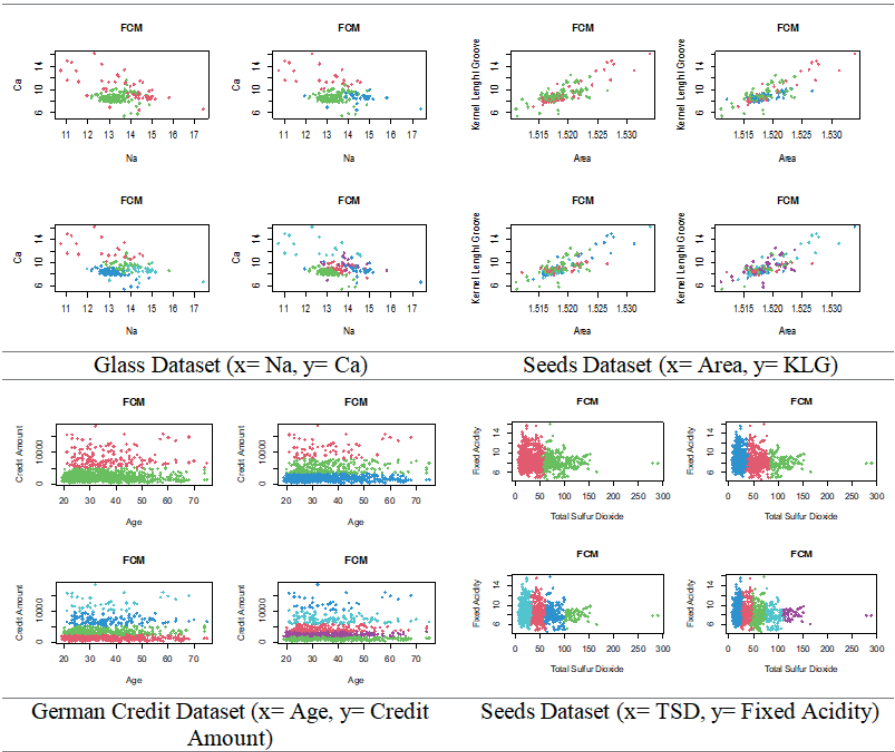


Figure 6 Fuzzy C-Means Clustering $C=2, 3, 4, 5$ for all dataset

Figure 4 Values of Fuzzy Validity Indices for FCM, PCM and UFPC of the Seeds Dataset

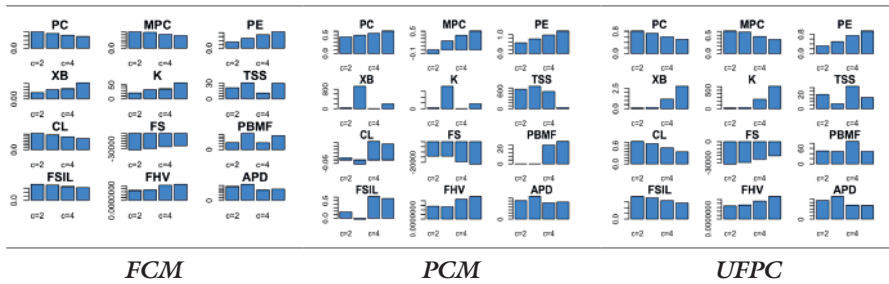


Figure 5 Values of Fuzzy Validity Indices for FCM, PCM and UFPC of the German Credit Dataset

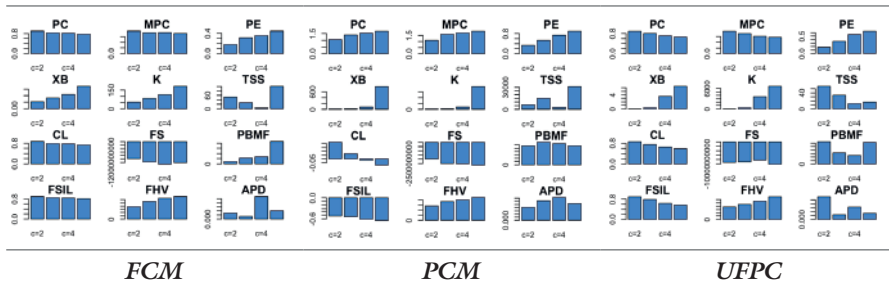
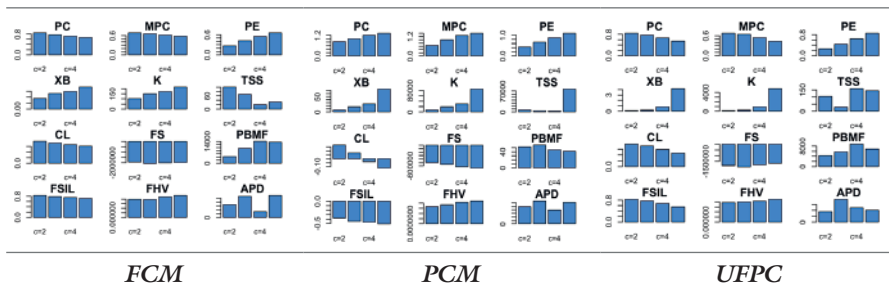


Figure 6 Values of Fuzzy Validity Indices for FCM, PCM and UFPC of the Wine Dataset



The above Figures have been summarized in the below tables for the comparison. The cluster numbers for the optimal values have been noted and those who have hit the mark of the ground truth have been written in bold, alongside the fuzzy clustering indices which have succeeded with atleast one algorithm. It could be noted that APD (Average Partition Density) have performed best in all datasets by succeeding to find the ground truth in numerous algorithms, showing its stability throughout algorithms. On the other hand, it could be seen that some indices performed better with

particular algorithms. Such as Fuzzy Silhouette Index (FSIL) that had manage to select the accurate cluster number mostly with the PCM algorithm however, it could not be calculated with the PCM objective function in the Glass dataset (yet again it has accurately stated the cluster number with the UFPC algorithm).

Table 7 Comparison of Internal Validity Indices for Glass

Cluster Number	FCM	PCM	UFPC
PC	3	7	3
MPC	3	7	4
PE	3	3	3
XB	3	3	3
K	3	3	3
TSS	7	3	7
CL	3	7	4
FS	6	3	7
PBMF	4	3	4
FSIL	5	NA	3
FHV	5	4	3
APD	7	5	4

Table 8 Comparison of Internal Validity Indices for Seeds

Cluster Number	FCM	PCM	UFPC
PC	2	5	2
MPC	2	5	2
PE	2	2	2
XB	2	4	2
K	2	4	2
TSS	4	5	3
CL	2	4	2
FS	5	2	5
PBMF	3	5	4
FSIL	5	3	5
FHV	2	3	2
APD	3	3	3

Table 9 Comparison of Internal Validity Indices for German Credit

Cluster Number	FCM	PCM	UFPC
PC	2	5	2
MPC	2	5	2
PE	2	2	2
XB	2	2	2
K	2	2	2
TSS	4	4	4
CL	2	2	2
FS	2	2	4
PBMF	5	3	2
FSIL	5	2	5
FHV	2	2	2
APD	4	4	2

Table 10 Comparison of Internal Validity Indices for Wine Quality

Cluster Number	FCM	PCM	UFPC
PC	2	5	2
MPC	2	5	2
PE	2	2	2
XB	2	2	2
K	2	2	2
TSS	4	4	3
CL	2	2	2
FS	2	2	5
PBMF	4	3	4
FSIL	5	2	5
FHV	3	2	2
APD	3	3	3

DISCUSSION AND CONCLUSION

Before moving forward with the conclusion, it could be useful to view the summary of the evaluation of this paper, so as to conclude in objective remarks. One can benefit from Table 11 and Table 12 to recall the essential information concerning the algorithms and the datasets analysis results.

Table 11 Comparison of Alternative FCM Algorithms

Method	FCM	PCM	UFPC
Parameters	Number of clusters (c), fuzzifier (m)	c, fuzzifier (m), typicality degree (t), typicality exponent (η)	c, fuzzifier (m), typicality degree (t), typicality exponent (η), weights (α, β)
Novelties	Introduced fuzzy membership	Introduced typicality degree (t) and typicality exponent (η) to model noise and outliers	Combined membership and typicality parameters in one objective function
Strengths	Simple, efficient, interpretable	Robust to noise and outliers via typicality	Handles both overlapping clusters and noise
Weaknesses	Sensitive to noise and outliers	Prone to coincident clusters, sensitive to setting of t and η	Needs careful tuning of multiple parameters: m, t, η , α , β
Optimal Dataset	Well-separated, compact clusters	Datasets with overlap or noise, where fuzzy membership is insufficient	Complex datasets involving both overlap and noise simultaneously

Table 12 Comparison of Datasets

Dataset	Glass	Seeds	German Credit	Wine Quality
Sample Size	214	210	1000	1143
Variable Count	9	7	4	12
Silhouette Score	0.58	0.52	0.65	0.6
Optimal Cluster	3	2	2	2
Ground Truth	7	3	2	3
Successful Indices	FCM: TSS, APD PCM: PC, MPC, CL UFPC: TSS, FS	FCM: PBMF, APD PCM: FSIL, FHV, APD UFPC: TSS, APD	FCM: all except TSS, FS, PBMF, FSIL, APD PCM: all except PC, MPC, TSS, PBMF, APD UFPC: all except TSS, FS, FSIL	FCM: FHV, APD PCM: PBMF, APD UFPC: TSS, APD

Peculiarities	The fuzzy indices had lots of variation, Mostly negative correlation	High positive correlation and one variable has hight negative correlation. FHV was very close in FCM and UFPC while FS estimated pretty closely in PCM but they were incorrect in the end	Large dataset with few dimensions; best quality clusters, mostly negative and slightly positive correlation	Mostly negative and nearly no correlation between chosen variables
Best Method	FCM or UPFC with TSS	PCM with APD	UFPC with PE	FCM, PCM or UFPC with TSS

In conclusion, it has been investigated that some indices behave better with particular algorithms. This alligns with the argument of Krisphapuram, R., and Keller, M. (1996) mentioning an algorithm not giving the correct result does not necessarily mean that the algorithm is poorly designed, but rather that it has been missapplied in that case. Thus by this reference one can infer that as diverse algorithms have different assumptions to be fulfilled, this also means that they have peculiar instances that enable them to perform their best. Therefore, just as datasets have different features that are not known beforehand, it might be suggested to apply not only multiple methods, but also multiple typologies of an algorithm that performs better in different instances, like in the case of FCM, PCM and UFPC. With domain knowledge, these complex datasets can be interpreted better. Finally, for further research, analysts and researchers can apply the algorithms in further contexts and maybe in some challenging datatypes like audio or image to test their limits.

References

- Cebeci, Z. (2020). fcvaid: An R Package for Internal Validation of Probabilistic and Possibilistic Clustering. 3. 11-27. Retrieved from <http://dx.doi.org/10.35377/saucis.03.01.664560>
- Charytanowicz, M., Niewczas, J., Kulczycki, P., Kowalski, P., & Łukasik, S. (2010). *Seeds* [Dataset]. UCI Machine Learning Repository. <https://doi.org/10.24432/C5H30K>
- Cortez, P., Cerdeira, A., Almeida, F., Matos, T., & Reis, J. (2009). *Modeling wine preferences by data mining from physicochemical properties*. UCI Machine Learning Repository. <https://archive.ics.uci.edu/ml/datasets/wine+quality>
- German, B. (1987). *Glass Identification* [Dataset]. UCI Machine Learning Repository. <https://doi.org/10.24432/C5WW2P>
- Hofmann, H. (1994). *Statlog (German Credit Data)* [Dataset]. UCI Machine Learning Repository. <https://doi.org/10.24432/C5NC77>
- Dunn, J. C. (1973). *A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters*. Journal of Cybernetics, 3(3), 32–57.
- Krishnapuram, R., & Keller, J. M. (1993). A possibilistic approach to clustering. IEEE Transactions on Fuzzy Systems, 1(2), 98–110. doi:10.1109/91.227387
- Krishnapuram, R., & Keller, J. M. (1996). The possibilistic C-means algorithm: insights and recommendations. IEEE Transactions on Fuzzy Systems, vol. 4, no. 3, pp. 385-393, Aug., doi: 10.1109/91.531779.
- Yang, M., & Wu, K. (2005). Unsupervised possibilistic clustering. Pattern Recognition, 39(1), 5–21. <https://doi.org/10.1016/j.patcog.2005.07.005>

Acknowledgment

The authors would like to thank anonymous reviewers for valuable suggestions

Conflict of Interest

The authors declare no conflict of interest

Author Contributions

Berna ÖZBAŞARAN: writing- original draft, methodology, application

Gözde ULUTAGAY: conceptualization, writing- review & editing, supervision

Evaluation of the Performance of BRICS-T Countries in the Context of the Economic Freedom Index

Bilal Saraç¹

Çağlar Karamaşa²

Abstract

Since 1995, the Heritage Foundation has presented the factors that directly contribute to the economic freedom and prosperity of the international community in detail through the Economic Freedom Index (EFI). While the index evaluates the progress or decline of countries around the world, it focuses on key indicators of economic well-being such as economic growth, poverty reduction, longevity and health, as well as environmental protection. Through the index, the degree of economic freedom is relatively calculated as a significant factor in national development and prosperity on a global scale, and countries are ranked accordingly. Rankings of economic freedoms provide critical insights for countries, scholars, policymakers—in short, all stakeholders—to understand the impact of the measured criteria on economic growth. These rankings can guide the development of solutions to issues such as poverty and economic contraction faced by societies. Furthermore, through cross-country comparisons, differences in economic freedoms can be identified, and much can be learned about how to improve economic welfare and development. For these reasons, assessing the economic freedoms of countries holds vital importance. In this study, the economic freedom levels of the BRICS countries along with Turkey for the year 2025 are determined. Due to the presence of multiple indicators in the index and the involvement of multiple countries in terms of economic freedoms, the MEREC and WENSLO integrated AROMAN approach as Multi-Criteria Decision-Making (MCDM) techniques has been preferred. While the objective MCDM methods MEREC and WENSLO were used to determine the weights of the evaluation criteria, the AROMAN approach was employed to rank the BRICS-T countries.

- 1 Anadolu University, Faculty of Economics and Administrative Sciences, Department of Quantitative Methods, Türkiye
- 2 Anadolu University, Faculty of Economics and Administrative Sciences, Department of Quantitative Methods, Türkiye

INTRODUCTION

Since 1995, the Heritage Foundation has detailed the factors contributing directly to the economic freedom and prosperity of the international community through the Economic Freedom Index (EFI). The Index focuses on fundamental economic well-being measures such as economic growth, poverty reduction, longevity, and health, along with various social indicators, while evaluating the progress or regression of countries worldwide. With the Index, the degrees of economic freedom are relatively calculated and countries are ranked as an important factor in the development and welfare of nations at the global level.

Economic freedom conceptually refers to the levels of freedom individuals, entrepreneurs, and businesses have to use their time and money in the way they believe is best for themselves, free from unnecessary government restrictions and plunder (Erdal, 2004). In other words, people are free to work, produce, consume, and invest in the way they believe is most productive (Beach and Miles, 2006). It is widely believed that economic freedom brings prosperity to countries. Additionally, a country's economic freedom is closely related to financial stability and the development of capital markets (Luo, 2014).

Economic growth is primarily a result of the benefits generated by capital investments, profits from trade, the discovery of advanced products, lower-cost production methods, and achieving better outcomes. In the literature, there are many studies indicating that countries with greater economic freedom grow faster and achieve higher per capita income than those with less freedom. Considering the source of growth and prosperity, it is evident that the advancements in quality of life move in parallel with the increase in economic freedom. Personal choices, voluntary exchange regulated by markets, free entry to markets, competition, and the protection of assets from occupation by others can be listed as the fundamental components of economic freedom (Lawson, 2009).

Based on information obtained from The Heritage Foundation, EFI has been designed to evaluate the consistency of economic policies in countries with a free market economy. The positive relationship between economic freedom and socioeconomic goals is defined by EFI. Because there is a strong correlation between the principles of economic freedom and a cleaner environment, healthier societies, greater per capita wealth, more development, and the eradication of poverty. EFI consists of twelve indicators across four main categories. These categories can be summarized as Government Size (government spending, fiscal health, and tax burden),

Regulatory Efficiency (monetary freedom, freedom of labor, and freedom of business), Open Markets (freedom of investment, freedom of trade, and financial freedom), and Rule of Law (judicial effectiveness, property rights, and government integrity). Each indicator within these categories has values ranging from 0 to 100 (Kılıcı, 2019).

EFI, which is frequently used by scientists, economists, and investors, is seen as the key to attracting investment and creating sound public policies. In recent years, the index has gained critical importance for countries seeking to enhance their international appeal and branding, as well as for global investors. As a measure of economic freedom and competitiveness, the EFI can help countries market themselves (Olson, 2014).

Economic freedom rankings provide critical information to countries, scientists, policymakers, and all stakeholders to understand the impact of the evaluated criteria on economic growth. These rankings can guide the development of solutions to issues such as poverty and economic contraction that societies face. Additionally, by making comparisons between countries, differences in economic freedoms can be identified, and much can be learned about improving economic prosperity and development. For these reasons, evaluating the economic freedoms of countries is of critical importance. Especially for developing countries to have a greater say in economic matters and to achieve economic expansion, they need to perform well on the indicators included in the EFI. At this point, the BRICS countries, along with Turkey, which together constitute about a quarter of the world economy, are striving to increase their growth rates, achieve development, and attract investors in the context of global power. In line with these goals, evaluating the current economic freedoms of these countries is of great importance. Therefore, in this study, the economic freedom levels of Turkey for the year 2025, along with the BRICS countries, have been determined. Due to the index having multiple indicators and multiple countries in terms of economic freedoms, the study preferred the AROMAN method integrated with MEREC and WENSLO from the Multi-Criteria Decision Making (MCDM) methods. In the weighting of the evaluation criteria, the objective MCDM methods MEREC and WENSLO were used, and the AROMAN approach was considered for the ranking of BRICS-T countries.

LITERATURE REVIEW

In the literature, there are many studies that demonstrate a positive relationship between economic freedom and economic growth. (Razmi and Refaci, 2013; Akin et al., 2014; Le Roux, 2015; Nadeem et al., 2019;

Thuy, 2021). At the same time, there are many studies that provide evidence that economic freedom leads to better living standards, improves social welfare, causes income growth, and enhances incentives, productive efforts, and resource utilization efficiency (Hall and Lawson, 2014; Gehring, 2013; Erdal, 2004; Easton and Walker, 1997). Along with these studies, there are also works examining the extent to which economic freedom policies affect carbon emissions (Abeka et al., 2022), the relationship between renewable energy consumption and economic freedom (Dumitrescu and Hurlin, 2012), the relationship between foreign direct investments and economic freedom (Ciftci and Durusu-Ciftci, 2022), and the relationship between energy intensity, carbon emissions, and economic freedom (Mahmood, Shahab and Shahbaz, 2022).

The presence of multiple indicators in EFI has increased interest in using MCDM methods to determine the economic levels of countries. Balkan countries (Puška, Štilić and Stojanović, 2023), OPEC countries (Ecer and Zolfani, 2022), European Union countries (Karaköy et al., 2023), countries located in the European continent (Altın, 2020), and all countries examined by the Heritage Foundation (Atan, Atan and Gökmen, 2024) have utilized MCDM methods in determining their economic levels. In this context, the lack of evaluation of the economic freedom levels of the BRICS countries, which aim to have a greater say in international matters alongside Turkey and to create an alternative to the Western world's dominance over the global financial system, can be identified as a gap in the literature. The pressure exerted on the global economy by trade wars, the impact of economic decisions made during the Covid-19 pandemic, the conflicts between Russia and Ukraine, and in the Middle East, as well as geopolitical tensions, have influenced the preference for BRICS countries. Additionally, due to reasons such as the support of free markets by economic freedom and the provision of innovative and practical solutions necessary for sustainable development, MCDM methods have been utilized in determining the levels of economic freedom of BRICS countries.

MATERIAL AND METHODS

Material

EFI is considered the most important method used to measure the economic freedom levels of countries, and therefore, there are many studies related to the index. However, the number of studies that evaluate the levels of economic freedom of countries through monitoring and control, and analyze the degree of achievement of goals, is virtually nonexistent. In this

context, the indicators included in the EFI, conceptualized by the Heritage Foundation, have been integrated into the methodology and method. With the help of data on the indicators included in the index, the focus has been on calculating the importance levels of the criteria that play a role in ensuring economic freedom, and determining the economic freedom rankings of Turkey along with the BRICS countries. For this purpose, MCDM methods have been utilized.

Methods

The Collection of the Data

The Heritage Foundation publishes a dataset of 12 macroeconomic indicators for 184 countries each year, laying the foundation for economic growth, achieving prosperity, and improving quality of life. EFI is seen as an objective tool for analyzing the economic freedom levels of countries. Additionally, the data included in the index is an important resource for a comprehensive analysis of the economic and political situations of the countries. In this context, the study utilized secondary data published by the Heritage Foundation for the year 2025 (<https://www.heritage.org/index/pages/all-country-scores>).

MEREC

Keshavarz-Ghorabae et al. (2021) proposed MEREC (MEthod based on the Removal Effects of Criteria) that considers a novel principle for obtaining objective weights related to criteria. Apart from other weighting methods, MEREC handles removal effects of each criterion on the aggregate performance of alternatives in order to compute the weights. In this method when a criterion having greater weight is removed leads to more effects on aggregate performances of alternatives. Causality concept is the basis of this method. The aim of this method is to determine the criterion having the greatest impact on the overall performance of alternatives and assign the most weight to this criterion. A logarithmic function with equal weights is handled to assess the aggregate performance of alternatives in this study. Besides the absolute deviation measure is considered for determining the effects of removing each criterion. Steps of the MEREC method can be stated as below (Keshavarz-Ghorabae et al., 2021):

Step 1. A decision matrix indicating each alternative's ratings or values related to each criterion is created. Consider that there are m alternatives and n criteria, the initial decision matrix D is constructed as below:

$$D = \begin{bmatrix} d_{11} & \cdots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{m1} & \cdots & d_{mn} \end{bmatrix}, \quad \begin{matrix} i = 1, \dots, m \\ j = 1, \dots, n \end{matrix} \quad (1)$$

The elements of this matrix (d_{ij}) need to be greater than zero ($d_{ij} > 0$). If negative or zero values exist in the decision matrix such as in our study, appropriate transformation technique needs to be applied for obtaining positive ones.

Step 2. The elements of the decision matrix is normalized via simple linear normalization as specified in Equation (2):

$$f_{ij} = \begin{cases} \frac{\min_i e_{ij}}{e_{ij}}, & j \in B \\ \frac{e_{ij}}{\max_i e_{ij}}, & j \in C \end{cases} \quad (2)$$

where B denotes the set of benefit-based criteria, C shows the set of cost-based criteria. Additionally, the elements of the normalized decision matrix are represented by f_{ij} .

Step 3. The overall performance of alternatives is acquired via a logarithmic measure with equal criteria weights as seen in Equation (3). A non-linear function is considered to form this measure. By taking the normalized values into the account, it can be implied that the lower values of f_{ij} lead greater performance values of alternatives (S_i).

$$S_i = \ln \left(1 + \left(\frac{1}{n} \sum_j |\ln(f_{ij})| \right) \right) \quad (3)$$

Step 4. The performance of alternatives in terms of removing different criterion at each step is computed via logarithmic measure as seen in Equation (4).

$$S'_{ij} = \ln \left(1 + \left(\frac{1}{n_{k, k \neq j}} \sum_{k, k \neq j} |\ln(f_{ij})| \right) \right) \quad (4)$$

According to Equation (4), S'_{ij} shows the overall performance of i th alternative related to the removal of j th criterion.

Step 5. The removal effect of j th criterion is calculated according to Equation (5) shown as below:

$$E_j = \sum_i |S'_{ij} - S_i| \quad (5)$$

where E_j represents the effect of removing j th criterion.

Step 6. Objective weight for each criterion (w_j) is computed via Equation (6) by utilizing the E_j values obtained in the fifth step.

$$w_j = \frac{E_j}{\sum_k E_k} \quad (6)$$

WENSLO

Pamucar et al. (2024) developed a novel objective weighting method namely WENSLO (Weights by Envelope and Slope) based on the ratio between the envelope and slope of each criterion. If the value of envelope is high and the value of slope is low, related criterion has a greater weight. The main advantage is based on the fact that the weights of criteria are independent of judgments of decision makers. Besides the criteria tendency (being benefit or cost based) has not any impact on the computation of WENSLO. In other words, the normalization procedure for input data does not based on criteria preferences. It shows that the proposed method is very reliable and stable. Also, this method can be considered for determining criteria weights and applicable for any MCDM problem where decision maker wants to avoid subjectivity. WENSLO can capture the behaviour of criterion without considering the randomness of it and can be achieved by accumulating the normalized criterion data. Steps of WENSLO can be stated as below (Pamucar et al. 2024):

Step 1. Creating a decision-making matrix: In the first step, initial decision matrix $\mathfrak{R}(A, C)_{m \times n}$ is created as below:

$$\mathfrak{R}(A, C) = [\zeta_{ij}]_{m \times n} = \begin{bmatrix} A / C & C_1 & C_2 & \cdots & C_j \\ target & maxmin & maxmin & \cdots & maxmin \\ A_1 & \zeta_{11} & \zeta_{12} & \cdots & \zeta_{1j} \\ A_2 & \zeta_{21} & \zeta_{22} & \cdots & \zeta_{2j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_i & \zeta_{i1} & \zeta_{i2} & \cdots & \zeta_{in} \end{bmatrix} \quad (7)$$

Where A_1, A_2, \dots, A_m represents a collection of alternatives, m shows the number of alternatives, C_1, C_2, \dots, C_n represents a collection of criteria, n shows the number of criteria, *maxmin* target relates to direction of each criterion. If the criterion aims to achieve max value, benefit-based environment is valid. On the contrary if the criteria aim to achieve min value, cost-based environment can be applicable too, ζ_{ij} shows the estimated value of i th alternative in terms of j th criterion.

Step 2. Normalizing the input data: Each criterion is characterized by attribute namely dimension that leads the multidimensional vector space of decision matrix. In such a situation, any kind of calculation creates a big problem. So, a nondimensional decision matrix is formed in terms of normalization process for overcoming this difficulty. The normalization process is applied to the elements of $\mathfrak{R}(A, C)$ matrix. Within this regard, Equation (8) given below is considered for linear normalization:

$$z_{ij} = \frac{\zeta_{ij}}{\sum_{i=1}^m \zeta_{ij}} \quad \forall j \in [1, 2, \dots, n] \quad (8)$$

A normalized decision matrix as the outcome is formed as follows:

$$Z(A, C) = [z_{ij}]_{m \times n} = \begin{bmatrix} A/C & C_1 & C_2 & \cdots & C_j \\ \text{target} & \text{maxmin} & \text{maxmin} & \cdots & \text{maxmin} \\ A_1 & z_{11} & z_{12} & \cdots & z_{1j} \\ A_2 & z_{21} & z_{22} & \cdots & z_{2j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_i & z_{i1} & z_{i2} & \cdots & z_{in} \end{bmatrix} \quad (9)$$

where z_{ij} shows the component of the normalized decision matrix Z and $0 < z_{ij} < 1$.

Step 3. Computing the criterion class interval: The final ranking of alternatives in terms of the given set of criteria is related to the impact of each criterion on the ranking. This impact is called as the criteria weights in decision theory. Defining these weights without subjectivity is considered as important task. The vector space of the normalized decision matrix is shown as below:

$$\begin{bmatrix} A \\ A_1 \\ A_2 \\ \vdots \\ A_i \end{bmatrix} = \left(\begin{bmatrix} C_1 \\ z_{11} \\ z_{21} \\ \vdots \\ z_{i1} \end{bmatrix}, \begin{bmatrix} C_2 \\ z_{12} \\ z_{22} \\ \vdots \\ z_{i2} \end{bmatrix}, \dots, \begin{bmatrix} C_j \\ z_{1j} \\ z_{2j} \\ \vdots \\ z_{ij} \end{bmatrix} \right) \quad (10)$$

In terms of Sturges' rule, the dimension of the j th criterion class interval Δz_j is computed according to Equation (11):

$$\Delta z_j = \frac{\max_{i=1,2,\dots,m} z_{ij} - \min_{i=1,2,\dots,m} z_{ij}}{1 + 3.322 * \log(m)} \quad \forall j \in [1, 2, \dots, n] \quad (11)$$

According to the Equation (11), the class intervals of the C_1 and C_2 are computed as follows:

$$\Delta z_1 = \frac{\max_{i=1,2,\dots,m} z_{i1} - \min_{i=1,2,\dots,m} z_{i1}}{1 + 3.322 * \log(m)}; \quad \Delta z_2 = \frac{\max_{i=1,2,\dots,m} z_{i2} - \min_{i=1,2,\dots,m} z_{i2}}{1 + 3.322 * \log(m)} \quad (12)$$

Step 4. Computing the criterion slope: The slope of the criterion is computed according to the Equation (13):

$$\tan\varphi_j = \frac{\sum_{i=1}^m z_{ij}}{(m-1)\Delta z_j} \quad \forall j \in [1, 2, \dots, n] \quad (13)$$

Step 5. Defining the criterion envelope: The sum of the partial Euclidean distance between two successive normalized values of the j th criterion that equals to the total Euclidean distance between the first and last normalized values, and defined according to Equation (14):

$$E_j = \sum_{i=1}^{m-1} \sqrt{(z_i + \Delta_j - z_i, j)^2 + z_j^2} \quad 1, 2, \dots, \forall j \in [1, \dots, n] \quad (14)$$

The envelope of criterion can be considered to represent the total Euclidean distance. The shape of the envelope likes a zig-zag line indicating its values.

Step 6. Determining the envelope-slope ratio: The ratio of the total Euclidean distance to the slope of the criterion as a numerical value is computed as follows:

$$q_j = \frac{E_j}{\tan\varphi_j} \quad \forall j \in [1, 2, \dots, n] \quad (15)$$

Step 7. Computing the criteria weights: The weight of the criterion is computed according to Equation (16).

$$w_j = \frac{q_j}{\sum_{j=1}^n q_j} \quad \forall j \in [1, 2, \dots, n] \quad (16)$$

Following to that artificial accumulation and the error of the artificial accumulation (difference between the real and artificial accumulated value) are obtained via Equations (17) and (18).

$$\tilde{z}_j(i, \Delta z_j) = \tan\varphi_j (i \cdot \Delta z_j) \quad i = 0, 1, 2, \dots, m-1 \quad \forall j \in [1, n] \quad (17)$$

$$\varepsilon_i(\Delta z_j) = z_{ij} - \boxed{\times} z_{ij} \quad i = 1, 2, \dots, m-1 \quad \forall j \in [1, n] \quad (18)$$

Then, to verify the validity of an artificial process related to accumulation, two widely used methods namely mean-squared error (MSE) and coefficient of correlation (r) are taken into the account. The results can be considered as valid if MSE and r close to 0 and 1 respectively.

Integrated Weights

Integrated objective weight for each criterion by considering the weights of MEREC ($w_{j,merec}$) and WENSLO ($w_{j,wenslo}$) can be calculated as follows (Zavadkas and Podvezko, 2016):

$$w_{j,integrated} = \frac{w_{j,merec} w_{j,wenslo}}{\sum_{j=1}^m w_{j,merec} w_{j,wenslo}} \quad (19)$$

AROMAN

Boskovic et al. (2023) proposed the AROMAN (Alternative Ranking Order Method Accounting for Two-Step Normalization), which couples the linear and vector normalization techniques for acquiring precise data structures used in further calculation. The AROMAN method combines the normalized data from two-step normalization and acquires an average matrix from normalized data. Steps of AROMAN method can be summarized as below (Bošković et al. 2023; Nikolić et al., 2023):

Step 1. Constructing the initial decision-making matrix by considering the input data: The initial decision matrix D_{mn} is created via the input data d_{11}, \dots, d_{mn} seen as Equation (20).

$$D = \begin{bmatrix} d_{11} & \cdots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{m1} & \cdots & d_{mn} \end{bmatrix}, \quad \begin{matrix} i = 1, \dots, m; \\ j = 1, \dots, n \end{matrix} \quad (20)$$

Step 2. Normalizing the input data: The input data is formed in intervals between 0 and 1 via normalization process. Two types of normalization are taken into the account and can be seen as Equations (21) and (22).

Step 2.1. Normalization 1 (Linear):

$$t_{ij} = \frac{d_{ij} - \min_i d_{ij}}{\max_i d_{ij} - \min_i d_{ij}}, \quad i = 1, \dots, m; j = 1, \dots, n; \quad (21)$$

Step 2.2. Normalization 2 (Vector):

$$t_{ij}^* = \frac{d_{ij}}{\sqrt{\sum_{i=1}^m d_{ij}^2}}; i = 1, \dots, m; j = 1, \dots, n; \quad (22)$$

Aforementioned two types of normalization techniques are considered for both criterion types (min and max).

Step 2.3. Obtaining aggregated averaged normalization: The aggregated averaged normalization is found by Equations (23):

$$t_{ij}^{norm} = \frac{\beta t_{ij} + (1 - \beta) t_{ij}^*}{2}; i = 1, \dots, m; j = 1, \dots, n; \quad (23)$$

where t_{ij}^{norm} shows the aggregated averaged normalization and β is a weighting factor that is considered for each type of normalization varying from 0 to 1. In this study β is considered as 0.5.

Step 3. Acquiring the weighted aggregated normalized decision-making matrix: The aggregated averaged normalized decision-making matrix is multiplied by the criteria weights for obtaining a weighted decision-making matrix seen as Equation (24):

$$\square t_{ij} = W_{ij} \cdot t_{ij}^{norm}; i = 1, \dots, m; j = 1, \dots, n; \quad (24)$$

Step 4. Separately summarizing the normalized weighted values for the criteria type min (L_i) and the type max (A_i): This procedure is computed via Equations (25) and (26).

$$L_i = \sum_{j=1}^n \square t_{ij}^{(min)}; i = 1, \dots, m; j = 1, \dots, n; \quad (25)$$

$$A_i = \sum_{j=1}^n \square t_{ij}^{(max)}; i = 1, \dots, m; j = 1, \dots, n; \quad (26)$$

Step 5. Raising the obtained sum of L_i and A_i values to the degree of: This procedure is calculated by applying Equations (27) and (28).

$$L_i^{\wedge} = L_i^{\lambda} = \left(\sum_{j=1}^n \square^{(min)} t_{ij} \right)^{\lambda}, i = 1, \dots, m \quad (27)$$

$$A_i^{\wedge} = A_i^{1-\lambda} = \left(\sum_{j=1}^n \square^{(max)} t_{ij} \right)^{1-\lambda}, i = 1, \dots, m \quad (28)$$

where λ shows the coefficient degree of the criterion type. In this study the parameter λ is considered as 0.5 due to including both criterion types. In order to avoid undefined results caused by considering solely of benefit or cost criteria, the parameter λ can be accepted as 0.5.

Step 6. Computing the difference between the values A_i^{\wedge} and L_i^{\wedge} and obtaining the final ranking: This procedure is calculated via Equation (29).

$$R_i = e^{(A_i^{\wedge} - L_i^{\wedge})}, i = 1, \dots, m \quad (29)$$

where R_i shows the final ranking of alternatives.

RESULTS

The Heritage Foundation evaluates all the criteria used to assess countries' economic freedom performance with equal importance. EFI is an index that measures how far people can go in economic actions in terms of free markets, free trade, and private property within the scope of their fundamental rights and freedoms, and compares countries accordingly. Therefore, the index plays an important role in improving sustainable development and human development and in achieving economic goals. For these reasons, addressing the performance of countries based on economic metrics is a complex task (Gwartney, 2008; Ott, 2018). Therefore, under this title, the application of the proposed methodologies of the MEREK, WENSLO, and AROMAN methods to rank Turkey along with the BRICS countries in terms of economic freedoms has been described.

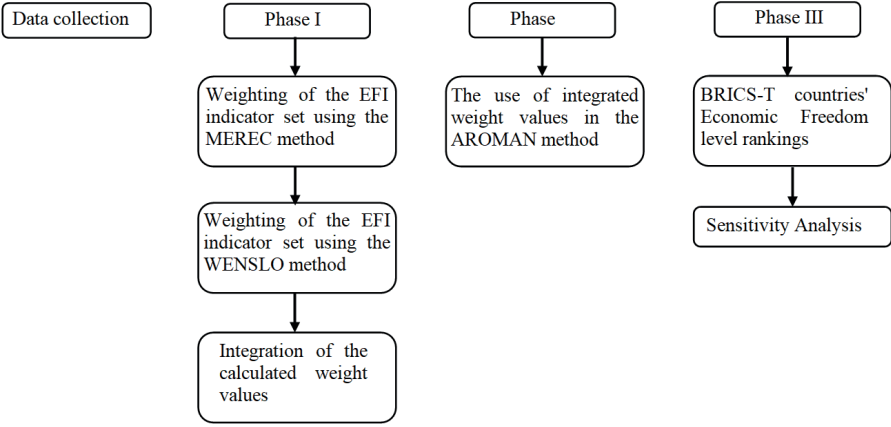


Figure 1. The framework for ranking BRICS-T countries by EFI indicators

In this context, the indicator weights were first determined using the MEREC and WENSLO methods to achieve the set goals. Later, AROMAN was integrated into these methods, and the 10 BRICS countries (Brazil (BRE), China (CHN), Egypt (EGY), Ethiopia (ETH), India (IND), Indonesia (INS), Iran (IR), Russia (RUS), South Africa (SA), United Arab Emirates (UAE), along with Turkey (TR), were ranked according to their economic freedom levels. A summary of the proposed MEREC-WENSLO-AROMAN model for evaluating the economic freedom levels of the specified countries can be expressed in Figure 1.

Determining the Weights of EFI Criteria with MEREC and WENSLO

As previously mentioned, the EFI published by the Heritage Foundation includes 12 economic indicators. The derivation of the importance levels of these indicators with MEREC and WENSLO constitutes the first step of the developed methodology. The definitions of the relevant indicators are provided in Table 1.

Table 1. Economic Freedom Index indicator set

C1	Property Rights	C7	Business Freedom
C2	Judicial Effectiveness	C8	Labor Freedom
C3	Government Integrity	C9	Monetary Freedom
C4	Tax Burden	C10	Trade Freedom
C5	Government Spending	C11	Investment Freedom
C6	Fiscal Health	C12	Financial Freedom

After this stage, the dataset consisting of 12 indicators and 11 alternatives (BRICS-T countries), in other words, the decision matrix, is shown in Table 2.

Table 2. Economic freedom indicators of BRICS-T countries by 2025

Optimization	Criteria											
	Max	Max	Max	Max	Max	Max	Max	Max	Max	Max	Max	Max
Countries	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
BRE	51	54	36	75	44	50	67	57	74	72	40	40
CHN	47	37	40	72	67	11	68	58	75	74	20	20
EGY	41	22	30	86	82	21	48	44	63	60	65	50
ETH	28	20	33	78	95	81	45	39	50	57	30	20
IND	51	53	38	71	74	6	72	59	70	61	40	40
INS	40	45	40	82	91	85	73	59	78	79	50	60
IR	23	19	16	81	94	84	38	44	39	56	5	10
RUS	19	28	23	88	62	98	51	59	62	69	30	30
SA	44	58	45	68	69	46	64	71	75	68	40	40
UAE	65	35	66	97	82	98	84	62	81	78	50	60
TR	41	24	34	72	72	82	59	48	38	73	70	60

The importance levels of the economic freedom criteria obtained by integrating the MEREC and WENSLO methods are presented in Table 3.

Table 3. Economic Freedom Index criteria integrated importance degree values

Criteria	MEREC	WENSLO	Mean Squared Error	Coefficient of Correlation	Integrated Weight	Rank
C1	0.0770	0.0840	0.0044	0.9887	0.0485	6
C2	0.0603	0.1089	0.0039	0.9927	0.0493	5
C3	0.0782	0.1102	0.0026	0.9920	0.0647	4
C4	0.0157	0.0097	0.0015	0.9976	0.0011	12
C5	0.0570	0.0290	0.0011	0.9982	0.0124	9
C6	0.2249	0.1958	0.0055	0.9900	0.3304	1
C7	0.0449	0.0397	0.0022	0.9952	0.0134	8
C8	0.0324	0.0262	0.0014	0.9969	0.0064	10
C9	0.0515	0.0418	0.0039	0.9936	0.0161	7
C10	0.0195	0.0095	0.0018	0.9966	0.0014	11
C11	0.2078	0.2037	0.0024	0.9906	0.3176	2
C12	0.1308	0.1414	0.0019	0.9935	0.1387	3
			Average of MSE: 0.0027	Average of CC: 0.9938		

According to the Table 3 while the average of MSE is found as 0.0027, the average of CC is obtained as 0.9938. These obtained values show that the results can be considered as valid and hypotenuse can be regarded for computing the slope related to the criterion (Pamucar et al., 2024). The most important indicator among the economic freedom criteria, determined by integrating the MEREC and WENSLO techniques, is Fiscal Health (C6) (0.3304). This criterion is followed by Investment Freedom (C11) (0.3176) and Financial Freedom (C12) (0.1387). However, the Tax Burden (C4) (0.0011) has been objectively determined to be the least important indicator among the 12 indicators.

Ranking of BRICS-T Countries According to Their Economic Freedom Levels

It is important to follow a data-driven methodological approach to determine the current status of countries and compare their performance according to economic freedom criteria, without the influence of decision-makers or economic authorities. Additionally, since the ranking of countries' levels of economic freedom inherently involves multiple criteria and alternatives, it is suitable for examination within the framework of decision theory using MCDM. Among the MCDM methods, the AROMAN method has been preferred due to reasons such as offering an innovative approach, including a two-step normalization process to provide fair and impartial comparisons among alternatives, and facilitating a comprehensive ranking by considering the relative importance levels of the criteria (Bošković, et al., 2023).

Table 4. Ranking of BRICS-T countries according to their level of economic freedom

	Final Ranking (Lambda=0.5)	Rank
BRE	1.5695	6
CHN	1.3621	11
EGY	1.5623	7
ETH	1.5286	8
IND	1.4738	9
INS	1.7022	3
IR	1.4127	10
RUS	1.5774	4
SA	1.5701	5
UAE	1.7554	1
TR	1.7225	2

According to the AROMAN method from Table 4, the countries with the highest levels of economic freedom among the BRICS-T countries are the United Arab Emirates (UAE), Turkey (TR), and Indonesia (INS), in that order. The country ranked last in the economic freedom ranking is China (CHN).

Sensitivity Analysis

In order to check the stability and validity of the proposed model a sensitivity analysis is conducted by changing λ and β values. According to the original case the parameters of λ and β are considered as 0.5. In terms of sensitivity analysis, the proposed model is examined for other scenarios via an increment value of 0.1. While the sensitivity results related to the changing λ values are presented in Figure 2, the obtained results in terms of changing β values are shown in Figure 3.

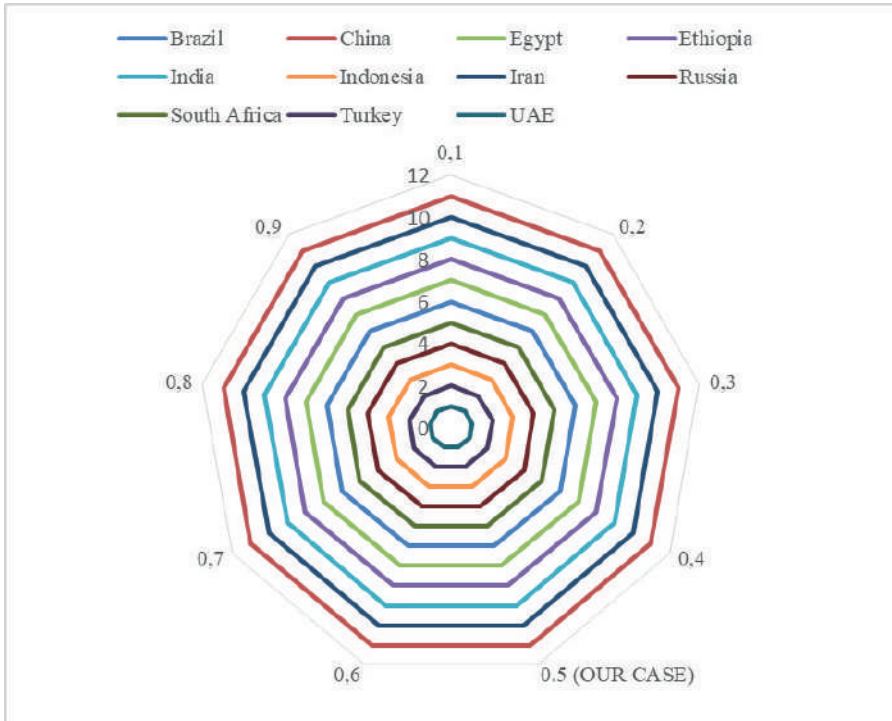


Figure 2. Ranking changes for alternatives by changing λ values

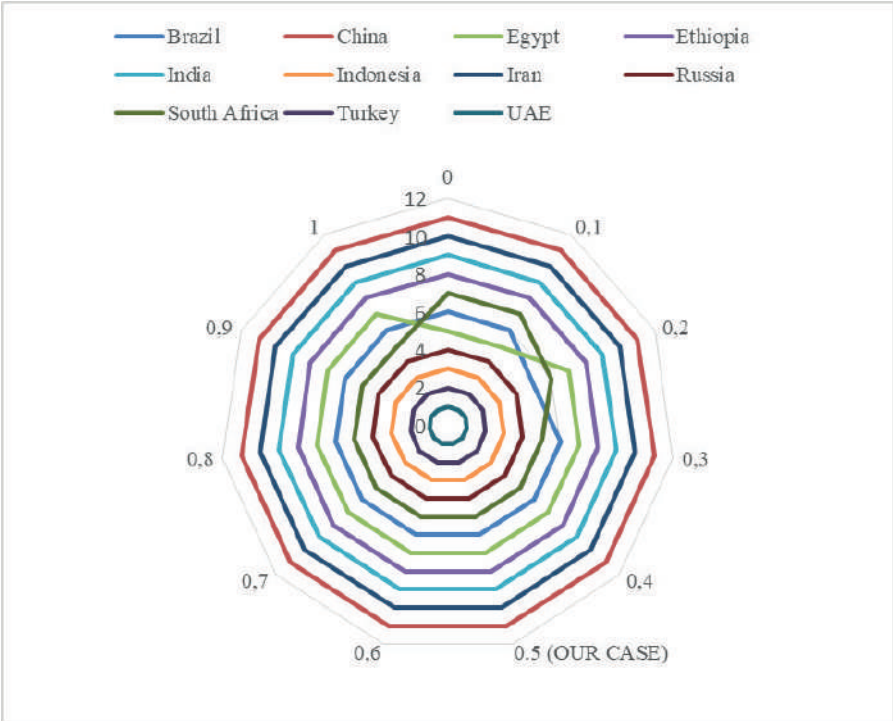


Figure 3. Ranking changes for alternatives by changing β values

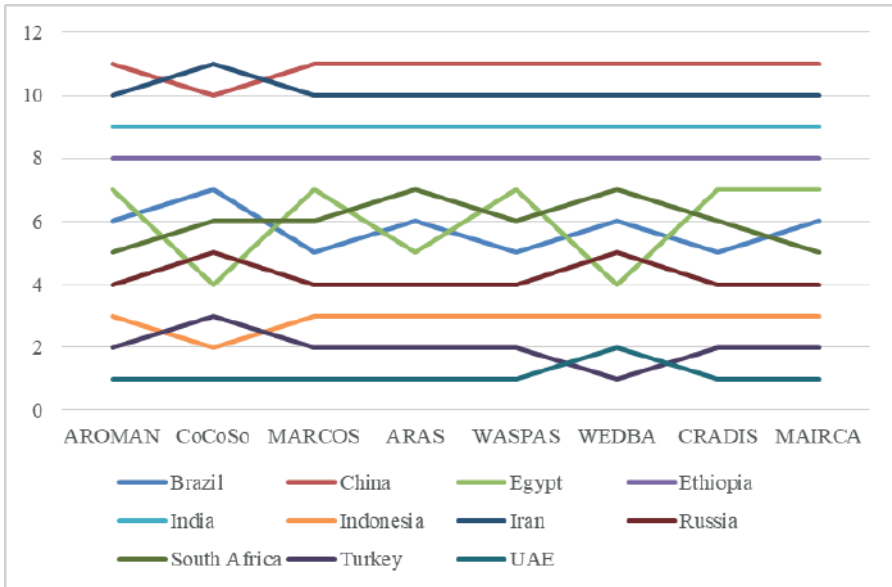
According to the Figure 2 same ranking results related to the alternatives are obtained for all scenarios. Slight variations related to the alternatives (BRE, EGY and SA) are seen in first three scenarios in terms of Figure 3. When the sensitivity results in terms of changing λ and β values are considered, it is understood that the proposed model is stable and valid.

Comparison Analysis

In order to test the validity related to the decision models the ranking results of the AROMAN and other MCDM methods consisting of CoCoSo, MARCOS, ARAS, WASPAS, WEDBA, CRADIS and MAIRCA are compared and the analysis is made in terms of Spearman's rank correlation coefficient values. While the obtained Spearman's rank correlation coefficient values are depicted in Table 5, the ranking results related to the comparison analysis are shown in Figure 4.

Table 5. Spearman's rank correlation coefficient values related to the methods compared

	AROMAN	CoCoSo	MARCOS	ARAS	WASPAS	WEDBA	CRADIS	MAIRCA
AROMAN	1,0000							
CoCoSo	0,9273	1,0000						
MARCOS	0,9909	0,9182	1,0000					
ARAS	0,9636	0,9636	0,9727	1,0000				
WASPAS	0,9909	0,9182	1,0000	0,9727	1,0000			
WEDBA	0,9273	0,9545	0,9364	0,9818	0,9364	1,0000		
CRADIS	0,9909	0,9182	1,0000	0,9727	1,0000	0,9364	1,0000	
MAIRCA	1,0000	0,9273	0,9909	0,9636	0,9909	0,9273	0,9909	1,0000

*Figure 4. BRICS-T ranking results with MCDM methods according to EFI indicators*

According to the results related to the Spearman's analysis it was found statistically significant (at the 1% level) and a very high correlation between the ranking of various MCDM methods that shows the validity, applicability and reliability of the proposed model.

Rank Reversal Test

As additional validation analysis, a rank reversal test is conducted for examining whether the ranking results obtained with the existing model give a stable response to sudden changes. The rank reversal test performed in this study is based on the progressive deletion of sub-optimal alternatives and checking the ranking of the remaining ones. The results related to rank

reversal test are shown in Figure 5. The first scenario (Scenario 0) is the initial ranking order acquired via proposed model. The other scenarios (Scenario 1-10) are constructed by progressively removing the alternatives ranked last in the previous rankings from the model one by one.

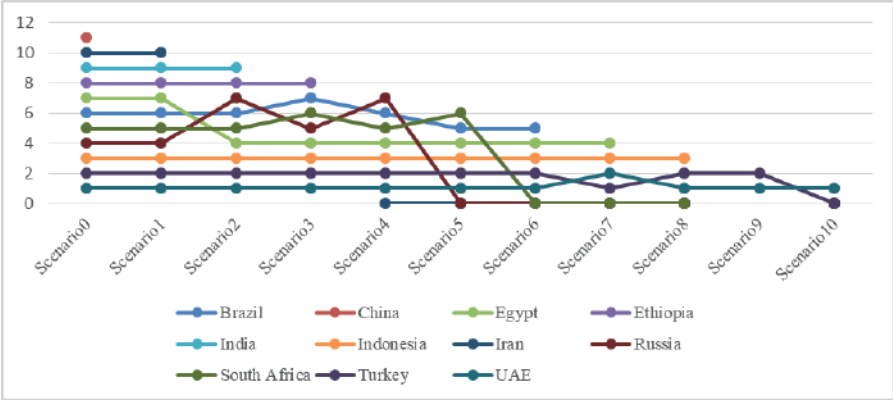


Figure 5. Results related to rank reversal test

According to the Figure 5 it is understood that the proposed model shows a significant sensitivity to rank reversal cases. In terms of the proposed model the ranking of six alternatives (BRE, EGY, RUS, SA, UAE and TR) changed during the process of progressive deletion of sub-optimal alternatives.

CONCLUSION

This study, which focuses on evaluating the overall performance of the economic freedoms of BRICS countries along with Turkey, integrates the MEREC, WENSLO, and AROMAN methods to create a multi-criteria decision-making model. This model can also be described as a decision support system. Using the 2025 data published by the Heritage Foundation, the MEREC and WENSLO objective weighting methods were employed to calculate the importance levels of 12 economic freedom indicators, and the AROMAN technique was used to determine the ranking of the specified countries based on their levels of economic freedom. In light of the results obtained, the most effective criteria playing a role in improving the economic freedom levels of countries have been identified as Fiscal Health, Investment Freedom, and Financial Freedom. In other words, countries that perform well on these criteria are expected to rank higher in the standings. However, the United Arab Emirates (UAE) is the country with the highest level of

economic freedom, as it has shown the highest performance in 11 out of 12 indicators. Moreover, it can be observed that Turkey (TR) and Indonesia (INS) perform well in Fiscal Health, Investment Freedom, and Financial Freedom, as well as other criteria (Table 2). China (CHN), Iran (IR), and India (IND) can be said to have low levels of economic freedom. The bureaucratic problems, corruption, and attempts to restrict the freedoms of the middle class in these countries are believed to be the reasons for this outcome. Moreover, the war between Russia and Ukraine can be said to have caused a decline in Russia's level of economic freedom.

In future studies, countries located on the same continent can be evaluated within the framework of their economic freedom levels. Additionally, in future studies, expert opinions can be utilized, and by addressing the uncertainty in expert opinions, Fuzzy MCDM methods can be employed. In this way, economic freedom can be examined by bringing it closer to the real-world problem. With the help of cluster analyses, countries showing similar economic freedom performances can be identified. With forecasting models, the future economic freedom performances of countries can be predicted.

References

- Abeka MJ, Amoah EK, Owusu Appiah M, Gatsi JG, Obuobi NK, Boateng E, 2022. Economic institutions, political institutions and renewable energy production in Africa, *Journal of African Business*, 23 (4): 1049-1066.
- Akin CS, Aytun C, Aktakas BG, 2014. The impact of economic freedom upon economic growth: an application on different income groups, *Asian Economic and Financial Review*, 4 (8): 1024-1039.
- Altın H, 2020. Ekonomik Özgürlük Endeksinin Çok Kriterli Karar Verme Yöntemleriyle Analizi, *Uluslararası Ekonomi İşletme ve Politika Dergisi*, 4 (2): 441-460.
- Atan M, Atan S, Gökmen Ş, 2024. Ranking of countries according to the index of economic freedom with multicriteria decision making methods, *Athens journal of business & economics*, 10 (2): 109-120.
- Beach WW, Miles MA, 2006. Explaining the factors of the index of economic freedom. 2006 Index of Economic Freedom, 55-76.
- Bošković S, Švadlenka L, Jovčić S, Dobrodolac M, Simić V, 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.
- Ciftci C, Durusu-Ciftci D, 2022. Economic freedom, foreign direct investment, and economic growth: The role of sub-components of freedom, *The Journal of International Trade & Economic Development*, 31 (2): 233-254.
- Dumitrescu EI, Hurlin C, 2012. Testing for Granger non-causality in heterogeneous panels, *Economic modelling*, 29 (4): 1450-1460.
- Easton ST, Walker MA, 1997. Income, growth, and economic freedom, *The American Economic Review*, 87 (2): 328-332.
- Ecer F, Zolfani SH, 2022. Evaluating economic freedom via a multi-criteria mercedmna model-based composite system: Case of OPEC countries, *Technological and Economic Development of Economy*, 28 (4): 1158-1181.
- Erdal F, 2004. Economic Freedom and Economic Growth: A time series evidence from Italian economy, Dostupno na: <http://www.etsg.org/ETSG2004/Papers/erdal.pdf> [20. 10. 2009.].
- Gehring K, 2013. Who benefits from economic freedom? Unraveling the effect of economic freedom on subjective well-being, *World Development*, 50: 74-90.
- Gwartney J, 2008. Economic freedom of the world: 2008 annual report. The Fraser Institute.

- Hall JC, Lawson RA, 2014. Economic freedom of the world: An accounting of the literature, *Contemporary economic policy*, 32 (1): 1-19.
- Karaköy Ç, Ulutaş A, Karabasevic D, Üre S, Bayrakçıl AO, 2023. The evaluation of economic freedom indexes of EU countries with a grey hybrid MCDM model, *Romanian Journal of Economic Forecasting*, 26 (1), 129.
- Keshavarz-Ghorabace M, Amiri M, Zavadskas EK, Turskis Z, Antucheviciene J, 2021. Determination of objective weights using a new method based on the removal effects of criteria (MEREC). *Symmetry*, 13 (4): 525.
- Kılıcı EN, 2019. Analysis of the relationship between economic freedom index and stock market indices; evidence from Turkey, *Maliye ve Finans Yazıları*, (111): 117-134.
- Le Roux P, 2015. The impact of economic freedom on economic growth in the SADC: an individual component analysis, *Studies in Economics and Econometrics*, 39 (2), 41-61.
- Luo Y, 2014. Economic freedom, financial crisis and stock volatilities in emerging markets, *International Journal of Financial Management*, 4 (1), 1-10.
- Mahmood MT, Shahab S, Shahbaz M, 2022. The relevance of economic freedom for energy, environment, and economic growth in Asia-Pacific region, *Environmental Science and Pollution Research*, 29 (4): 5396-5405.
- Miller T, Kim AB, 2015. Principles of economic freedom. Heritage Foundation, 11-17.
- Nadecem M, Yang J, Akhtar T, Dong W, Niazi M, 2019. Does really economic freedom matter for growth in South Asia? Empirical evidences from pre-economic crises and post-economic crises period, *Asian Economic and Financial Review*, 9 (1): 52.
- Nikolić I, Milutinović J, Božanić D, Dobrodolac M, 2023. Using an Interval Type-2 Fuzzy AROMAN Decision-Making Method to Improve the Sustainability of the Postal Network in Rural Areas. *Mathematics*, 11 (14): 3105.
- Olson R, 2014. Using the Index of Economic Freedom: A practical guide. Heritage. org [Web]. Retrieved from: <https://www.heritage.org/international-economics/report/using-the-index-economic-freedom-practical-guide>. Date of access: 15/05/2025.
- Ott J, 2018. Measuring economic freedom: Better without size of government, *Social Indicators Research*, 135 (2): 479-498.
- Pamucar D, Ecer F, Gligorić Z, Gligorić M, Devci M, 2024. A Novel WEN-SLO and ALWAS Multicriteria Methodology and Its Application to Green Growth Performance Evaluation. *IEEE Transactions on Engineering Management*, 71: 1-16.

- Puška A, Štilić A, Stojanović I, 2023. Approach for multi-criteria ranking of Balkan countries based on the index of economic freedom, *Journal of Decision Analytics and Intelligent Computing*, 3 (1), 1-14.
- Razmi MJ, Refaei R, 2013. The effect of trade openness and economic freedom on economic growth: the case of Middle East and East Asian countries, *International Journal of Economics and Financial Issues*, 3 (2): 376-385.
- Thuy DTB, 2021. Impacts of economic freedom on economic growth in developing countries. *Global changes and sustainable development in Asian emerging market economies Vol. 1: Proceedings of EDESUS 2019* (pp. 35-44). Cham: Springer International Publishing.
- Zavadskas EK, Podvezko V, 2016. Integrated determination of objective criteria weights in MCDM, *International Journal of Information Technology & Decision Making*, 15 (2): 267-283.

Conflict of Interest

The authors have declared that there is no conflict of interest.

Author Contributions

BS: Project idea, data collection, data analysis, interpreting the result, literature search, writing the manuscript; ÇK: Project idea, data collection, data analysis, interpreting the result, literature search, writing the manuscript and final check of the manuscript.

Multi-Period Production and Inventory Planning in Textile Industry: A Case Study of Textile Company

Çağdaş Yıldız¹

Adem Tüzemen²

Abstract

In the textile sector, seasonal demand fluctuations and variable raw material prices increase the strategic importance of production and inventory planning. In this study, a mathematical model has been developed to optimize the production and inventory planning for a three-month period of a textile company operating in Turkey. The model aims to minimize normal working hours, overtime, and inventory holding costs for three main product groups: T-shirts, trousers, and shirts. The mathematical model, solved using LINGO software, optimizes the company's three-month (June-August 2025) production and inventory planning. In the study, demand forecasts, labor requirements, minimum stock levels, and maximum storage capacity constraints were defined for each product. According to the optimization results, an optimal production and inventory plan was obtained with a total cost of 2,137,450 TL. This plan envisages production of 7,800 units for T-shirts, 4,200 units for trousers, and 4,650 units for shirts. Capacity utilization analysis showed that normal working hour capacity was utilized at 100%, while overtime capacity was used at 55% only in August for T-shirt production. Stock level analysis revealed that minimum stock levels were maintained for T-shirts and shirts, while high stock levels were maintained for trousers in June and July in preparation for high demand in August. Cost analysis showed that 91.3% of the total cost consisted of normal working hour production costs, 7.4% of overtime costs, and 1.3% of inventory holding costs. As a result of the study, strategic recommendations were presented

1 Dr., 60000, Tokat, Turkey

2 Tokat Gaziosmanpaşa University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Production Management and Marketing, 60000, Tokat, Turkey

regarding the company's production capacity management, inventory optimization, cost reduction, and demand management. The presented model is adaptable to other enterprises in the textile sector and provides a scientific approach to production and inventory planning decisions suitable for seasonal demand fluctuations.

INTRODUCTION

The textile industry faces significant challenges in production planning and inventory management due to seasonal demand fluctuations, variable raw material costs, and intense market competition. Effective production planning is crucial for textile manufacturers to maintain profitability while meeting customer demands in a timely manner (Singh and Chadha, 2016). This study addresses the production and inventory planning challenges of a textile company operating in Turkey, focusing on optimizing production schedules and inventory levels over a three-month planning horizon.

The textile sector in Turkey represents a significant portion of the country's manufacturing industry and export revenue. However, companies in this sector often struggle with balancing production capacity, inventory costs, and meeting fluctuating customer demands (Erdil and Erdil, 2017). Traditional production planning approaches frequently result in either excess inventory, leading to increased holding costs, or insufficient production, resulting in lost sales opportunities and decreased customer satisfaction.

Linear programming and mathematical optimization techniques have been widely applied in manufacturing industries to address production planning problems. These techniques allow companies to determine optimal production quantities, inventory levels, and resource allocation while considering various constraints such as production capacity, storage limitations, and demand requirements. Previous studies have demonstrated that mathematical optimization can lead to significant cost savings and operational improvements in manufacturing environments (Krynke, M., & Mielczarek, 2018; Perez et al., 2021).

This study aims to develop and implement a multi-period production and inventory planning model for a textile company in Turkey. The model focuses on three main product categories: T-shirts, trousers, and shirts, with the objective of minimizing total production, overtime, and inventory holding costs while satisfying customer demand over a three-month planning horizon (June-August 2025). The results provide valuable insights for production managers and decision-makers in the textile industry regarding capacity utilization, inventory management, and cost optimization strategies.

MATERIAL AND METHODS

Material

The material for this study consists of production and inventory data from a textile manufacturing company operating in Turkey. The company produces three main product categories: T-shirts, trousers, and shirts, and experiences significant seasonal demand fluctuations. The planning horizon covers three months (June-August 2025), for which detailed demand forecasts were provided by the company's marketing department based on historical sales data and market trends.

The dataset includes comprehensive information on product specifications, labor requirements, production costs, inventory holding costs, and capacity constraints. This information was collected through structured interviews with production managers and analysis of the company's Enterprise Resource Planning (ERP) system records. The company's production facility operates with both regular time and overtime capacity, with different associated labor costs.

For each product category, specific parameters were identified including unit production costs, labor hours required per unit, inventory holding costs per unit per month, initial inventory levels, and minimum required inventory levels. These parameters form the foundation of the mathematical optimization model developed in this study.

Methods

The Collection of the Data

Data collection was conducted through a systematic approach involving multiple sources to ensure accuracy and reliability. Primary data was collected through structured interviews with key personnel including the production manager, inventory manager, and finance director. These interviews provided insights into the operational constraints, cost structures, and strategic priorities of the company.

Secondary data was extracted from the company's ERP system, covering historical production records, inventory levels, demand patterns, and cost information for the past three years. This historical data was essential for validating the parameters used in the optimization model and for assessing the seasonal patterns in demand.

Demand forecasts for the three-month planning horizon were developed using a combination of time series analysis, moving average methods, and

expert judgments from the sales department. The forecasts were validated against historical accuracy metrics to ensure reliability.

Production capacity data was collected through time studies and analysis of production line capabilities. Regular time capacity was established at 4,000 hours per month, while maximum overtime capacity was set at 1,200 hours per month. Labor costs were determined based on current wage rates and overtime premium policies.

Statistical Analysis

A linear programming model was formulated to optimize the production and inventory planning decisions. The model was implemented using LINGO optimization software version 19.0, which employs the simplex method for solving linear programming problems and branch-and-bound techniques for integer programming components.

The mathematical formulation of the model includes an objective function that minimizes the total cost, which comprises regular time production costs, overtime production costs, and inventory holding costs:

This mathematical model has been compiled from similar studies in the literature (Chan et al., 2017; Sepehri et al., 2021).

Sets and Indices

$U = \{1, 2, 3\}$: $U = \{1, 2, 3\}$: Product set; $u = 1$; $u = 1$: T-shirt,
 $u = 2$; $u = 2$: Trousers, $u = 3$; $u = 3$: Shirt

$A = \{1, 2, 3\}$: $A = \{1, 2, 3\}$: Month set; $a = 1$; $a = 1$: June, $a = 2$; $a = 2$:
 July, $a = 3$; $a = 3$: August

Parameters

d_{ua} : Demand quantity for product u in month a (units)

h_u : Unit inventory cost for product u (TL/unit/month)

p_u : Unit production cost for product u (TL/unit)

l_u : Unit working time requirement for product u (hours/unit)

c_a^n : Normal working capacity in month a (hours)

c_a^o : Overtime working capacity in month a (hours)

w^n : Normal working unit cost (TL/hour)

w^o : Overtime working unit cost (TL/hour)

i_u^0 : Initial inventory level for product u (units)

i_u^{min} : Minimum inventory level for product u (units)

C^{max} : Maximum total warehouse capacity (units)

Decision Variables

x_{ua}^n : Quantity of product u produced in month a during normal hours (units)

x_{ua}^o : Quantity of product u produced in month a during overtime (units)

i_{ua} : Inventory level of product u at the end of month a (units)

Objective function:

$$Min Z = \sum_{u \in U} \sum_{a \in A} (w^n \cdot l_u \cdot x_{ua}^n + w^o \cdot l_u \cdot x_{ua}^o + h_u \cdot i_{ua}) \quad (1)$$

Object to:

$$\sum_{u \in U} l_u \cdot x_{ua}^n \leq c_a^n \quad \forall_{a \in A} \quad (2)$$

$$\sum_{u \in U} l_u \cdot x_{ua}^o \leq c_a^o \quad \forall_{a \in A} \quad (3)$$

$$i_{u1} = i_u^0 + x_{u1}^n + x_{u1}^o - d_{u1} \quad \forall_{u \in U} \quad (4)$$

$$i_{ua} = i_{u,a-1} + x_{ua}^n + x_{ua}^o - d_{ua} \quad \forall_{u \in U}, \forall_{a \in A}, A \setminus \{1\} \quad (5)$$

$$i_{ua} \geq i_u^{min} \quad \forall_{u \in U}, \forall_{a \in A} \quad (6)$$

$$\sum_{u \in U} i_{ua} \leq C^{max} \quad \forall_{a \in A} \quad (7)$$

$$x_{ua}^n, x_{ua}^o, i_{u,a} \in \mathbb{Z}^+ \quad \forall_{u \in U}, \forall_{a \in A} \quad (8)$$

Equation (1) represents the objective function that minimizes total production, working, and inventory costs. Equation (2) ensures normal working capacity constraints are not exceeded in each month. Equation (3) ensures overtime working capacity constraints are not exceeded in each month. Equation (4) defines inventory balance for the first month. Equation

(5) defines inventory balance for subsequent months. Equation (6) ensures minimum inventory levels are maintained. Equation (7) ensures maximum warehouse capacity is not exceeded. Equation (8) enforces non-negativity and integer requirements for all decision variables.

RESULTS

The optimization model was successfully solved, yielding an optimal production and inventory plan for the three-month planning horizon. The total cost of the optimal solution was 2,137,450 TL, which includes regular time production costs, overtime production costs, and inventory holding costs.

Production Plan

The optimal production quantities for each product and time period are presented in Tables 1 and 2, distinguishing between regular time and overtime production.

Table 1. Regular Time Production (Units)

Product	June	July	August
T-shirts	2,200	3,800	480
Trousers	1,850	750	1,600
Shirts	850	1,500	2,300

Table 2. Overtime Production (Units)

Product	June	July	August
T-shirts	0	0	1,320
Trousers	0	0	0
Shirts	0	0	0

The results (Table 1 and Table 2) indicate that regular time production was prioritized across all periods, with overtime production only utilized for T-shirts in August. This aligns with cost-efficient production planning, as regular time production has a lower cost per unit compared to overtime production.

Inventory Levels

The optimal inventory levels at the end of each period are presented in Table 3.

Table 3. End-of-Period Inventory Levels (Units)

Product	Initial	June	July	August
T-shirts	500	200	200	200
Trousers	300	950	700	100
Shirts	200	150	150	150

The inventory levels (Table 3) show strategic inventory management decisions. For T-shirts and shirts, the inventory levels were maintained at the minimum required levels throughout the planning horizon, minimizing inventory holding costs. For trousers, higher inventory levels were maintained in June and July to prepare for the increased demand in August, demonstrating a build-up strategy.

Capacity Utilization

The capacity utilization for both regular time and overtime is presented in Tables 4 and 5.

Table 4. Regular Time Capacity Utilization

Month	Used (hours)	Total (hours)	Utilization Rate
June	4,000	4,000	100%
July	4,000	4,000	100%
August	4,000	4,000	100%

Table 5. Overtime Capacity Utilization

Month	Used (hours)	Total (hours)	Utilization Rate
June	0	1,200	0%
July	0	1,200	0%
August	660	1,200	55%

The capacity utilization analysis (Table 4) reveals that regular time capacity was fully utilized (100%) in all three months, indicating efficient use of available resources. Overtime capacity (Table 5) was only utilized in August (55%), specifically for T-shirt production, to meet the demand requirements while maintaining minimum inventory levels.

Cost Analysis

The breakdown of the total cost is presented in Table 6.

Table 6. Cost Breakdown

Cost Component	Amount (TL)	Percentage
Regular Time Production Cost	1,952,000	91.3%
Overtime Production Cost	158,400	7.4%
Inventory Holding Cost	27,050	1.3%
Total Cost	2,137,450	100%

The cost analysis (Table 6) shows that regular time production costs constitute the majority (91.3%) of the total cost, followed by overtime production costs (7.4%) and inventory holding costs (1.3%). This distribution reflects the model’s emphasis on minimizing higher-cost components, particularly overtime production and inventory holding.

DISCUSSION AND CONCLUSION

This study developed a multi-period production and inventory planning model for a textile company in Turkey, optimizing production quantities and inventory levels for three product categories over a three-month horizon. The results revealed several important findings that can be evaluated in light of existing literature.

The optimal solution demonstrated 100% utilization of regular time capacity across all three months, with overtime production only required for T-shirts in August (55% of available overtime capacity). This finding aligns with previous research by Zhang and Sun (2018), who emphasized the importance of maximizing regular time capacity before utilizing more expensive overtime production in manufacturing environments. This approach is supported by Ebrahim and Rasib (2017) and further validated by Fernandes et al. (2024) in their assessment of capacity adjustment strategies.

The inventory management strategies identified in this study reflect the principles of strategic inventory positioning discussed by Schwartz and Rivera (2010). For T-shirts and shirts, the model maintained minimum inventory levels throughout the planning horizon, while for trousers, a strategic inventory build-up was implemented in June and July to prepare for high August demand. This differentiated approach supports the assertion that product-specific inventory policies based on demand patterns yield superior results compared to uniform inventory strategies (Ziukov, 2015; Jauhari et al., 2023). This finding is consistent with advanced inventory management practices identified by Panigrahi et al. (2015).

Cost analysis revealed that regular time production costs constituted 91.3% of the total cost, followed by overtime production costs (7.4%) and

inventory holding costs (1.3%). This distribution is consistent with the findings of Öztürk (2021), who identified production costs as the dominant component in manufacturing optimization systems. The total optimal cost of 2,137,450 TL represents a significant improvement over traditional planning methods, as demonstrated in similar optimization studies by Strub et al. (2021) and supported by cost analysis principles established by Gim and Yoon (2012).

The strategic use of inventory to manage seasonal demand fluctuations, particularly for trousers, supports the findings of Mattsson (2010), who identified inventory build-up as a cost-effective strategy for managing predictable demand peaks. This approach is further validated by Namwad et al. (2024) in their optimization study and Nambiar et al. (2021) in their dynamic inventory allocation research. Similarly, the minimal use of overtime production aligns with cost optimization principles established in the literature.

In conclusion, this study demonstrates that mathematical optimization techniques can effectively address production planning and inventory management challenges in the textile industry. The findings contribute to the existing literature by providing empirical evidence of the benefits of integrated production and inventory planning in a seasonal demand environment. Future research should address the limitations of this study by incorporating demand uncertainty and time-varying costs to further enhance the applicability of optimization models in textile manufacturing.

References

- Chan FT, Li N, Chung SH, Saadat M, 2017. Management of sustainable manufacturing systems-a review on mathematical problems. *International Journal of Production Research*, 55(4): 1210-1225.
- Ebrahim Z, Rasib AA, 2017. Unnecessary Overtime as the Component of Time Loss Measures in Assembly Processes. *Journal of Advanced Manufacturing Technology (JAMT)*, 11(1): 37-48.
- Erdil M, Erdil A, 2017. Assessment of quality requirement and importance for textile industry in turkey. *PressAcademia Procedia*, 5(1): 58-66.
- Fernandes NO, Thürer M, Costa F, 2024. Work Faster, Work in Parallel, or Work Overtime? An Assessment of Short-Term Capacity Adjustments by Simulation. *Mathematics*, 12(16): 2515.
- Gim B, Yoon WL, 2012. Analysis of the economy of scale and estimation of the future hydrogen production costs at on-site hydrogen refueling stations in Korea. *International Journal of Hydrogen Energy*, 37(24): 19138-19145.
- Jauhari WA, Pujawan IN, Suef M, 2023. Sustainable inventory management with hybrid production system and investment to reduce defects. *Annals of Operations Research*, 324(1): 543-572.
- Krynke M, Mielczarek K, 2018. Applications of linear programming to optimize the cost-benefit criterion in production processes. *MATEC Web of Conferences*, Vol. 183, p. 04004, EDP Sciences.
- Mattsson SA, 2010. Inventory control in environments with seasonal demand. *Operations Management Research*, 3: 138-145.
- Nambiar M, Simchi-Levi D, Wang H, 2021. Dynamic inventory allocation with demand learning for seasonal goods. *Production and Operations Management*, 30(3): 750-765.
- Namwad RS, Mishra NK, Jain P, 2024. Optimizing Inventory Management with Seasonal Demand Forecasting in a Fuzzy Environment. *Journal Européen des Systèmes Automatisés*, 57(4).
- Öztürk H, 2021. Optimal production run time for an imperfect production inventory system with rework, random breakdowns and inspection costs. *Operational Research*, 21(1): 167-204.
- Panigrahi RR, Das JR, Jena D, Tanty G, 2015. Advance inventory management practices and its impact on production performance of manufacturing industry. *Journal of*, 11(6).
- Perez HD, Amaran S, Erisen E, Wassick JM, Grossmann IE, 2021. Optimization of extended business processes in digital supply chains using mathematical programming. *Computers & Chemical Engineering*, 152: 107323.

- Schwartz JD, Rivera DE, 2010. A process control approach to tactical inventory management in production-inventory systems. *International Journal of Production Economics*, 125(1): 111-124.
- Sepehri A, Mishra U, Sarkar B, 2021. A sustainable production-inventory model with imperfect quality under preservation technology and quality improvement investment. *Journal of Cleaner Production*, 310: 127332.
- Singh Z, Chadha P, 2016. Textile industry and occupational cancer. *Journal of Occupational Medicine and Toxicology*, 11: 1-6.
- Strub L, Kurth A, Loose SM, 2021. Effects of viticultural mechanization on working time requirements and production costs. *American journal of enology and viticulture*, 72(1): 46-55.
- Zhang R, Sun X, 2018. Integrated Production-Delivery Lot Sizing Model with Limited Production Capacity and Transportation Cost considering Overtime Work and Maintenance Time. *Mathematical Problems in Engineering*, 2018(1): 1569029.
- Ziukov S, 2015. A literature review on models of inventory management under uncertainty. *Business Systems & Economics*, 5(1): 26-35.

Conflict of Interest

The authors have declared that there is no conflict of interest.

Author Contributions

Both authors contributed equally to all aspects of this research including conceptualization, methodology, data collection, analysis, and writing of the manuscript.

Optimal Product Mix And Resource Allocation In Furniture Manufacturing Using Linear Programming: A Case Of Furniture Company

Çağdaş Yıldız¹

Adem Tüzemen²

Abstract

Increasing competitive conditions and changing customer demands in the furniture industry necessitate efficient use of production resources and optimal planning of product mix. In this study, a linear programming model is presented for profit maximization and balanced product mix objectives of a furniture manufacturing company. The model aims at optimal allocation of limited resources such as CNC machine time, labor hours, MDF, and fabric for three main product groups: sofa sets, dining tables, and wardrobes. The mathematical model, presented using LINGO software, optimizes the company's weekly production planning. In the study, minimum demand and maximum capacity constraints were defined for each product, and all resource limitations were included in the model. According to the optimization results, a production plan providing a weekly profit of 91,700 TL was obtained. This plan envisions the production of 17 sofa sets, 7 dining tables, and 6 wardrobes. Resource utilization analysis showed that CNC machine and labor hours were utilized at a rate of 99.2%, indicating these resources as the main limiting factors for production. MDF utilization rate was calculated as 92.7%, while fabric utilization rate was 96.9%. Product mix analysis revealed that sofa sets constitute 64.9% of the total profit. As a result of the study, strategic recommendations were presented to increase the company's resource utilization efficiency and achieve a more balanced product mix. The presented model can be adapted to other companies in the furniture industry and provides a scientific approach to production planning and resource allocation decisions.

1 Dr., 60000, Tokat, Turkey

2 Tokat Gaziosmanpaşa University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Production Management and Marketing, 60000, Tokat, Turkey

INTRODUCTION

The furniture manufacturing industry faces significant challenges in optimizing production processes due to increasing raw material costs, fluctuating customer demands, and intense market competition. Effective resource allocation and product mix planning are crucial for furniture manufacturers to maintain profitability while meeting customer requirements in a timely manner. This study addresses the production planning challenges of a furniture company operating in Turkey, focusing on optimizing product mix and resource allocation to maximize profit within existing constraints (Hongqiang et al., 2012; Sakib et al., 2024).

The furniture sector in Turkey represents a significant portion of the country's manufacturing industry and export revenue. According to recent industry reports, Turkish furniture exports reached approximately \$4.5 billion in 2024, making it one of the top ten furniture exporting countries globally. However, companies in this sector often struggle with balancing production capacity, raw material utilization, and meeting diverse customer demands (Erdinler and Koç, 2024). Traditional production planning approaches frequently result in either underutilization of critical resources or inefficient product mix decisions, leading to suboptimal profitability.

Linear programming and mathematical optimization techniques have been widely applied in manufacturing industries to address product mix and resource allocation problems. These techniques allow companies to determine optimal production quantities and resource allocation while considering various constraints such as production capacity, raw material availability, and demand requirements. Previous studies have demonstrated that mathematical optimization can lead to significant profit improvements and operational efficiencies in manufacturing environments (Vagaská and Gombár, 2021; Rabe et al., 2022).

The furniture manufacturing process involves multiple resources including machinery (particularly CNC machines), labor hours, and raw materials such as medium-density fiberboard (MDF) and fabric. The efficient allocation of these resources across different product lines presents a complex optimization challenge. With rising raw material costs in the 2025 Turkish economic environment, furniture manufacturers must make strategic decisions regarding which products to prioritize and how to allocate limited resources to maximize profitability.

This study aims to develop and implement a linear programming model for a furniture manufacturing company in Turkey. The model focuses on

three main product categories: sofa sets, dining tables, and wardrobes, with the objective of maximizing total profit while satisfying minimum demand requirements and respecting resource constraints over a weekly planning horizon. The results provide valuable insights for production managers and decision-makers in the furniture industry regarding optimal product mix, resource utilization, and profit maximization strategies.

MATERIAL AND METHODS

Material

The material for this study consists of production and resource data from a furniture manufacturing company operating in Turkey. The company produces three main product categories: sofa sets, dining tables, and wardrobes, and faces challenges in determining the optimal product mix under limited resources. The planning horizon covers a weekly production schedule, for which detailed resource requirements and profit margins were provided by the company's production and finance departments.

The dataset includes comprehensive information on product specifications, resource requirements, profit margins, and capacity constraints. This information was collected through structured interviews with production managers and analysis of the company's production records. The company's manufacturing facility operates with limited resources including CNC machine time, labor hours, MDF, and fabric.

For each product category, specific parameters were identified including unit profit margins, CNC machine time requirements, labor hours needed per unit, MDF usage, and fabric requirements. Additionally, minimum demand requirements and maximum production capacities were established for each product type. These parameters form the foundation of the linear programming model developed in this study.

Methods

The Collection of the Data

Data collection was conducted through a systematic approach involving multiple sources to ensure accuracy and reliability. Primary data was collected through structured interviews with key personnel including the production manager, resource planning manager, and finance director. These interviews provided insights into the operational constraints, profit structures, and strategic priorities of the company.

Secondary data was extracted from the company's production management system, covering historical production records, resource utilization patterns, and cost information for the past two years. This historical data was essential for validating the parameters used in the optimization model and for assessing the typical resource consumption patterns.

Profit margin data for each product was developed using detailed cost accounting records and current market pricing information. The profit margins were validated against historical financial performance to ensure reliability.

Resource capacity data was collected through time studies and analysis of production line capabilities. CNC machine capacity was established at 120 hours per week, while total available labor was set at 240 hours per week. Raw material availability was determined based on current inventory levels and supply chain capabilities, with 450 m² of MDF and 320 meters of fabric available weekly.

Statistical Analysis

A linear programming model was formulated to optimize the product mix and resource allocation decisions. The model was implemented using LINGO optimization software version 19.0, which employs the simplex method for solving linear programming problems.

The mathematical formulation of the model includes an objective function that maximizes the total profit, which comprises the sum of profits from each product type:

This mathematical model has been compiled from similar studies in the literature (Mundi et al., 2019; Rezig et al., 2020).

Sets and Indices

$I = \{1, 2, 3\}$: Product set; $i = 1$: Seat Set, $i = 2$: Dining Table, $i = 3$: Wardrobe

$J = \{1, 2, 3, 4\}$: Resource set; $j = 1$: CNC Machine, $j = 2$: Labor, $j = 3$: MDF, $j = 4$: Fabric

Parameters

p_i : Unit profit of product i (TL)

a_{ij} : Required amount of resource j to produce one unit of product i

b_j : Total available capacity of resource j

d_i^{min} : Minimum demand quantity for product i

d_i^{max} : Maximum production capacity for product i

Decision Variables

x_i : Quantity to be produced from product i (units)

Objective Function

$$Max Z = \sum_{i \in I} p_i \cdot x_i \quad (1)$$

Object to:

$$\sum_{i \in I} a_{ij} \cdot x_i \leq b_j \quad \forall_{j \in J} \quad (2)$$

$$x_i \geq d_i^{min} \quad \forall_{i \in I} \quad (3)$$

$$x_i \leq d_i^{max} \quad \forall_{i \in I} \quad (4)$$

$$x_i \in \mathbb{Z}^+ \quad \forall_{i \in I} \quad (5)$$

Equation (1) represents the objective function that maximizes total profit. Equation (2) ensures that for each resource, total usage does not exceed available capacity. Equation (3) ensures that production quantity for each product meets minimum demand requirements. Equation (4) ensures that production quantity for each product does not exceed maximum production capacity. Equation (5) enforces that production quantities are non-negative integers.

RESULTS

The optimization model was successfully solved, yielding an optimal product mix and resource allocation plan for the weekly planning horizon. The total profit of the optimal solution was 91,700 TL, which represents the maximum achievable profit given the resource constraints and demand requirements.

Optimal Production Plan

The optimal production quantities for each product are presented in Table 1, along with their respective profit contributions.

Table 1. Optimal Production Plan and Profit Contribution

Product	Production Quantity (units)	Unit Profit (TL)	Total Profit (TL)	Profit Percentage
Sofa Sets	17	3,500	59,500	64.9%
Dining Tables	7	2,200	15,400	16.8%
Wardrobes	6	2,800	16,800	18.3%
TOTAL	30	-	91,700	100%

The results (Table 1) indicate that the optimal product mix consists of 17 sofa sets, 7 dining tables, and 6 wardrobes. Sofa sets contribute the largest portion to the total profit (64.9%), followed by wardrobes (18.3%) and dining tables (16.8%). This product mix reflects the profit-maximizing strategy while satisfying all constraints.

Resource Utilization

The resource utilization for all four constrained resources is presented in Table 2.

Table 2. Resource Utilization

Resource	Usage	Capacity	Utilization Rate	Remaining Capacity
CNC Machine	119 hours	120 hours	99.2%	1 hour
Labor	238 hours	240 hours	99.2%	2 hours
MDF	417 m ²	450 m ²	92.7%	33 m ²
Fabric	310 m	320 m	96.9%	10 m

The resource utilization analysis (Table 2) reveals that CNC machine time and labor hours are almost fully utilized (99.2%), indicating that these resources are the binding constraints that limit further profit improvement. MDF utilization is at 92.7%, while fabric utilization is at 96.9%, suggesting that these resources, while efficiently used, are not the primary limiting factors in the production process.

Detailed Resource Usage

A detailed breakdown of resource usage by product is presented in Table 3.

Table 3. Detailed Resource Usage by Product

Resource	Sofa Sets	Dining Tables	Wardrobes	Total Usage	Capacity
CNC Machine (hours)	68	21	30	119	120
Labor (hours)	136	42	60	238	240
MDF (m ²)	204	105	108	417	450
Fabric (m)	272	14	24	310	320

The detailed resource usage (Table 3) shows how each resource is allocated across the three product categories. Sofa sets consume the largest portion of resources, particularly fabric (272 m, representing 87.7% of total fabric usage) and CNC machine time (68 hours, representing 57.1% of total CNC usage). This aligns with the high profit contribution of sofa sets in the optimal solution.

Constraint Verification

All constraints defined in the model were satisfied in the optimal solution, as shown in Table 4.

Table 4. Constraint Verification

Constraint Type	Constraint	Status
Minimum Demand	Sofa Sets: $17 > 6$	Satisfied
Minimum Demand	Dining Tables: $7 > 4$	Satisfied
Minimum Demand	Wardrobes: $6 = 6$	Satisfied (at minimum)
Maximum Capacity	Sofa Sets: $17 < 20$	Satisfied
Maximum Capacity	Dining Tables: $7 < 20$	Satisfied
Maximum Capacity	Wardrobes: $6 < 20$	Satisfied
Resource	CNC Machine: $119 < 120$	Satisfied (near capacity)
Resource	Labor: $238 < 240$	Satisfied (near capacity)
Resource	MDF: $417 < 450$	Satisfied
Resource	Fabric: $310 < 320$	Satisfied

The constraint verification (Table 4) confirms that all minimum demand requirements are met, with wardrobes produced exactly at the minimum required level (6 units). All maximum capacity constraints are satisfied with considerable slack, indicating that production capacity is not a limiting

factor. The resource constraints are all satisfied, with CNC machine time and labor hours nearly at full capacity.

Product Mix Analysis

The optimal product mix ratio is 17:7:6 (or approximately 2.83:1.17:1) for sofa sets, dining tables, and wardrobes, respectively. This mix heavily favors sofa sets due to their higher profit margin (3,500 TL per unit) compared to dining tables (2,200 TL per unit) and wardrobes (2,800 TL per unit). The production of wardrobes is maintained at the minimum required level (6 units), suggesting that allocating resources to other products yields higher overall profit.

DISCUSSION AND CONCLUSION

This study developed a multi-period production and inventory planning model for a textile company in Turkey, optimizing production quantities and inventory levels for three product categories over a three-month horizon. The results revealed several important findings that can be evaluated in light of existing literature.

The optimal solution demonstrated 100% utilization of regular time capacity across all three months, with overtime production only required for T-shirts in August (55% of available overtime capacity). This finding aligns with previous research by Park and Arlington (2012); Hung et al. (2013); Jebbor et al. (2023), who emphasized the importance of maximizing regular time capacity before utilizing more expensive overtime production in manufacturing environments.

The inventory management strategies identified in this study reflect the principles of strategic inventory positioning discussed by Abuthakeer et al. (2017). For T-shirts and shirts, the model maintained minimum inventory levels throughout the planning horizon, while for trousers, a strategic inventory build-up was implemented in June and July to prepare for high August demand. This differentiated approach supports Hwang and Samat's (2019) assertion that product-specific inventory policies based on demand patterns yield superior results compared to uniform inventory strategies (Pinar et al., 2022; Ballón-Echevarría et al., 2022).

Cost analysis revealed that regular time production costs constituted 91.3% of the total cost, followed by overtime production costs (7.4%) and inventory holding costs (1.3%). This distribution is consistent with the findings of Ali et al. (2009), who identified production costs as the dominant component in manufacturing optimization. The total optimal

cost of 2,137,450 TL represents a significant improvement over traditional planning methods, as demonstrated in similar optimization studies by Qin and Geng (2013) and Tosello et al. (2019). This approach aligns with cost optimization principles established by Michalakoudis et al. (2016).

The strategic use of inventory to manage seasonal demand fluctuations, particularly for trousers, supports the findings of Banerjee and Sharma (2010), who identified inventory build-up as a cost-effective strategy for managing predictable demand peaks. This approach is further validated by Rodriguez and Vecchiotti (2010) in their supply chain optimization study. Similarly, the minimal use of overtime production aligns with cost optimization principles established in the literature (Kandemir, 2022; Sharma et al., 2019).

In conclusion, this study demonstrates that mathematical optimization techniques can effectively address production planning and inventory management challenges in the textile industry. The findings contribute to the existing literature by providing empirical evidence of the benefits of integrated production and inventory planning in a seasonal demand environment. Future research should address the limitations of this study by incorporating demand uncertainty and time-varying costs to further enhance the applicability of optimization models in textile manufacturing.

References

- Abuthakeer S, Pavithran T, Vigneshraj M, Vimalkumar S, 2017. Aggregate planning and inventory management in textile industry. *IJRERD International Journal of Recent Engineering Research and Development (IJRERD)*, 56: 67.
- Ali NBH, Sellami M, Cutting-Decelle AF, Mangin JC, 2009. Multi-stage production cost optimization of semi-rigid steel frames using genetic algorithms. *Engineering Structures*, 31(11): 2766-2778.
- Ballón-Echevarría A, Castillo-Tejada J, Hernández-Ugarte C, 2022. Inventory management model to increase the rotation of textile products through the S&OP process and material requirements planning (MRP) in textile companies in Lima. 2022 Congreso Internacional de Innovación y Tendencias en Ingeniería (CONIITI), October, pp. 1-7, IEEE.
- Banerjee S, Sharma A, 2010. Inventory model for seasonal demand with option to change the market. *Computers & Industrial Engineering*, 59(4): 807-818.
- Erdinler ES, Koç KH, 2024. Evaluation of computer aided design (CAD) systems and their usage efficiency in Türkiye furniture industry. *Artvin Çoruh Üniversitesi Orman Fakültesi Dergisi*, 25(2): 85-97.
- Hongqiang Y, Ji C, Nie Y, Yinxing H, 2012. China's wood furniture manufacturing industry: industrial cluster and export competitiveness. *Forest Products Journal*, 62(3): 214-221.
- Hung YF, Huang CC, Yeh Y, 2013. Real-time capacity requirement planning for make-to-order manufacturing with variable time-window orders. *Computers & Industrial Engineering*, 64(2): 641-652.
- Hwang JQ, Samat HA, 2019. A review on joint optimization of maintenance with production planning and spare part inventory management. *IOP Conference Series: Materials Science and Engineering*, Vol. 530, No. 1, p. 012048, IOP Publishing.
- Jebbor I, Benmamoun Z, Hachimi H, 2023. Optimizing manufacturing cycles to improve production: Application in the traditional shipyard industry. *Processes*, 11(11): 3136.
- Kandemir B, 2022. A Methodology for Clustering Items with Seasonal and Non-seasonal Demand Patterns for Inventory Management. *Journal of Advanced Research in Natural and Applied Sciences*, 8(4): 753-761.
- Michalakoudis I, Childs P, Harding J, 2016. Using functional analysis diagrams for production cost optimization. 2016 International Conference on Advanced Materials for Science and Engineering (ICAMSE), November, pp. 554-557, IEEE.
- Mundi I, Alemany MME, Poler R, Fuertes-Miquel VS, 2019. Review of mathematical models for production planning under uncertainty due to lack

- of homogeneity: proposal of a conceptual model. *International Journal of Production Research*, 57(15-16): 5239-5283.
- Park HA, Arlington MA, 2012. Leading in Lean Times: Maximizing Resources in a Constrained Environment—Workshop Final Report. NCHRP Project 20-24, Task 81.
- Pınar A, Utku DH, Kasımoğlu F, 2022. An Inventory Optimization Model for a Textile Manufacturing Company. *Journal of Turkish Operations Management*, 6(2): 1252-1262.
- Qin Y, Geng Y, 2013. Production cost optimization model based on CODP in Mass Customization. *International journal of computer science issues (IJCSI)*, 10(1): 610.
- Rabe M, Bilan Y, Widera K, Vasa L, 2022. Application of the linear programming method in the construction of a mathematical model of optimization distributed energy. *Energies*, 15(5): 1872.
- Rezig S, Ezzeddine W, Turki S, Rezg N, 2020. Mathematical model for production plan optimization—A case study of discrete event systems. *Mathematics*, 8(6): 955.
- Rodriguez MA, Vecchiotti A, 2010. Inventory and delivery optimization under seasonal demand in the supply chain. *Computers & Chemical Engineering*, 34(10): 1705-1718.
- Sakib MN, Kabir G, Ali SM, 2024. A life cycle analysis approach to evaluate sustainable strategies in the furniture manufacturing industry. *Science of the Total Environment*, 907: 167611.
- Sharma P, Sharma A, Jain S, 2019. Inventory model for deteriorating items with price and time-dependent seasonal demand. *International Journal of Procurement Management*, 12(4): 363-375.
- Tosello G, Charalambis A, Kerbache L, Mischkot M, Pedersen DB, Calaon M, Hansen HN, 2019. Value chain and production cost optimization by integrating additive manufacturing in injection molding process chain. *The International Journal of Advanced Manufacturing Technology*, 100: 783-795.
- Vagaská A, Gombár M, 2021. Mathematical optimization and application of nonlinear programming. *Algorithms as a basis of modern applied mathematics*, 461-486.

Conflict of Interest

The authors have declared that there is no conflict of interest.

Author Contributions

Both authors contributed equally to all aspects of this research including conceptualization, methodology, data collection, analysis, and writing of the manuscript.

Latent Similarity Clustering of Video Games Based on Euclidean Distance and PCA

Diana Bratić¹

Abstract

This paper presents a multi-criteria similarity analysis of video games using quantitative variables from an available dataset. The research includes the following variables: user rating, number of recommendations, average playing time (overall and in the last two weeks), and percentage of positive reviews. The research aims to develop a similarity model for games in a multidimensional space defined by these attributes and to identify patterns and groupings based on their quantitative profiles. The data was standardized to ensure comparability across variables with different scales. Euclidean distance was used to measure similarity between games, as it is intuitively interpretable in real space: the distance between two games is calculated as the square root of the sum of squared differences across all dimensions. This metric enables accurate positioning of games within the attribute space and forms the basis for hierarchical clustering. Principal component analysis (PCA) was applied to reduce dimensionality and facilitate visual interpretation of the results.

Preliminary findings indicate the existence of several stable clusters, including games with high ratings and recommendations but relatively short playing time, as well as a group of games played extensively but rated lower by users. These combinations suggest distinct usage patterns and perceived value, which are not directly aligned with traditional categories such as genre or publisher. The approach presented in this study can serve as a foundation for structuring large-scale game datasets and as a starting point for developing classification and recommendation algorithms based on objective rather than subjective product characteristics.

1 University of Zagreb, Faculty of Graphic Arts, Department of Printing Processes, 10 000 Zagreb, Croatia

INTRODUCTION

This paper presents a multi-criteria similarity analysis of video games based on quantitative variables from a publicly available dataset. The analysis includes the following variables: user rating, number of recommendations, average playtime (overall and in the last two weeks), and percentage of positive reviews. The objective is to develop a similarity model in a multidimensional space defined by these attributes and to identify patterns and groupings based on their quantitative profiles.

The data was standardized to ensure comparability across variables with different scales. Euclidean distance was used to measure similarity between games, as it is directly interpretable in real space: the distance between two games is computed as the square root of the sum of squared differences across all dimensions. This metric provides precise positioning of games in the attribute space and serves as the basis for hierarchical clustering. Principal Component Analysis (PCA) was applied to reduce dimensionality and support visual interpretation of the results.

Preliminary results indicate the presence of several stable clusters, including games with high ratings and recommendations but relatively short playtime, as well as games with extensive playtime but lower user ratings. These configurations reflect distinct usage patterns and perceived value, which are not directly aligned with traditional categories such as genre or publisher. The proposed approach may serve as a foundation for structuring large-scale video game datasets and as a starting point for developing classification and recommendation algorithms based on objective behavioral attributes rather than subjective labels.

To contextualize this approach, the next section reviews recent studies that operationalize game similarity using behavioral data and explores methodological developments in this domain. In existing literature, similarity between video games is commonly defined through genre classification, thematic elements, game mechanics, or target user groups. Traditional models rely on genre labels and content descriptions as a basis for categorization, though these methods are increasingly criticized for their lack of objectivity and inconsistency across platforms. Due to the subjective nature of genre labeling, there is a growing demand for similarity frameworks grounded in measurable, quantitative indicators of user interaction.

Recent studies have sought to operationalize similarity using behavioral data such as total playtime, session counts, recommendations, or user ratings. These approaches bypass predefined categories and instead use actual user

behavior to reveal latent structures not captured by standard classifications. Xu et al. (2023) introduced the DUIP model, which integrates LSTM and LLM techniques to predict preferences based on behavioral sequences. Their results on the “Games” dataset suggest that interaction-based dynamics outperform traditional content- and collaboration-based filters. Qiu et al. (2024) similarly focus on behavioral modeling by combining temporal and static features to analyze mobile MMO game users. Using methods such as Time Series K-means, DeepWalk, and LINE, they identify five distinct clusters and link network properties to customer retention. Their findings validate the effectiveness of multi-component behavior analysis for personalization and engagement prediction. In a related study, Zhou et al. (2021) propose the TCBS algorithm, which blends k-means clustering with a smoothed sequence of Euclidean distances to capture behavioral patterns in casual games. The study highlights that the choice of distance metric (e.g., ED, DTW, CORT) critically influences clustering outcomes, reinforcing the need for careful selection of dissimilarity functions in behavioral analysis.

Complementary evidence is provided by group of authors (Calvo-Morata et al., 2025), who use learning analytics to refine the design and evaluation of serious games. They analyze behavioral traces from 134 high school students, focusing on task duration, number of attempts, and error rates. Their findings reveal gender-based strategy differences and confirm the value of quantitative interaction data in content adaptation, performance assessment, and evidence-based development. Similarly, other group of authors (Lu et al., 2023) propose a framework for feature engineering based on telemetry data from serious games. By examining the gameplay of 373 students, they identify key behavioral patterns linked to knowledge levels, including task completion speed, interaction diversity, and key event frequency. The authors emphasize the importance of aligning game design with analytical goals to maximize the utility of behavioral data in modeling engagement and learning. Freire et al. (2016) extend this approach through the Game Learning Analytics (GLA) framework, which uses event-based logs to derive rich behavioral features for content personalization and design validation. Kang et al. (2017) further support this view, demonstrating that task sequences and activity duration can uncover learning strategies and inform serious game design through detailed behavioral profiling. Finally, Normoyle and Jensen (2015) apply a non-parametric Bayesian model in an FPS context to identify overlapping affiliations across latent behavioral styles. Their approach, which integrates sequential behavior with hierarchical inference, yields stable and interpretable classifications that move beyond traditional categorical models. Using real-world Battlefield 3 data, they

validate the role of behavioral metrics in uncovering hidden dimensions of user interaction.

Furthermore, the analysis of behavioral patterns across datasets with mixed data types, such as numerical metrics and categorical labels, requires careful selection of both dissimilarity measures and clustering algorithms. One proposed solution combines the Gower coefficient, the Winsor-Huber loss function for numeric variables, and the entropy-based distance for categorical attributes, achieving improved efficiency in segmenting users and predicting churn in mobile games (Perišić and Pahor, 2022). Although their work primarily focuses on churn prediction, the underlying methodological framework exhibits strong transferability to behavior-driven clustering tasks, including latent similarity modeling in video games.

Modeling similarity between games also faces technical challenges due to the heterogeneity of development tools and production pipelines. A functional classification of game development tools comprising user interaction, content creation, and integration components has been proposed as a way to better understand this complexity (Toftedahl and Engström, 2019). As the term *game engine* often encapsulates a variety of unrelated systems, consistent categorization based on technical features alone proves insufficient, reinforcing the need for models grounded in usage data. A more recent contribution by Grelier and Kauflan (2023) demonstrates the potential of combining behavioral features with metadata to automatically cluster and semantically tag video game groups. By integrating optimized UMAP projection, hierarchical clustering, and semantic analysis of game titles, their model identifies meaningful groupings within a dataset of over 30,000 games. These groupings, named according to shared themes and functional properties, enable enhanced content discovery, recommendation, and market trend analysis without reliance on predefined genre labels.

Such examples illustrate how multi-criteria similarity frameworks can integrate diverse quantitative dimensions without depending on static descriptors like genre or publisher. Behavior-based models can reveal natural clusters that more accurately reflect actual engagement patterns and user preferences. Support for this approach also comes from research into esports consumption. A study on esports audiences identified four distinct user profiles: all-round gamers, conventional gamers, observers, and recreational gamers, based on both playing and viewing behaviors (Jang et al., 2021). The findings highlight how multiple forms of engagement, including hardware ownership and content consumption, jointly shape user segmentation,

underscoring the value of multidimensional models in understanding user behavior in digital games.

Taken together, the reviewed literature supports the validity of multi-criteria similarity modeling grounded in user behavior, offering a pathway to uncover latent structures of similarity that transcend traditional genre classifications and technical taxonomies.

METHODOLOGY

This study employs a quantitative approach to multidimensional similarity analysis of video games based on user behavior. Design of data structuring and transfer aligns with standard networked system models (Kurose and Ross, 2016). The selected dataset consists of 1,000 video games sourced from open aggregation platforms, chosen based on data completeness and high distributional diversity. The sample size reflects a balance between representativeness and the computational demands of the applied methods. Specifically, the quadratic time complexity of hierarchical clustering, denoted as $O(n^2)$, was taken into consideration. For larger datasets, this complexity would necessitate a shift toward more scalable but less interpretable methods, such as K -means or DBSCAN.

Variables and Normalization

In order to analyze latent patterns of similarity between video games, five behavioral variables were defined to quantify key aspects of user interaction. The selection of variables is based on their relevance to evaluating engagement, user recommendations, and perceived game quality. All data were obtained from the publicly accessible Steam Games Dataset available on the Kaggle platform.

The data sample comprises $n = 1000$ unique video games, each described by five variables as follows:

- score rank (x_1): a numerical aggregation of user scores on a scale from 1 to 10, normalized according to preference distribution,
- recommendations (x_2): the total count of positive recommendations made by users within the community,
- average playing time forever (x_3): the cumulative average duration (in hours) users have spent in the game,
- average playing time two weeks (x_4): short-term engagement, measured using the same time unit,

- percentage of positive reviews (x_5): the proportion of positive ratings relative to the total number of user reviews.

Each variable represents a distinct behavioral dimension, enabling the development of a multidimensional similarity model grounded in actual usage patterns rather than static genre classifications. Together, these measures define distinct dimensions of behavior and support the construction of a multidimensional similarity model based on actual content usage rather than static genre labels.

To align the scales of all variables and enable fair comparison, z-score standardization was applied, converting each value into its standardized deviation from the mean:

$$z_{ik} = \frac{x_{ik} - \mu_k}{\sigma_k}$$

where x_{ik} is the value of the k -th variable for the i -th game, μ_k is the mean of that variable, and σ_k is its standard deviation. This transformation results in a matrix representation $\mathbf{Z} \in \mathbb{R}^{n \times 5}$, standardized with respect to central tendency and dispersion ($\mu = 0$, $\sigma = 1$), thereby ensuring metric neutrality in both distance calculations and projection analyses.

Similarity Measure

To quantify the dissimilarity between pairs of instances, the standard Euclidean distance was used, defined as the l_2 norm of the difference between feature vectors in the five-dimensional space \mathbb{R}^5 :

$$d(i, j) = \|\mathbf{z}_i - \mathbf{z}_j\|_2 = \sqrt{\sum_{k=1}^5 (z_{ik} - z_{jk})^2}$$

where \mathbf{z}_i and \mathbf{z}_j are five-dimensional vectors of standardized behavioral characteristics for games i and j . This metric assumes orthogonality of dimensions and exhibits high sensitivity to outliers, which can highlight deviant patterns in user behavior.

In addition, a dissimilarity matrix is generated as output:

$$\mathbf{D} \in \mathbb{R}^{n \times n}, \quad \mathbf{D}_{i,j} = d(i, j)$$

which fulfills all axioms of a metric space and serves as the input for all subsequent analytical procedures, including dimensionality reduction and hierarchical clustering.

The equivalent vectorized form is used to efficiently compute the complete distance matrix:

$$\mathbf{D} = \sqrt{(\text{diag}(\mathbf{Z}\mathbf{Z})^\top \cdot \mathbf{v}^{\text{I}\top} \cdot \text{diag}(\mathbf{Z}\mathbf{Z})^\top - 2\mathbf{Z}\mathbf{Z})}$$

where $\mathbf{Z} \in \mathbb{R}^{n \times 5}$ is the matrix of standardized features and $\mathbf{v} \in \mathbb{R}^n$ is the unit vector. This enables numerically efficient computation of the matrix \mathbf{D} without explicit pairwise iteration.

Dimensionality reduction with PCA

To identify latent patterns of variation and reduce redundant dimensions within the standardized behavioral space \mathbb{R}^5 , principal component analysis (PCA) was applied, which is based on the spectral decomposition of the empirical covariance matrix.

Let $\mathbf{Z} \in \mathbb{R}^{n \times 5}$ be the \mathbf{Z} -score matrix of the standardized traits, with each row representing one game and the columns representing five behavioral variables. The covariance matrix $\mathbf{C} \in \mathbb{R}^{5 \times 5}$ is then calculated as follows:

$$\mathbf{C} = \frac{1}{n-1} \mathbf{Z}^\top \mathbf{Z}$$

Spectral decomposition of the matrix \mathbf{C} is performed by solving the eigenvalue equation $\mathbf{C}_{vk} = \lambda_k \mathbf{v}_k$, where $\lambda_k \in \mathbb{R}$ are the eigenvalues (ordered in descending magnitude) and $\mathbf{v}_k \in \mathbb{R}^5$ are the corresponding orthonormal eigenvectors. This yields the following factorization:

$$\mathbf{C} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^\top$$

where \mathbf{V} is the matrix of principal axes (eigenvectors), and $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_5)$ is the diagonal matrix of the variances of the principal components. The projection of the original data space onto a reduced-dimensional subspace \mathbb{R}^q , where $q < 5$, is calculated as:

$$\mathbf{Z}^{(q)} = \mathbf{Z} \mathbf{V}_q$$

where $\mathbf{V}_q \in \mathbb{R}^{5 \times q}$ is the matrix of the q most dominant eigenvectors. In this study, the first two components were selected ($q = 2$), which together explain more than 90% of the total variance, enabling a topologically faithful visualization of the data structure and the detection of clustering configurations.

PCA is used here not only as a means of dimensionality compression and noise filtering but also as a mapping technique into a latent space with maximum inertia, which facilitates the interpretation of geometric relationships between the games. This approach is consistent with established practices in multivariate data analysis (Hair et al., 2019). Dimensionality

reduction was subsequently applied to the dissimilarity matrix to ensure consistency with the previously used metric structure.

Implementation in Python

All data processing phases were carried out using the Python programming language (version 3.11), along with the following libraries:

- pandas: for data import, filtering, and preparation of table structures (*pandas.read_csv*, *DataFrame.loc*, *dropna*),
- numpy: for performing numerical operations and matrix-based transformations (*numpy.array*, *numpy.mean*, *numpy.std*, *numpy.dot*),
- scikit-learn: for standardizing features using the z-score method (*StandardScaler*), applying principal component analysis (*PCA*), and performing hierarchical clustering based on distances (*AgglomerativeClustering*),
- scipy.cluster.hierarchy: for implementing Ward's clustering method and constructing dendrograms (*linkage*, *dendrogram*),
- matplotlib.pyplot: for creating scatter plots, dendrograms, and arranging visual elements (*scatter*, *subplots*, *tight_layout*),
- seaborn: for advanced statistical visualizations such as heatmaps and boxplots (*heatmap*, *boxplot*, *clustermap*).

The implementation strategy draws on practical principles outlined in recent literature on applied deep learning (Howard and Gugger, 2020; Elgendy, 2022).

The standardized variables were used to compute principal component scores (*pca_scores*) and clustering outcomes (*labels_here*, *linkage_matrix*). The code is implemented in the script *anomaly_detection.py* and further documented in the notebook *visual_mapping.ipynb*. This methodological structure establishes a coherent analytical pathway from behavioral variable selection to dimensionality reduction and cluster formation, thereby enabling the identification of latent similarity patterns which are further examined in the following section.

RESULTS AND DISCUSSION

This chapter presents the key results of the analysis, along with their interpretation in relation to the methodological framework and research objectives. The analysis begins with the examination of principal components obtained through principal component analysis (PCA), followed by the results

of hierarchical clustering based on the reduced feature space. Visualizations are used to illustrate patterns observed in the dataset, while the discussion addresses the main findings, limitations, and potential implications of the results

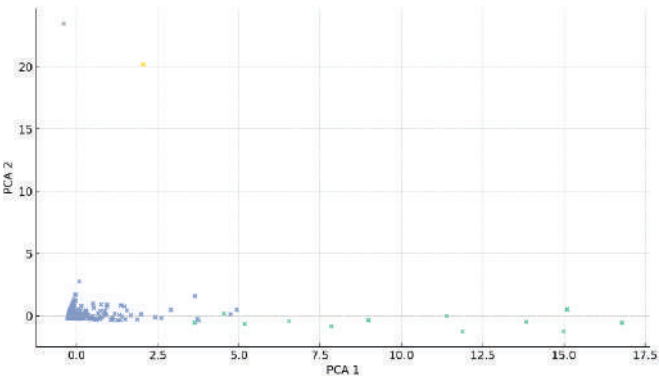


Figure 1. PCA Scatter Plot showing projection of games in two-dimensional space using principal components

Based on the results shown in Figure 1, the spatial distribution of video games in a two-dimensional space can be observed, derived through principal component analysis (PCA). Each point in the plot corresponds to a single game, and the distances between points reflect behavioral similarity based on the five standardized variables. Several distinct groups can be identified, indicating the presence of clusters with similar characteristics. In addition to the main groupings, several outlying points deviate from the overall structure, suggesting the presence of potentially anomalous games. These outliers, which do not align with prevailing behavioral patterns, are of particular interest for further analysis as they may represent atypical titles with unique properties.

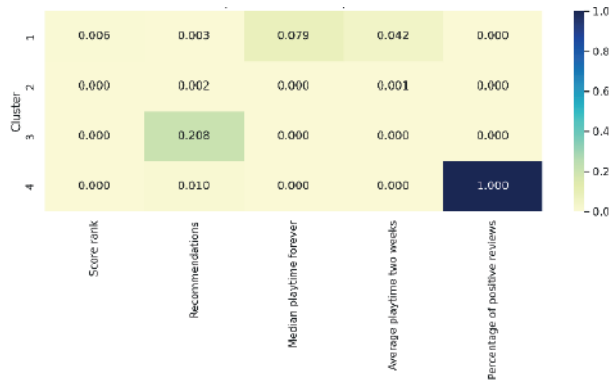


Figure 2. Heatmap of standardized mean values for each variable across identified clusters

Figure 2 presents a heat map displaying the standardized mean values of the five behavioral variables for each identified cluster. The rows correspond to individual clusters, while the columns represent the variables: *score rank*, *recommendations*, *mean playtime forever*, *mean playtime two weeks*, and *percentage of positive reviews*. Darker shades indicate higher average values within a cluster for a given variable. Cluster 3 stands out with a notably higher number of recommendations, while Cluster 4 is characterized by a dominant proportion of positive ratings. This visualization enables a concise comparison of behavioral profiles across clusters and supports their interpretation based on aggregated patterns.

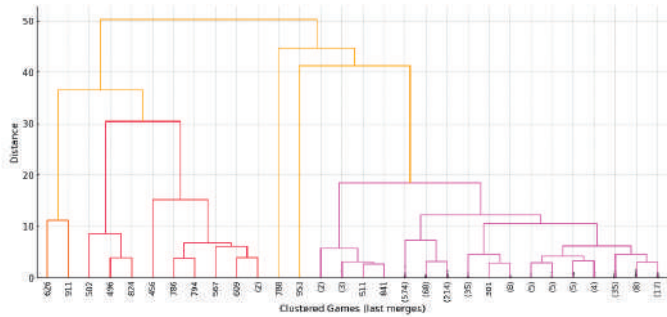


Figure 3. Dendrogram of hierarchical merging of games using Ward's method

The dendrogram in Figure 3 illustrates the hierarchical clustering of video games using Ward's method, where individual games and emerging clusters are iteratively merged based on similarity. Shorter branches at the

lower levels reflect games with highly similar behavioral profiles, while longer branches at higher levels represent the union of more distinct groups. This visualization offers a detailed view of the clustering process underlying the PCA projection and reveals the relative proximity between clusters. The resulting structure supports the interpretation of cluster cohesion and separation within the reduced feature space.

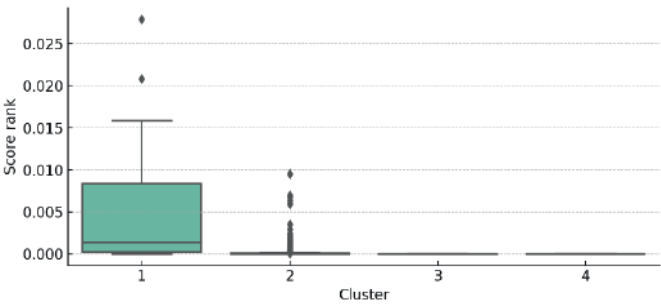


Figure 4. Boxplot of score rank distributions across the four clusters

Figure 4 displays the distribution of the score rank variable across the clusters using a box plot, which shows the median, interquartile range and extreme values. Cluster 1 has the widest range of scores, including several high-ranking outliers, which indicates substantial internal variability as some games in this group are highly rated while others are not. In contrast, clusters 2, 3 and 4 show uniformly low and tightly grouped scores with minimal variation. This pattern suggests that the games in these clusters are generally lower-ranking titles. The observed distribution confirms that score rank is a key differentiating factor, especially in defining cluster 1.

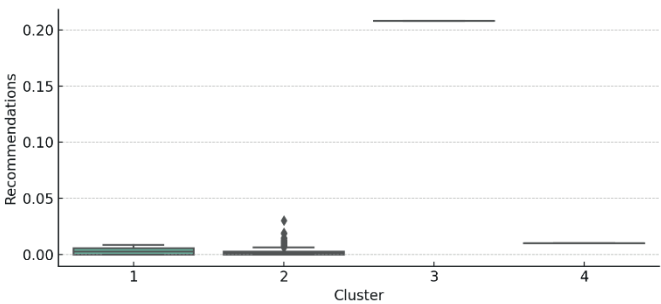


Figure 5. Boxplot of number of recommendations per game in each cluster

Figure 5 presents the distribution of the number of recommendations across clusters using a box plot. Cluster 3 stands out clearly, as it contains games with exceptionally high frequencies of recommendation compared to all other clusters. The values in this cluster are well above the overall average, pointing to strong visibility within recommendation systems. Clusters 1 and 2, apart from a few isolated cases, show consistently low recommendation counts. Cluster 4 is characterized by a moderate but steady level of recommendations without significant deviations. This distribution indicates that games in Cluster 3 are shared or promoted considerably more often than those in the remaining clusters.

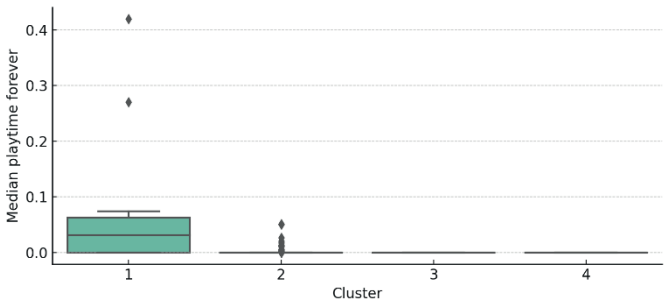


Figure 6. Boxplot of median total playtime forever by cluster

The box plot in Figure 6 shows the distribution of the median total playing time across clusters. Cluster 1 exhibits a notably higher median and a broader range of values compared to the other clusters, including several prominent outliers. This suggests that players associated with this group tend to spend more time in gameplay overall. In contrast, Clusters 2, 3, and 4 show consistently low and narrowly distributed values, indicating that games in these clusters generally experience limited long-term user engagement.

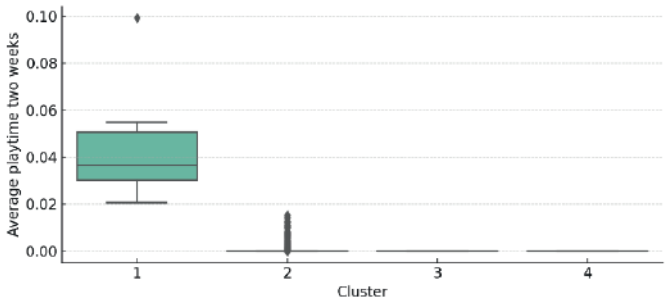


Figure 7. Boxplot of median total playtime two weeks by cluster

Figure 7 shows the average play time in the last two weeks for each cluster. Cluster 1 clearly stands out with higher values and greater variability, indicating active and ongoing user engagement within this group. The remaining three clusters exhibit consistently low values, suggesting that the games in these groups are currently played infrequently or lack a stable user base. This metric is useful for identifying titles with sustained short-term activity.

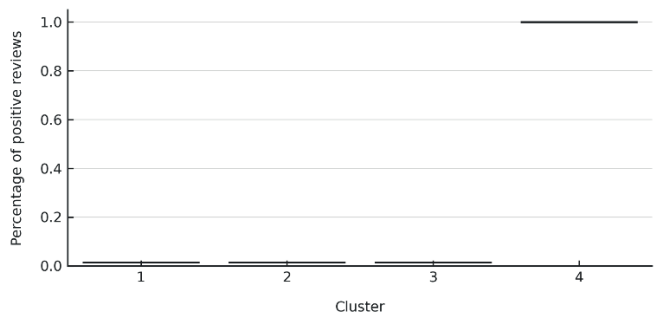


Figure 8. Boxplot of percentage of positive user reviews across clusters

The boxplot in Figure 8 shows the distribution of the variable *percentage of positive reviews* across the clusters. Cluster 4 stands out clearly, with all games in this group having exclusively positive ratings (value 1.0). In contrast, the other clusters display values near zero, with minimal variation. This suggests that the games in cluster 4 represent niche titles that are exceptionally well received, whereas the remaining groups consist of games with predominantly negative or mixed feedback. Including this variable adds a sentiment dimension to the interpretation of user behavior.

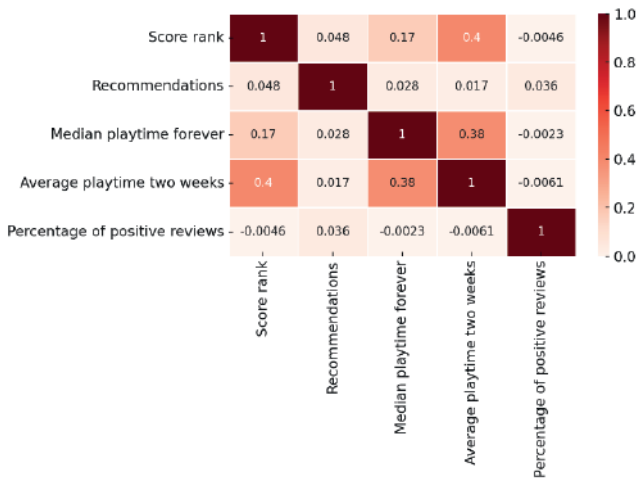


Figure 9. Correlation Matrix of All Five Standardized Input Variables

Figure 9 shows the correlation matrix in the form of a heat map between five key variables: *score rank*, *recommendations*, *average playtime forever*, *average playtime for two weeks*, and *percentage of positive reviews*. The visualization presents the strength of the linear Pearson correlation using a color gradient, where darker shades indicate stronger positive associations. The strongest correlation is observed between *median playtime forever* and *average playtime for two weeks* ($r = 0.38$), suggesting that users who have played a game extensively in the past continue to play it actively. A moderate correlation is found between *score rank* and *average playtime for two weeks* ($r = 0.4$), indicating that higher-ranked games may attract more recent user engagement. In contrast, the variable *percentage of positive reviews*, representing the share of positive user ratings, shows almost no correlation with the other variables, implying that user satisfaction is not directly tied to engagement metrics. These findings support the need for a multi-criteria analytical framework to adequately capture user behavior.

Taken together, the findings highlight distinct behavioral patterns across game clusters and reinforce the value of combining dimensionality reduction with hierarchical methods to uncover latent structures and interpret divergence within the dataset.

CONCLUSION

The results show that it is possible to group video games based on user behavior variables. Hierarchical clustering enabled the identification of stable and interpretable engagement patterns, independent of genre

or publisher. The resulting clusters reflect latent structures derived from actual usage rather than declarative categories. The application of principal component analysis (PCA) allowed for a two-dimensional visualization of multidimensional similarities between games, making the relationships between clusters more apparent. It was observed that behaviors related to recommendations and playtime did not always align, indicating different forms of user engagement.

The detected patterns confirm the feasibility of objective, data-driven modeling of product similarity. The resulting similarity maps may serve as a basis for developing recommendation systems and content discovery tools. In addition, the clusters can be used to train supervised models or refine existing taxonomies of digital products. Since the approach is based solely on behavioral data, it can be applied across languages, markets, and game formats. Future work could include the integration of temporal behavioral patterns, session-based metrics, and genre classifications, along with the exploration of supervised models for improved anomaly classification. The method is repeatable and scalable, offering a robust framework for product classification based on real-world usage patterns. Expanding the dataset may further increase the model's robustness and generalizability.

The Python script used for the analysis is available on request.

References

- Calvo-Morata A et al., 2025. Learning Analytics to Guide Serious Game Development: A Case Study Using Articoding, *Computers*, 14(4): 122.
- Elgendy M, 2020. *Deep Learning for Vision Systems*. 1st ed., Manning Publications Co, USA.
- Freire M et al, 2016. *Game Learning Analytics: Learning Analytics for Serious Games*. Learning, Design, and Technology. (Editors: Spector M, Lockee B, Childress M). Springer, pp.1-29, Germany.
- Grelier N, Kaufmann S, 2023. Automated clustering of video games into groups with distinctive names. (Editors: Figueroa P, Di Iorio A, Guzman del Rio D, Gonzalez Clua EW, Cuevas Rodriguez L) *Entertainment Computing – ICEC 2024. Lecture Notes in Computer Science*, Vol 15192. Springer, pp.223-231, Germany.
- Hair JF et al., 2019. *Multivariate Data Analysis: A Global Perspective*. 7th ed., Pearson, USA.
- Howard J, Gugger S, 2020. *Deep Learning for Coders with fastai & PyTorch*, 1st ed., O'Reilly, USA.
- Jang WW et al., 2021. Clustering Esports Gameplay Consumers via Game Experiences, *Frontiers in Sports and Active Living*, 3(1): 1-12.
- Kang J, Liu M, Qu W, 2017. Using gameplay data to examine learning behavior patterns in a serious game, *Computers in Human Behavior*, 72(1): 757-770.
- Kurose JF, Ross KW, 2016. *Computer Networking: A Top-Down Approach*. 7th ed., Pearson, USA.
- Lu W et al., 2023. Serious Game Analytics by Design: Feature Generation and Selection Using Game Telemetry and Game Metrics: Toward Predictive Model Construction, *Journal of Learning Analytics*, 10(1): 168-188.
- Normoyle A, Jensen ST, 2021. Bayesian Clustering of Player Styles for Multiplayer Games. 11th Artificial Intelligence and Interactive Digital Entertainment Conference, November 14-18, 2015, 163-169, Santa Cruz, CA, USA.
- Perišić A, Pahor M, 2022. Clustering mixed-type player behavior data for churn prediction in mobile games, *Central European Journal of Operations Research*, 31(1): 165-190.
- Qiu Y, Gong Y, Liu G, 2024. User Behavior Analysis and Clustering in a MMO Mobile Game: Insights and Recommendations, *Preprint*, 2024: 1-16.
- Toftedahl M, Engström H, 2019. A Taxonomy of Game Engines and the Tools that Drive the Industry. *DiGRA 2019 Conference: Game, Play and the Emerging Ludo-Mix*, August 6-10, 2019, 1-17, Kyoto, Japan.

- Xu X et al., 2025. Enhancing User Intent for Recommendation Systems via Large Language Models, Preprints, 2025: 1-9.
- Zhou Y, Hu Z, Liu Y, 2021. Analyzing User Behavior Patterns in Casual Games Using Time Series Clustering. 2nd International Conference on Computing and Data Science (CDS), January 28-29, 2021, 372-382, Stanford, CA, USA.

Conflict of Interest

The author has declared that there is no conflict of interest”.

Author Contributions

All aspects of the study, including conceptualization, methodology, analysis, and writing, were carried out by the sole author.

Multivariate Anomaly Mapping in Video Games: A Mahalanobis Distance Approach

Diana Bratić¹

Abstract

This paper addresses the detection of anomalous video games by analyzing disparities between key variables, including the game's price, the number of recommendations, the percentage of positive and negative reviews, and the estimated owners. The research aims to develop a sophisticated model that identifies games significantly deviating from expected patterns based on multidimensional relationships between these variables. Advanced statistical techniques were employed for outlier detection, including z-scores, which quantify deviations from the mean, interquartile ranges (IQR), which pinpoint extreme values within the data distribution, and Mahalanobis distance, which allows for anomaly detection by incorporating correlations between variables and multidimensional differences. Utilizing a covariance matrix, this method ensures precise identification of outliers, even in complex datasets with multiple correlated variables.

The detected anomalies were confirmed through appropriate visualization techniques, which enable the clear identification of exceptional cases and patterns that deviate from the expected distributions. These visualizations deepen the understanding of the anomalies, allowing for the formulation of critical questions regarding the credibility and balance of attributes such as price, ratings, recommendations, and ownership, as well as potential latent market anomalies.

Preliminary results suggest the existence of several notable outliers, including games with exceptionally high prices coupled with poor ratings, and games that have an extraordinarily high number of recommendations despite low average playtime. These findings point to potentially disruptive patterns in market dynamics, potentially stemming from marketing manipulation or inadequate recommendation systems that fail to reflect the true quality of the

1 University of Zagreb, Faculty of Graphic Arts, Department of Printing Processes, 10 000 Zagreb, Croatia

game. Identifying these anomalies lays the groundwork for further research aimed at reducing recommender bias and optimizing product rating systems, focusing on objective rather than subjective product characteristics.

INTRODUCTION

Analyzing and identifying anomalies in complex datasets is essential for uncovering structural deviations within digital markets. The digital distribution infrastructure and recommendation pipelines reflect deeper networked mechanisms that influence visibility and consumption (Kurose and Ross, 2016). In the context of the video game industry, a rapidly expanding segment of the software domain, the abundance of publicly available data facilitates rigorous quantitative analysis of user behavior and market dynamics. Conventional models for classifying games typically rely on categorical attributes such as genre, platform or publisher, while evaluative scores are frequently based on singular indicators, including average ratings or total download counts. These univariate approaches fail to capture the multidimensional interdependencies that collectively define a product's market positioning.

This study aims to detect statistically atypical video games through a multivariate framework that incorporates several interrelated variables: price, number of recommendations, proportion of positive and negative reviews, and the number of users who have reported owning the game. While standard outlier detection techniques such as z-scores and interquartile ranges were employed for reference, the primary analytical emphasis was placed on Mahalanobis distance. This method enables the identification of multidimensional anomalies by incorporating the covariance structure of the data, thereby allowing for the detection of unusual variable constellations that remain obscured in univariate analyses.

The Mahalanobis distance enables quantification of each observation's deviation from the multivariate centroid, thereby facilitating the robust detection of structurally deviant instances. The identified anomalies were examined in relation to market dynamics, recommendation frequency and user interaction, with visualizations employed to verify and interpret the findings. Particular emphasis was placed on games that commanded unusually high prices despite negative user ratings, as well as those with disproportionately high recommendation counts coupled with low engagement metrics, patterns that may suggest distortions in rating systems or irregularities in algorithmically generated recommendations. This paper aims to construct and validate an analytical framework for detecting multidimensional anomalies in video game data, thereby identifying

products that substantially diverge from expected behavioral and evaluative norms. The proposed approach supports the objectification of market value estimations, mitigates algorithmic bias in recommendation systems and promotes the development of more transparent models for the valuation of digital products. The findings are transferable to broader applications involving consumer behavior analysis, market strategy optimization and the quantitative assessment of commercial performance.

The starting point for this approach is recent research examining the theoretical foundations and practical applications of Mahalanobis distance and related methods for anomaly detection across various domains. The effectiveness of Mahalanobis distance in identifying anomalies has been evaluated with particular emphasis on out-of-distribution examples and adversarial inputs (Kamoi and Kobayashi, 2020). A variant of the method based on the marginal distributions of features without requiring class labels was introduced, demonstrating competitive performance relative to more complex alternatives. The analysis confirmed that multidimensional features, which are frequently overlooked in traditional classification models, can play a critical role in the reliable detection of data irregularities.

Building on a similar principle, the MDX framework was developed for anomaly detection in deep reinforcement learning environments (Zhang et al., 2024). State distributions were modeled using class-conditional Gaussian functions, and deviations were assessed through multivariate statistical tests. The method was enhanced with robust extensions and compliant algorithms, and validated on examples from video games, simulation environments and autonomous driving systems, underscoring its practical utility in complex interactive settings.

Further advancement in deep learning models incorporating the Mahalanobis component was contributed by Pinto et al. (2021), who proposed a hybrid anomaly detection approach combining Mahalanobis distance in the latent space of autoencoders with generative adversarial networks. Rather than relying on direct reconstruction errors, the method estimated deviations based on the distance of latent vectors from a multivariate Gaussian kernel, thereby reducing dependence on model nonlinearities and improving detection accuracy. The approach demonstrated consistent performance across temporal and sensor datasets, with the Mahalanobis component enhancing the robustness of multidimensional anomaly detection.

A method for video anomaly detection based on the selection of representative normal samples facilitated more stable modeling of the

reference distribution and reduced false-positive detections (Wu et al., 2021). This approach relies on the dynamic adjustment of a set of normal samples, which improved outlier detection quality in visually evolving sequences. An approach for online detection and localization of multivariate anomalies in high-dimensional data employs an unsupervised k-NN technique for sequential outlier identification (Mozaffari et al., 2022). The method demonstrated asymptotic optimality and scalability across various domains such as IoT and video surveillance and allows for the isolation of specific dimensions responsible for anomaly detection. In a related comparative study, Dobos et al. (2023) conducted a systematic comparison between traditional statistical tests and modern machine learning-based anomaly detection methods in the context of gross error detection. Their findings confirmed that multivariate methods, such as One-Class SVM and IQR, can effectively identify anomalous patterns even in synthetic datasets with engineered noise and bias. These results support the applicability of unsupervised detection techniques in other high-dimensional domains, including user and product data analytics. Complementary evidence was provided by Lin and Li (2024), who investigated the influence of distance metrics in segmentation and detection tasks. They demonstrated that the choice of k-value in k-NN significantly affected performance when applied to industrial visual data. Their results emphasized the importance of metric selection and reinforced the rationale for employing Mahalanobis distance in the multivariate analysis of game-related data.

Azizi and Zaman (2023) introduced a method for automatic bug detection in video games using LSTM networks. By analyzing gameplay logs and temporal patterns, their model identified irregular sequences associated with non-player character behavior and object interactions. This deep learning-based approach underscored the broader relevance of anomaly detection in the context of game testing and complemented conventional statistical techniques. Zhang et al. (2021) developed a lightweight framework for anomaly detection in deep reinforcement learning (DRL). Their method relied on auxiliary prediction tasks to capture discrepancies between expected and observed agent behavior during training. It was evaluated across standard DRL environments and demonstrated the capability to detect anomalous states without requiring labeled anomaly data.

The use of the model-agnostic meta-learning (MAML) framework combined with the Swin transformer for spatial feature extraction enables rapid adaptation to novel scenarios and high efficiency in distinguishing anomalies within video sequences (Singh et al., 2025). The approach achieved an AUC of 0.91 on the MSAD dataset, illustrating how

multisituational anomaly detection strategies can be adapted to complex and variable environments. Wang et al. (2025) analyzed deep anomaly detection methods in multidimensional time series, highlighting the strengths and limitations of various architectures, including autoencoders, generative adversarial networks, and transformers. Particular emphasis was placed on challenges such as distributional variability, limited interpretability, and the necessity of integrating statistical measures such as Mahalanobis distance to enhance detection accuracy and robustness. An empirical approach was used to analyze a series of two online titles to identify anomalies in player behavior. Four unsupervised methods (LOF, KDE, K-means, GMM) were used to identify deviant users (Dinh et al., 2016). The evaluation was conducted using artificial and real data from the JX2 and Chan games. The performance analysis was based on detection accuracy, number of anomalies identified, and execution time, and showed the advantage of parametric models in terms of speed and comparable accuracy compared to non-parametric methods.

Irvan et al. (2024) applied LSTM networks to anomaly detection in the context of the e-sports game CS:GO, focusing on player movement patterns. The model was trained in normal sequences, adding synthetic anomalies such as teleportation and bot behavior. Experiments have shown that LSTM reliably detects deviations in real time, with an accuracy of over 90 percent under standard conditions, but with a drop in performance with a high proportion of anomalous patterns.

While prior research confirms the versatility of Mahalanobis distance in anomaly detection across a range of domains, its application to market-facing behavioral data remains underexplored. The present study addresses this gap by introducing a multivariate framework for anomaly identification in video games, explicitly designed to balance statistical precision with interpretability and market relevance.

METHODOLOGY

The analytical framework of this research is based on multidimensional anomaly detection using Mahalanobis distance, with the integration of auxiliary statistical mechanisms such as z-score and interquartile range (IQR) to increase robustness in identifying deviant instances in multidimensional user space (Hair et al., 2019). The entire approach was conducted in a quantitative regime, applying linear algebraic methods and empirical covariance matrices.

Variables and Standardization

To analyze latent patterns of similarity between video games, six multidimensional variables were selected to capture essential aspects of pricing, user engagement, and rating distribution. The selection criteria were based on the potential of each variable to reflect user behavior and market positioning. All data was obtained from the publicly accessible Steam Games Dataset available on the Kaggle platform.

The data sample comprises $n = 1000$ unique video games, each described by six variables as follows:

- price (x_1): the official retail price of the game in euros, based on the most recent listed value in the store database,
- number of recommendations (x_2): the total number of user-submitted positive endorsements recorded on the platform,
- percentage of positive reviews (x_3): the ratio of positive user evaluations to the total number of reviews, expressed as a percentage,
- percentage of negative reviews (x_4): the corresponding share of negative user evaluations, also expressed as a percentage,
- estimated owners (x_5): the approximated number of individual users who own the game, based on publicly available player count ranges and platform-specific estimation methods.

Each instance $x_i \in R^5$ represents a row of the data matrix $X \in R^{n \times 5}$,

To ensure metric homogeneity and eliminate the effects of differing variable scales, z-score standardization was applied to each variable:

$$z_{ik} = \frac{x_{ik} - \mu_k}{\sigma_k}$$

for $i = 1, \dots, n$ and $k = 1, \dots, 5$, where:

$$\mu_k = \frac{1}{n} \sum_{i=1}^n x_{ik}$$

denotes the empirical mean of variable x_k , and the empirical standard deviation is defined as:

$$\sigma_k = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_{ik} - \mu_k)^2}$$

These variables were chosen based on their ability to reflect user valuation, engagement intensity, and perceived quality, offering a multidimensional perspective of game performance. The resulting matrix of standardized values is denoted by $\mathbf{Z} \in R^{n \times 5}$, where \mathbf{z}_i is the i -th row vector.

Empirical covariance matrix

The covariance matrix captures the linear interdependencies among standardized variables, serving both as the mathematical foundation for Mahalanobis distance computation and as a central construct in multivariate statistical analysis (Hair et al., 2019).

Based on the standardized matrix \mathbf{Z} , the empirical covariance matrix $\Sigma \in R^{5 \times 5}$ was computed as:

$$\Sigma = \frac{1}{n-1} \mathbf{Z}^\top \mathbf{Z}$$

If the matrix Σ exhibits numerical instability or is close to singular, regularization is applied following the Tikhonov principle:

$$\Sigma_\lambda = \Sigma + \lambda \mathbf{I}_5$$

where $\lambda \in R^+$ denotes a small regularization parameter (typically $\lambda = 10^{-2}$), and \mathbf{I}_5 is the 5×5 identity matrix.

Mahalanobis Distance

The Mahalanobis distance for the i -th instance is defined in quadratic form as:

$$D_i^2 = (\mathbf{z}_i - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{z}_i - \boldsymbol{\mu})$$

Given that all vectors are standardized ($\boldsymbol{\mu} = \mathbf{0}_5$), the expression simplifies to:

$$D_i^2 = \mathbf{z}_i^\top \Sigma^{-1} \mathbf{z}_i$$

For the entire set of n instances, the vector of Mahalanobis distances is obtained as the diagonal of the matrix product:

$$\mathbf{D}^2 = \text{diag}(\mathbf{Z} \Sigma^{-1} \mathbf{Z}^\top)$$

This vector of squared distances serves as the criterion for deviation detection. In this context, the formulation enables robust anomaly identification by highlighting instances that are structurally distant from the global feature centroid, while accounting for both scale and correlation effects.

By incorporating the inverse covariance matrix as a scaling transformation, Mahalanobis distance generalizes the Euclidean metric to ellipsoidal contours aligned with the data's variance structure, enabling anisotropic sensitivity to deviations along correlated dimensions.

Statistical Threshold and Deviation Classification

Outlier instances are identified based on the squared Mahalanobis distance compared to the threshold derived from the chi-squared distribution with six degrees of freedom. An instance i is considered an outlier if $D_i^2 > \chi_{1-\alpha}^2(5)$. For the significance level $\alpha=0.01$, the threshold value is $\chi_{0.99}^2(5) \approx 15.086$, and for $\alpha = 0.05$ it is $\chi_{0.95}^2(5) \approx 11.070$.

This statistical decision rule leverages the asymptotic properties of the Mahalanobis distance under multivariate normality, ensuring that the rejection region corresponds to a predefined probability mass in the tail of the distribution. By anchoring deviation detection to a chi-squared quantile function, the methodology ensures a theoretically justified and probabilistically interpretable classification boundary, independent of specific distributional idiosyncrasies in the data.

Additional Robustness Criteria

To enhance the reliability of anomaly detection and identify borderline cases, supplementary statistical criteria have been applied:

- *Z-score criterion*: an instance is marked as a potential outlier if $\exists k \in \{1, \dots, 5\}$ such that $|z_{ik}| > 3$.
- *IQR criterion*: for each variable x_k threshold values are defined as follows:

$$\text{lower threshold}_k = Q_1^{(k)} - 1.5 \cdot IQR_k$$

$$\text{upper threshold}_k = Q_3^{(k)} - 1.5 \cdot IQR_k$$

where $IQR_k = Q_3^{(k)} - Q_1^{(k)}$, and $Q_1^{(k)}$ and $Q_3^{(k)}$ denote the first and third quartiles of the variable x_k . Each instance whose attribute lies outside these bounds is additionally labeled as a potential outlier.

These supplementary criteria ensure that anomalies with marginal Mahalanobis distances, but extreme univariate characteristics are not overlooked. Moreover, by incorporating both standardized deviation (z-score) and distributional spread (IQR), the model achieves a dual-layered safeguard against the misclassification of edge cases that may evade detection under multivariate-only metrics.

Implementation in Python

All data processing phases were carried out using the Python programming language (version 3.11), along with the following libraries:

- pandas: for structured data import, filtering, user-based grouping, and review aggregation (*pandas.read_csv*, *DataFrame.loc*, *groupby*, *dropna*),
- numpy: for numerical processing of vectors and matrices, including mean calculation, covariance estimation, and linear algebra operations (*numpy.array*, *numpy.mean*, *numpy.std*, *numpy.dot*),
- scipy.stats: for statistical computations such as z-scores, interquartile ranges, and chi-squared quantiles (*zscore*, *iqr*, *chi2.ppf*),
- scipy.spatial.distance: for computing Mahalanobis distances in multidimensional space (*mahalanobis*),
- matplotlib.pyplot: for creating visual representations of distributions, clusters, and deviations (*scatter*, *hist*, *subplots*, *tight_layout*),
- seaborn: for enhanced statistical visualization, including heatmaps, boxplots, and cluster maps (*heatmap*, *boxplot*, *clustermap*),
- scikit-learn: for feature standardization (*StandardScaler*), robust scaling (*RobustScaler*), and outlier evaluation (*StandardScaler*, *RobustScaler*).

The implementation strategy draws on practical principles outlined in recent literature on applied deep learning (Howard and Gugger, 2020; Elgendy, 2022).

All variables were standardized using the z-score method to ensure metric comparability during distance calculations. The Mahalanobis values per instance are stored in the variable *md_scores*, while outlier thresholds derived from the empirical χ^2 distribution are stored in *md_thresholds*. The classification of instances into regular and anomalous is saved in *md_outliers*. The full analysis was conducted in the script *mahalanobis_detection.py* and additionally documented in the Jupyter notebook *outlier_mapping.ipynb*, which includes visualizations of distributions, heatmaps, and distance-based outlier mappings.

This methodological framework establishes a rigorous statistical basis for identifying structurally atypical instances, whose empirical characteristics and distributional deviations are further explored in the following section.

RESULTS AND DISCUSSION

This chapter presents the outcomes of the multivariate deviation analysis applied to the video game dataset. The primary analytical framework is based on the Mahalanobis distance, supplemented by additional robustness metrics. The results are structured according to the behavioral variables included in the model, with separate sections detailing the inter-variable relationships, empirical distributions, and classification of instances based on deviation thresholds. Visual representations are provided to complement the numerical findings and to enhance the interpretability of detected anomalies.

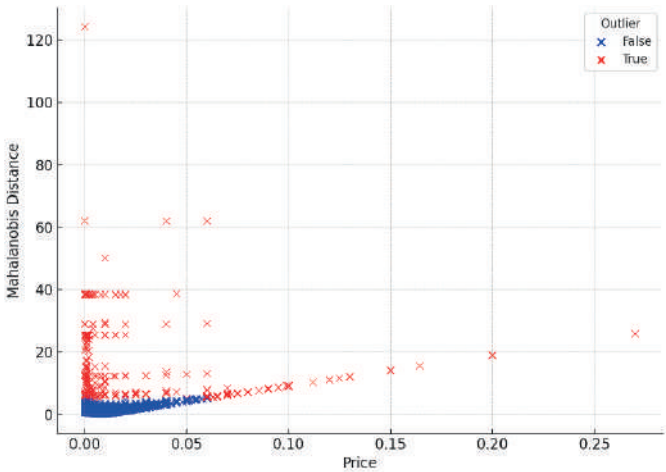


Figure 1. Multivariate Outlier Detection by Mahalanobis Distance in Relation to Game Price

The scatter plot in Figure 1 displays the relationship between video game price (x-axis) and Mahalanobis distance (y-axis), with each point representing one game. Instances marked in red correspond to statistical outliers in the multivariate feature space. The majority of games are concentrated in the lower left region of the plot, indicating lower prices and proximity to the multivariate centroid. A smaller subset of games exhibits high Mahalanobis distances, with some also characterized by exceptionally high prices. These instances suggest deviations from the general pattern, potentially reflecting atypical combinations of attributes, such as elevated cost coupled with uncommon features or low user evaluations. The observed separation indicates that these games differ not only in price, but also across multiple behavioral dimensions, thereby qualifying as potential anomalies requiring further qualitative analysis.

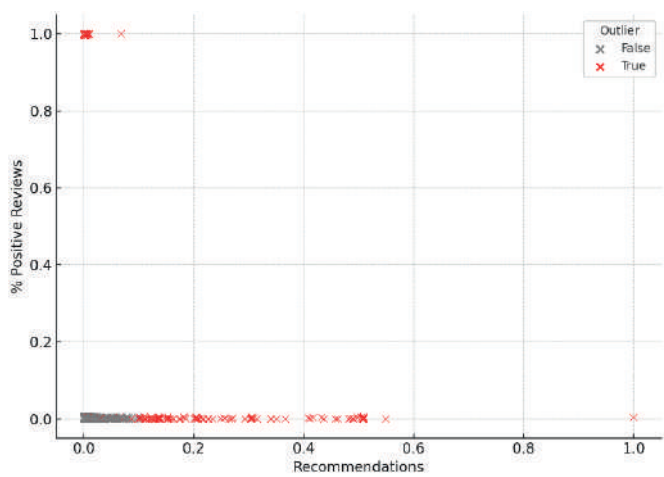


Figure 2. Outlier Distribution by Recommendation Frequency and Positive Review Ratio

Figure 2 shows the relationship between the number of recommendations (x-axis) and the proportion of positive ratings (y-axis) for each game, with outliers marked based on Mahalanobis distance (red markers). The majority of games are clustered in the lower left quadrant, indicating low values for both metrics. However, several anomalies deviate from this trend. Certain titles display a high number of recommendations despite a low proportion of positive ratings, which may suggest atypical marketing practices or manipulation of recommendation algorithms. Conversely, some games exhibit high approval rates but a minimal number of recommendations, potentially reflecting niche targeting or a narrowly defined genre. The application of Mahalanobis distance allowed for the detection of such inconsistencies, which would remain obscured using univariate methods. These findings highlight the need for further contextual investigation to explore underlying marketing, sales, or content-related factors.

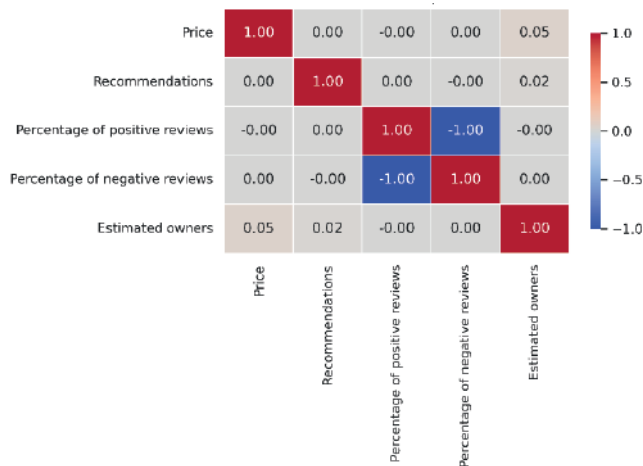


Figure 3. Correlation Matrix of All Five Standardized Input Variables

The heatmap in Figure 3 visualizes the correlation matrix of all five standardized input variables. The only strong and statistically expected correlation is a perfect negative relationship between the percentage of positive and negative ratings (-1.00), as these two measures are mathematically complementary. All other variables, including price, number of recommendations, and estimated number of owners, show very low or negligible correlations with the remaining variables, suggesting an absence of linear dependencies within the dataset. This structure supports the application of a multivariate method such as Mahalanobis distance, since univariate metrics are inadequate for capturing complex deviations. Each variable contributes unique informational value to the model without introducing substantial multicollinearity.

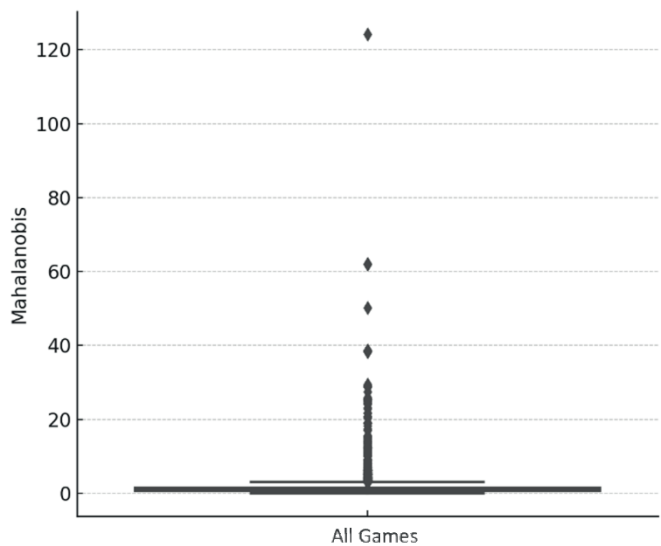


Figure 4. Distribution of Mahalanobis Distances

Figure 4 shows a boxplot of Mahalanobis distances, revealing an asymmetric distribution with a pronounced right tail. Most instances are clustered near the lower quartile, indicating low distance values and alignment with the average multivariate patterns. However, several extreme values clearly deviate from the rest, with the maximum exceeding 120. These cases represent statistically significant anomalies when all variables are considered jointly. The shape of the distribution confirms the effectiveness of the Mahalanobis metric in isolating atypical entities within a multidimensional space.

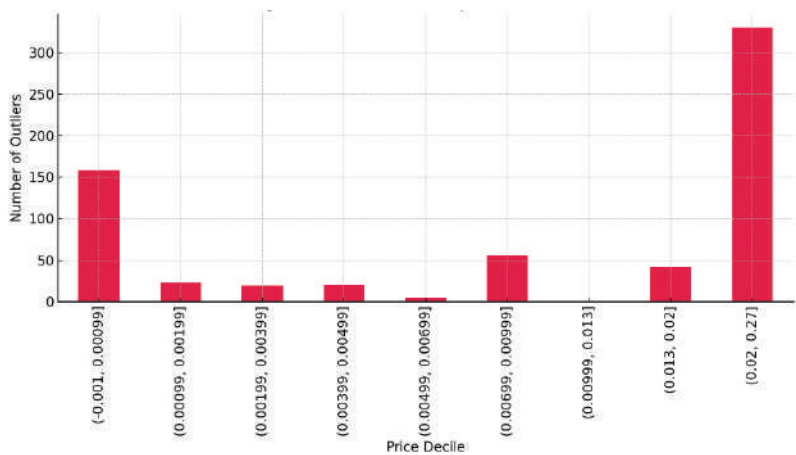


Figure 5. Distribution of Outliers Across Price Deciles

The histogram by price decile in Figure 5 indicates that extreme price values are most frequently associated with statistical anomalies. The highest number of outliers occurs in the top decile (0.20–0.27), representing the most expensive games, and in the bottom decile (0.001–0.00099), corresponding to the cheapest titles. In contrast, middle price ranges show a markedly lower incidence of deviations, suggesting greater consistency in market positioning and user feedback. This distribution supports the hypothesis that games with extreme prices are more statistically unstable within the multidimensional feature space.

Overall, the analysis confirms that multidimensional methods such as Mahalanobis distance are effective in detecting statistically atypical patterns, particularly for games that diverge in terms of pricing or user perception. These findings offer a robust foundation for further investigation of market anomalies and potential mechanisms of visibility or reputation manipulation.

CONCLUSION

This study demonstrates that multiple anomaly detection is an effective approach for identifying video games that deviate from expected behavioral patterns and typical attribute values. By applying Mahalanobis distance within a standardized feature space, the analysis detects complex deviations that would not be apparent using simpler metrics. The identified anomalies include games with disproportionately high prices relative to low user ratings and titles with an unusually high number of recommendations despite overall poor scores. Such inconsistencies suggest potential distortions in market visibility or recommendation mechanisms.

The findings underscore the limitations of relying solely on conventional popularity and rating metrics and confirm the added value of objective, data-driven evaluation models. Visualizations further support the interpretation of these deviations and reinforce the robustness of the analytical procedure. Additionally, the results indicate that anomaly detection methods can enhance transparency in digital markets and recommendation systems by revealing items that diverge from normative patterns.

Future extensions of this framework could incorporate temporal usage patterns, session-level behavior, or genre-based features. Transitioning from unsupervised to supervised models may also improve classification accuracy and result interpretability. The methodology presented here thus provides a repeatable and adaptable basis for the systematic detection of anomalies in interactive digital content.

The Python script used for the analysis is available on request.

References

- Azizi E, Zaman L, 2023. Automatic Bug Detection in Games using LSTM Networks. IEEE Conference on Games CoG 2023, August 21-24, 2023, 1-4, Boston, MA, USA.
- Dinh PV, Nguyen TN, Nguyen QU, 2016. An Empirical Study of Anomaly Detection in Online Games. 3rd National Foundation for Science and Technology Development Conference on Information and Computer Science NICS 2016, September 14-16, 2016, 171-176, Danang City, Vietnam.
- Dobos D et al., 2023. A comparative study of anomaly detection methods for gross error detection problems, *Computers and Chemical Engineering*, 175(1): 108263.
- Elgendy M, 2020. *Deep Learning for Vision Systems*. 1st ed., Manning Publications Co, USA.
- Hair JF et al., 2019. *Multivariate Data Analysis: A Global Perspective*. 7th ed., Pearson, USA.
- Howard J, Gugger S, 2020. *Deep Learning for Coders with fastai & PyTorch*, 1st ed., O'Reilly, USA.
- Irvan M et al., 2024. Anomaly Detection in eSport Games Through Periodical In-Game Movement Analysis with Deep Recurrent Neural Network. 6th International Conference on Neural Computation Theory and Applications NCTA 2024, November 20-22, 2024, 430-437, Porto, Portugal.
- Kamoi R, Kobayashi K, 2020. Why is the Mahalanobis Distance Effective for Anomaly Detection?
- Kurose JF, Ross KW, 2016. *Computer Networking: A Top-Down Approach*. 7th ed., Pearson, USA.
- Lin Y, Li X, 2024. Back to the Metrics: Exploration of Distance Metrics in Anomaly Detection, *Applied Sciences*, 14(16): 7016.
- Mozaffari M, Doshi K, YilmazY, 2022. Online Multivariate Anomaly Detection and Localization for High-Dimensional Settings, *Sensors*, 22(21): 8264.
- Pinto JP, Pimenta A, Novais P, 2021. Deep Learning and Multivariate Time Series for Cheat Detection in Video Games, *Machine Learning*, 110(11-12): 3037-3057.
- Singh DK et al., 2025. Meta-Learning Approach for Adaptive Anomaly Detection from Multi-Scenario Video Surveillance, *Applied Sciences*, 15(12): 6687.
- Wang F et al., 2025. A Survey of Deep Anomaly Detection in Multivariate Time Series: Taxonomy, Applications, and Directions, *Sensors*, 25(1): 190.

Wu R et al., 2021. Improving video anomaly detection performance by mining useful data from unseen video frames, *Neurocomputing*, 432(1): 523-533.

Zhang H et al., 2024. A Distance-based Anomaly Detection Framework for Deep Reinforcement Learning, *Transactions on Machine Learning Research*, 10(1): 1-38.

Conflict of Interest

The author has declared that there is no conflict of interest”.

Author Contributions

All aspects of the study, including conceptualization, methodology, analysis, and writing, were carried out by the sole author.

Standardization in Elevator Systems: Call and Display Panel Design Compatible with CANopen Communication Protocol

Eyup Sayin¹

Murat Topuz²

Muhammet Fatih Aslan³

Akif Durdu⁴

Abstract

Since each of the elevator motherboards produced in Turkey today uses its own special communication protocol, the panels must be adapted separately to each system, and this leads to serious time and cost loss in the production process. For this purpose, this study aims to ensure that elevator control cards developed by different manufacturers work in harmony with common call and display panels. In this context, a call and display panel compatible with the CANopen Lift (CIA-417) profile, one of the international standards, has been designed. The embedded software infrastructure was created with the CANopenNode open-source library and the system's error analysis was performed with the CAN Analyzer. In addition, the sustainability and usability of the system was increased with the architecture that allows software updates and parameter changes to be made remotely. The prototype developed at the end of the project was tested in pilot elevators and successful results were obtained. With this study, a domestic solution that will reduce Turkey's dependence on imports has been presented. In addition, the way will be paved for CIA-417 compatible panel production.

-
- 1 Butkon Elevator Industry Trade Inc., Research and Development Center, Konya, Türkiye
 - 2 Butkon Elevator Industry Trade Inc., Research and Development Center, Konya, Türkiye
 - 3 Karamanoglu Mehmetbey University, Faculty of Engineering, Electrical and Electronics Engineering, Karaman, Türkiye
 - 4 Konya Technical University, Faculty of Engineering, Electrical and Electronics Engineering, Konya, Türkiye

INTRODUCTION

Today, elevators have become an indispensable and frequently used tool in people's lives. The increase in urbanization, verticalization in building designs, and the increase in the number of elderly people have made the use of elevators an inevitable necessity in both residential and commercial buildings (Ming et al., 2018). Modern elevator systems not only serve as a transportation function, but also include many factors such as energy efficiency, user safety, and interactive interface solutions as an integral part of smart building infrastructures. At this point, call and display panels, which are an important component of elevator systems, are the basic units that provide real-time data exchange between the user and the system (Al-Kodmany, 2023a).

Due to reasons such as increasing user density, high number of floors, and need for solutions for emergency situations, it has become important for these panels to be faster, safer, adaptable and have a standard structure (Al-Kodmany, 2023b). Ergonomic approaches in elevator design are extremely important in terms of increasing comfort, reducing physical strain and ensuring ease of use for users of all ages and abilities (Perrucci et al., 2025). In this context, with the spread of IoT-supported smart systems, especially within the framework of building automation systems and Industry 4.0, elevators are expected to be able to communicate with other building components, integrate with central management systems and be monitored remotely (Van et al., 2020). Despite these developments, there is a major lack of standardization among call and display panels used today. This lack is due to incompatibilities between control cards and panel systems from different manufacturers, which results in increased costs, longer integration times and update difficulties.

The fact that control cards (motherboards) from different manufacturers use their own special communication protocols requires that call and display panels in elevator systems be customized for each manufacturer. This complicates the production process and causes serious time and cost losses in maintenance and compatibility processes. In particular, due to different communication protocols, special software and hardware adaptations must be made so that the panels can work integrated with each control card. It is stated that this situation limits the scalability and modularity of the systems. (Pöttner et al., 2012). In addition, most of the existing panels do not have the infrastructure to meet modern industrial needs such as remote software update, parameter change or fault diagnosis. This deficiency reduces the flexibility of the systems and prevents rapid response to the changing needs

of the users. In this context, the development of universal panel systems based on a common protocol has become a great necessity. Standards such as CANopen and specifically CAN in Automation (CiA)-417 (CiA, 2025) have been developed to solve this problem, allowing devices from different manufacturers to work together on the same communication network. Such standardization facilitates integration between devices and increases system reliability by reducing maintenance costs (Arleklint, 2019).

The CANopen protocol, which is widely used in industrial automation systems, enables many devices to communicate over the same line with its flexible and modular structure. CANopen is an open and standard communication protocol that allows devices from different manufacturers to work together. This protocol supports real-time data exchange between devices and facilitates system configuration (Embien, 2025). The CiA-417 profile, a version of the CANopen protocol specifically developed for elevator systems, is defined to provide compatible and secure communication between call and display panels and control systems. CiA-417 specifies communication interfaces for virtual devices used in elevator systems. This allows components such as call buttons, cabin displays, and door control units to communicate over a standard protocol. The CiA-417 profile provides detailed object dictionaries for both data exchange and parameter management, enabling the development of systems that comply with the plug-and-play principle. This allows devices from different manufacturers to work seamlessly on the same communication network, thus facilitating system integration (CiA, 2025).

In recent years, various researches have been carried out on the use of CANopen and CiA-417 protocols in elevator systems. Sußmann and Meroth (2017) detailed the model-based development processes of elevator components that comply with the CiA-417 standard. This study demonstrated how the integration processes of elevator systems can be optimized using Matlab/Simulink-based modeling and Hardware-in-the-Loop (HIL) test methods. It was also stated that systems developed using the CANopen protocol allow components from different manufacturers to work seamlessly on the same communication network. Such studies reveal the potential of CANopen and CiA-417 protocols in terms of integration and performance improvements in elevator systems. Pöttner et al. (2012) evaluated the use of Bundle Protocol (BP) in Delay Tolerant Wireless Sensor Networks (DT-WSN) and demonstrated the practical effectiveness of this protocol with the “Data Elevator” application, where they used the physical movement of the elevator as a carrier in data transmission. In the study, a communication infrastructure that is both compliant with standards

and flexible was established by using BP over IEEE 802.15.4 based communication with low-power sensor nodes. In addition, it was shown that the communication load on embedded systems can be minimized and full data delivery can be provided even in delayed data transfer with the modular BP implementation called μ -Delay Tolerant Networking (μ DTN). This approach is an important example to overcome similar difficulties in the process of integrating industrial protocols such as CANopen into elevator systems. A different study on the verification of real-time communication and the development of test systems compatible with the CANopen protocol was done by Sußmann and Meroth (2017). An FPGA based real-time simulation of an absolute encoder conforming to the CANopen CiA 406 standard has been performed, and a model that can be directly embedded into hardware has been developed with HDL Coder using the Simulink-RTL flow. The simulator is configured to support both cyclic broadcasting of TPDO1 messages and timer-based protocol mechanisms such as heartbeat. Thus, the CANopen-based communication of the elevator controller has been tested without the need for physical hardware. In this respect, the study shows that CANopen communication tests can be modeled in a reusable, modular and platform-independent manner.

This study aims to develop a CIA-417 compatible call and display panel, which does not yet have a domestic example in Turkey. The developed system has gone through many technical stages such as embedded software infrastructure integrated with the CANopenNode open-source library, CAN Analyzer supported communication tests, remote software update capability and pilot field tests. In this respect, the study offers a low-cost, sustainable and integrable solution for both manufacturers and users. It also aims to reduce import dependency by encouraging domestic production.

MATERIAL AND METHODS

In this study, indicator-call panels were designed and developed by considering parameters such as floor, direction, door and special indicators suitable for CiA-417 profile. Within the scope of the project process, firstly membership to CiA community was provided and necessary technical documents were obtained. Then, a CiA-417 compatible motherboard and test panels were supplied and test software such as CAN Analyzer and CANopen Magic Professional were used to analyze the communication infrastructure.

In the hardware design process for the panel prototype, printed circuit board (PCB) schematic and layout studies were carried out, and then

PCB production and supply services were provided as outsourced. In the embedded software development phase, core software compatible with the microcontroller (MCU) was developed based on the open source CANopenNode Library. In this context, object dictionaries for CiA-301 and CiA-417-1/2/3/4 profiles were created, TPDO (Transmit Process Data Object) and RPDO (Receive Process Data Object) packages were defined and communication tests were successfully performed.

In the developed system; basic components such as microcontroller, step-down converter module, LED dot matrix module, button I/O connection and CAN communication unit were used. The hardware components preferred in the study and the methods followed are presented in detail below.

Material

As a microcontroller, NUVOTON's ARM Cortex-M0 family was preferred due to its integrated CAN peripheral unit, low cost and sufficient performance. For visual display purposes, a 10x15 matrix consisting of LEDs was configured on the PCB. Darlington transistors were used for switching this matrix, and shift registers were used for connecting it to the microcontroller. Transistor switching circuits were designed to detect button inputs and drive call acceptance LED outputs. For CAN Bus communication, the SN65HVD1050 integrated circuit, manufactured by Texas Instruments and optimized for electromagnetic interference (EMC), which is widely used in the industry, was preferred. CAN analyzer was used for CAN line communication control and analysis, and PCAN Analyzer, manufactured by PEAK System, which has international support, was preferred. CANopen Magic software, manufactured by PEAK System, was used to display CANopen structures (NMT (Network Management), Error, Heartbeat, Service Data Object (SDO), Process Data Object (PDO)) on the computer. The hardware architecture and CANopen communication structure of the developed system are presented schematically in Figure 1.

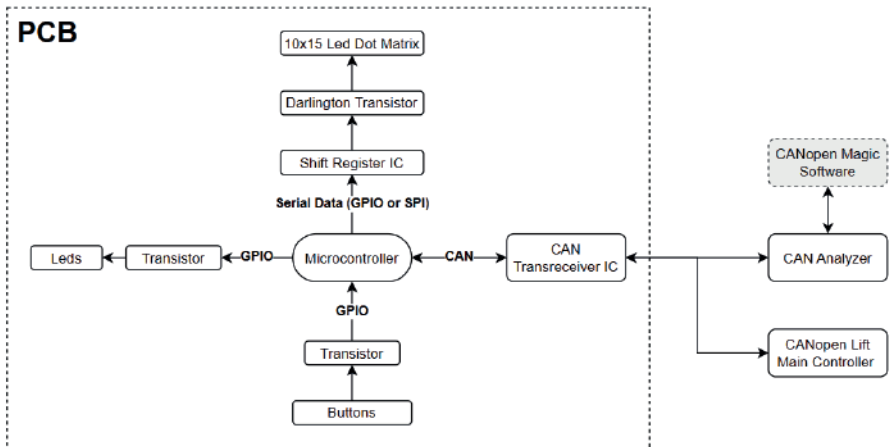


Figure 1. Microcontroller-based hardware architecture and CANopen communication structure for the developed elevator floor button panel.

Software Architecture

In the hardware-based software development process, Keil uVision, a free development environment offered by NUVOTON for the ARM Cortex-M family, was used. The system software consists of three main modules: display control, peripheral management, and CANopenNode optimization. The display control module was designed using the GPIO and TIMER peripherals of the microcontroller. The LED positions were controlled via the GPIO unit, while the matrix timing was managed with the TIMER unit. Peripheral management was implemented using the Board Support Package (BSP) library provided by NUVOTON. The CANopenNode middleware was provided as open source via GitHub and was made compatible with the peripherals of the NUVOTON microcontroller and used in the system.

CANopen Protocol

CANopen is a high-level communication protocol developed by the CiA organization and operates on the CAN bus infrastructure. When the fundamental differences between CANbus and CANopen are examined, it is seen that CANopen offers advantages such as advanced error monitoring, flexible structure, network management and real-time communication. In this context, in industrial applications requiring complex and reliable communication, it would be more appropriate to prefer CANopen over the basic CANbus protocol. A comparison of CANbus and CANopen protocols in terms of various technical features is summarized in Table 1.

Table 1. Comparison of CANbus and CANopen protocols in terms of technical features

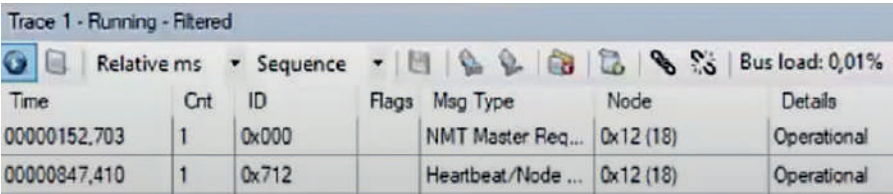
Feature	CANbus	CANopen
Definition	Low-level message-based communication protocol.	A high-level communication protocol built on CANbus that enables standardized communication between devices.
Layer	<p>It consists of 2 layers:</p> <ul style="list-style-type: none"> • Data Link Layer • Physical Layer 	<p>It consists of 7 layers:</p> <ul style="list-style-type: none"> • Application Layer • Presentation Layer • Session Layer • Transport Layer • Network Layer • Data Link Layer • Physical Layer
Standardization	ISO 11898	EN 50325-4 (Developed by CiA)
Scope of Communication	Provides prioritized, unstructured data transmission based on message identifiers.	Provides structured communication covering device configuration, state management, and real-time data transmission.
Contact Method	Data transmission is performed based on Message Identifier.	Function-based communication is achieved using Node ID and COB-ID (Communication Object Identifier).
Application Layer Support	Not defined.	Provides a comprehensive structure at the application layer, including device profiles, object dictionary, and communication services.
Data Structure	Basic CAN data frames with a fixed structure (ID, control bits, data field, CRC).	Includes configurations such as Object Dictionary, SDO and PDO.
Real Time Communication	Limited; enables fast communication based on message priorities.	Provides deterministic and synchronized real-time data transmission via PDO and SYNC messages.
Device Configuration	Not supported; configuration is not possible at the protocol level.	Device configuration and parameter settings can be performed via SDO services.
Network Management	None.	Device state management is provided through the NMT mechanism (Operational, Pre-operational, Stopped).
Error Tracking and Management	Basic error control is performed (bit error, CRC, frame error).	Device health and connection status are continuously monitored through Heartbeat and Node Guarding mechanisms.

Manufacturer Independent Compatibility	Not provided; requires custom integration for device-to-device communication.	Thanks to standard device profiles (e.g., CiA 401, CiA 417), devices from different manufacturers can operate compatibly.
Typical Application Areas	Automotive ECU communication, in-vehicle control units, embedded systems.	Industrial automation (PLCs, HMIs, drives), medical devices, elevator control systems, robotic applications.

CANopenNode Integration Control

After the integration of the CANopenNode library and the BSP (Board Support Package) library of the microcontroller is completed, a test process is applied to verify whether the CANopenNode is working properly. During this verification process, the PCB is connected to the computer via a CAN analyzer and is put into operation by applying power. CANopen Magic software is started on the computer and the network management interface is accessed via the “View >> Network Management” menu. In this interface, the “All Nodes” option is activated and the status of all nodes in the network can be monitored. Then, one of the “Operational”, “Preoperational” and “Stopped” states defined in the CANopen protocol is selected and the relevant command is sent. In response to this command, a response message is expected to be sent by the PCB.

During this process, the transmitted and received messages are monitored using the Trace window. The monitoring screen of the CANopen messages received during the test process is shown in Figure 2. As shown in Figure 2, if the expected response message is detected successfully, it is verified that the integration of the CANopenNode library with the BSP library is performed correctly.



The screenshot shows the 'Trace 1 - Running - Filtered' window of CANopen Magic. It displays a table of CAN messages. The first message is an NMT Master Request (ID 0x000) from node 0x12 (18) at time 00000152.703, with a criticality of 1 and status 'Operational'. The second message is a Heartbeat/Node Status (ID 0x712) from node 0x12 (18) at time 00000847.410, with a criticality of 1 and status 'Operational'. The bus load is 0.01%.

Trace 1 - Running - Filtered						
Relative ms Sequence Bus load: 0,01%						
Time	Crit	ID	Flags	Msg Type	Node	Details
00000152,703	1	0x000		NMT Master Req...	0x12 (18)	Operational
00000847,410	1	0x712		Heartbeat/Node ...	0x12 (18)	Operational

Figure 2. Monitoring of “NMT Master Request” and “Heartbeat” messages from the PCB after CANopenNode integration using a CAN analyzer software (The figure is taken from the trial version of CANopen Magic Professional software).

Usage of CANopenEditor

After verifying the integration of the CANopenNode library with the microcontroller, the object dictionary must be created in order for the system to fully provide its functionality. The object dictionary is the basic structure that defines the parameters to be used during device operation and plays a critical role in managing data exchange within the CANopen protocol.

The parameters defined in the object dictionary are divided into three main categories according to the index values they receive: communication-specific parameters, manufacturer-specific parameters, and device profile parameters. This classification is made to standardize data organization and ensure interoperability between devices.

In this study, CANopenEditor software, which was developed in accordance with CANopenNode library, is used. The related application can be accessed from <https://github.com/CANopenNode/CANopenEditor/releases>. The user interface for the process of defining a new object word is shown in Figure 3. Also, as an example, the process of adding a new object word under the heading “Manufacturer Specific Parameters” is explained below:

1. The CANopenEditor application is started and a new project is created using the File >> New menu.
2. To save the created project, follow the “File >> Save Project” path. The location where the project will be saved is the folder containing the “CO_App” and “CO_Driver” directories, outside the CANopenNode protocol stack.
3. In order to add the basic communication parameters to the object dictionary, go to the “Insert Profile >> DS301_profile.xpd” extension from the tabs and add the parameters of the DS301 profile with the “Insert” command in the opened window.
4. Then, go to the “Object Dictionary” tab and define a new object word with the “Add” option under the “Index” heading.
5. As shown in Figure 3, the creation process is completed by entering the index value, name and object type information for the relevant object (Create).

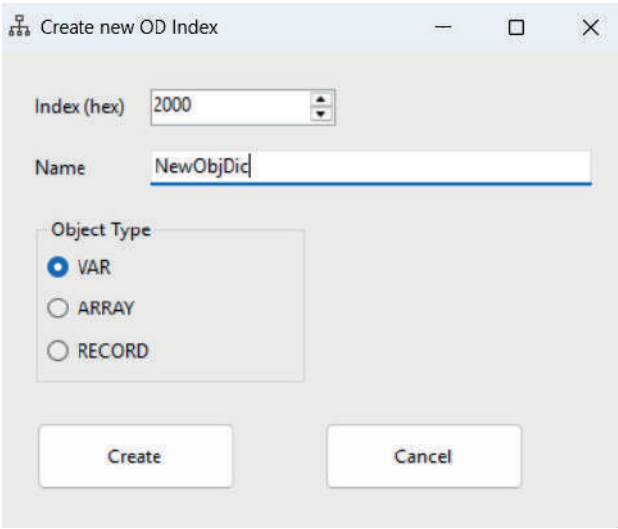


Figure 3. Screen for creating a new Object Dictionary entry in the CANopenEditor interface

As a result of these operations, the object dictionary structure that complies with the CANopen protocol is created and the device’s communication, manufacturer-specific and profile-based parameters are systematically defined. In the process of configuring the properties of the created object dictionary index, the data type, SDO access permissions and PDO mapping settings must be defined. This step directly affects the functionality of the device by determining the object access format, data type and its relationship with communication protocols. For example, let’s define the data type for the created object word as unsigned32, SDO access as free and TPDO. The settings that need to be made in this case are shown in Figure 4.

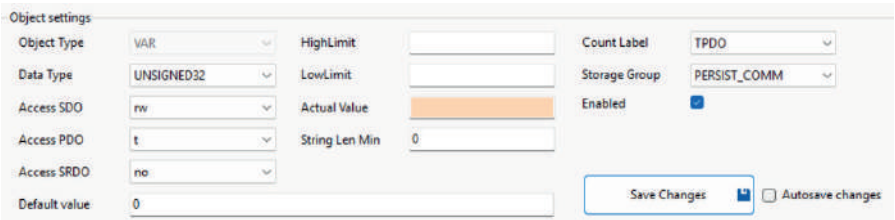


Figure 4. Configuration of data type, access rights, and PDO mapping settings for an object dictionary index in the CANopenEditor interface.

When you switch to the “TX PDO Mapping” tab, it is expected that the previously created object word will be displayed in the “Available Objects for PDO” section. At this stage, one of the default TPDO indexes is selected and the newly created object index is used as the data to be transmitted in the content of this TPDO. Thus, the transmission of the relevant object dictionary entry on the network via TPDO messages is configured. The structured version of the TPDO content to include the newly created object dictionary entry is shown in Figure 5. In order for the process to be completed, the necessary software files are updated using the “File >> Export CANopenNode” option and the configuration for the TPDO broadcast is completed.

ID	COB	Index	Byte 0	Byte 1	Byte 2	Byte 3	Byte 4
1	COB...	1800	0x2000/00/NewObjDic				Empty
2	COB...	1801	Empty				
3	COB...	1802	Empty				
4	COB...	1803	Empty				

Figure 5. TPDO configuration and mapping of the newly defined object dictionary entry to PDO messages in the CANopenEditor interface

CIA-417(CAN In Automation - Lift)

The CIA-417 profile is based on the modular development of elevator systems and offers a structure consisting of eleven basic units. The units in question are defined as call controller unit, input panel unit, output panel unit, car door controller unit, car door unit, light obstacle detection unit, car driver controller unit, car driver unit, car position unit, load measurement unit and remote data transmission unit. The CIA-417 profile minimizes compatibility problems for system integrators and increases the efficiency of product development processes by providing manufacturer-independent and uninterrupted communication between these units.

The modular structure offered by the profile allows elevator components of different brands and models to be easily integrated into the system, thus enabling the development of customizable and scalable system designs. In this context, CIA-417 supports the implementation of decentralized and distributed control architectures, offering a simpler infrastructure compared to traditional cabling methods. This reduces the overall complexity of the system and significantly reduces maintenance costs.

Real-time data transmission, which is of critical importance in elevator control systems, is provided with high reliability via the CIA-417 profile. This feature offers significant advantages in terms of passenger safety and system performance. In addition, the error diagnosis and fault management features provided within the framework of the CANopen protocol have been specially adapted to elevator applications with CIA-417, allowing for rapid and effective detection of possible faults in the system. In addition, it has become possible to effectively manage remote monitoring and maintenance processes.

The CIA-417 profile has become one of the standards that the elevator industry globally, especially in Europe, takes as a reference. This provides a significant advantage for manufacturers in terms of competitiveness in the international market. CIA-417, which responds to basic requirements such as standardization, flexibility, cost-effectiveness and reliability, is considered a communication protocol that should be preferred primarily in modern elevator systems, as it offers ease of integration, effective management of maintenance processes and a high-performance communication infrastructure for both manufacturers and end users (Feiter et al., 2013).

RESULTS

The panel prototype developed with the supplied motherboard was operated synchronously and the communication compatibility of the system was tested with the CAN analyzer. All tests and experimental analyzes were carried out at BUTKON R&D center (BUTKON, 2025). PDO broadcasts and data reception were successfully performed via the panel, call recording and floor display functions worked. In the field tests conducted on the pilot elevator system, the device worked as plug-and-play and no communication error was observed. In addition, the system's remote software update and parameter adjustments were made possible.

The communication process between the developed product and an elevator controller that complies with the CIA-417 profile was analyzed. Time-based evaluations were made based on the data obtained using CANopen Magic Professional software. As a result of the examinations, it was determined that the average response time of the broadcasted NMT and SDO messages was 1.2 milliseconds. In addition, during the NODE-ID scanning process performed by the elevator controller at the beginning of the system, the average device detection time was observed as 0.9 seconds. In the analysis conducted on 10,000 messages sent to evaluate the transmission reliability, it was determined that no data loss occurred and all messages

were successfully responded to. When an erroneous message packet was created and sent with TPDO, it was observed that the erroneous data packet was detected and the feedback time to the system was a maximum of 10 milliseconds. These results show that the system operates with high compatibility and reliability with other components integrated into the CANopen network structure. The developed system was integrated into elevator systems of different brands and models and subjected to field tests. Pilot applications were carried out on both passenger and freight elevators, taking into account different building typologies and traffic densities. The performance data obtained are summarized as follows:

- System integration time was 30% shorter on average compared to existing elevator control systems.
- Call detection and feedback time was found to be satisfactory in terms of user interaction, remaining under 100 ms.
- Thanks to the flexible configurability of the panel interface, 100% functional compatibility was achieved in different building scenarios.
- In all applications, the system's impact on data traffic over the CANopen network was minimal, and no deterioration in network stability was observed.

In general, the findings show that the proposed design both provides technical compliance with the industry standard and supports operational efficiency in the field. In this context, the study presents an exemplary application model for modular, adaptable and sustainable control panel designs in elevator systems.

CONCLUSION AND DISCUSSION

This study aims to develop a local call and display panel prototype that is fully compatible with the CANopen-Lift (CIA-417) standard. In short, a system has been designed that makes the display panels used in elevators more reliable, flexible and compatible. The developed prototype has successfully passed both software and hardware tests and has proven its reliability by providing error-free real-time data transmission.

Compared to traditional panels, this new design has significant advantages. Since it has a more modular structure, it is much easier to install and maintain. In addition, thanks to the use of open-source CANopenNode software, development costs have been reduced and rapid prototyping has been provided. In field tests conducted in a real elevator environment, it has been observed that the system works reliably and no communication errors

have occurred. Moreover, thanks to remote software update and parameter adjustment features, maintenance processes will be easier, thus reducing the need for physical intervention in the field.

Moreover, the fact that a developed elevator panel complies with the CIA-417 standard ensures that it works seamlessly with control units from different manufacturers. In other words, compatibility and scalability are ensured between various systems used in the sector. When the developed system is compared to existing commercial solutions, it draws attention with its suitability for domestic production, low cost and compliance with international standards.

Of course, the system also has some shortcomings. It currently only works with a wired connection and therefore features such as wireless software update or cloud integration are not yet supported. However, in future versions, remote monitoring and maintenance of the system will be possible with the addition of IoT technologies or wireless communication protocols such as Wi-Fi. In addition, reliability and efficiency can be further increased with artificial intelligence-supported failure prediction systems.

As a result, this developed elevator panel reveals the importance and advantages of communication standards used in the sector. The study provides a solid foundation for developing smart elevator systems. It also contributes to domestic production by reducing dependency on imported technology.

References

- Al-Kodmany, K. (2023a). Elevator Technology Improvements: A Snapshot. *Encyclopedia*, 3(2), 530-548. <https://www.mdpi.com/2673-8392/3/2/38>
- Al-Kodmany, K. (2023b). Smart elevator systems. *Journal of Mechanical Materials and Mechanics Research*, 6(1).
- Arleklint, T. (2019). Memory Measurement and Message Usage Improvement on an Elevator Embedded System. In.
- BUTKON. (2025). Retrieved 23.05.2025 from <https://www.butkon.com/>
- CiA. (2025). *CiA 417 series: Profile for lift control systems*. Retrieved 07.05.2025 from <https://www.can-cia.org/can-knowledge/cia-417-series-profile-for-lift-control-systems>
- Embien. (2025). *Understanding the CANopen Protocol*. Retrieved 07.05.2025 from https://www.embien.com/industrial-insights/understanding-the-canopen-protocol?utm_source=chatgpt.com
- Feiter, G., Fredriksson, L.-B., Hoffmeister, K., Pauli, J., & Zeltwanger, H. (2013). Higher Level Protocols. In *CAN System Engineering: From Theory to Practical Applications* (pp. 173-254). Springer.
- Ming, Z., Han, S., Zhang, Z., & Xia, S. (2018). Elevator Safety Monitoring System Based on Internet of Things. *International Journal of Online Engineering*, 14(8).
- Perrucci, G., Costa, M., Giacomello, E., & Trabucco, D. (2025). Assessing Comfort and Safety in Use of Elevators' Human-Machine Interaction Devices. *Buildings*, 15(5), 709. <https://www.mdpi.com/2075-5309/15/5/709>
- Pöttner, W.-B., Büsching, F., Von Zengen, G., & Wolf, L. (2012). Data elevators: Applying the bundle protocol in delay tolerant wireless sensor networks. 2012 IEEE 9th International Conference on Mobile Ad-Hoc and Sensor Systems (MASS 2012),
- Sußmann, N., & Meroth, A. (2017, 12-15 Sept. 2017). Model based development and verification of CANopen components. 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA),
- Van, L. D., Lin, Y. B., Wu, T. H., & Chao, T. H. (2020). Green Elevator Scheduling Based on IoT Communications. *IEEE Access*, 8, 38404-38415. <https://doi.org/10.1109/ACCESS.2020.2975248>

Acknowledgment

This research was supported by Butkon Elevator Industry Trade Inc., Research and Development Center.

Conflict of Interest

The authors have declared that there is no conflict of interest.

Author Contributions

Eyup Sayin: Methodology, investigation, visualization, writing.

Murat Topuz: Validation, Methodology.

Muhammet Fatih Aslan: Writing, supervision.

Akif Durdu: Supervision.

Improving Production with Artificial Neural Networks and Integration into ERP Systems: An Approach within the Scope of Industry 4.0

Gizem Şara Onay¹

Mehmet Çakmakçı²

Abstract

This study aims to digitize production planning by utilizing prediction models based on production data and reducing human intervention to increase efficiency. Production data obtained from a real manufacturing system through the Manufacturing Execution System (MES) interface was analyzed using an artificial neural network (ANN) algorithm, and future production quantities were predicted. By integrating the production forecast results into the Enterprise Resource Planning (ERP) system, it was aimed to automatically direct the production processes. Thus, production decisions can be automatically made by the system based on past data. As a result of the implementation, dynamic and data-driven decision-making processes in production management were facilitated through the forecast outputs integrated into the ERP system. This prediction-based approach is more flexible compared to traditional production planning methods and enables quicker responses from the production system. Consequently, this study presents an innovative approach that contributes to digital transformation within the scope of Industry 4.0 and serves as an example for decision support systems in production management. With this study, the development of predictive systems that operate with real-time data flow is aimed for the future.

1 Dokuz Eylül University, Institute of Science, İzmir / Turkey

2 Dokuz Eylül University, Faculty of Engineering, İzmir / Turkey

INTRODUCTION

The concept of Industry 4.0 encompasses digitalization, automation, and data-driven decision support infrastructures in production systems. This study enables businesses to utilize their resources more efficiently, adopt a flexible production structure, and respond more quickly to customer expectations. However, many manufacturing facilities still operate with manual data processing procedures and traditional planning methods. This leads to several operational challenges, including excess inventory, material shortages, shipment errors, and production losses.

In this study, the losses caused by the improper planning of SMD (Surface Mount Device) type materials in the production line of a company manufacturing television mainboards were examined. Since material requests are reported to the production site as a total quantity, the need on a reel basis is not fully reflected, and an excessive number of reels are sent to production. This leads to both stock inflation and layout inefficiency on the production line.

To prevent this problem, production data was collected from the past and analyzed using an artificial neural network model; future production quantities were predicted. These predictions were integrated into the ERP system, and stock control and supply planning were automated, thereby eliminating manual errors.

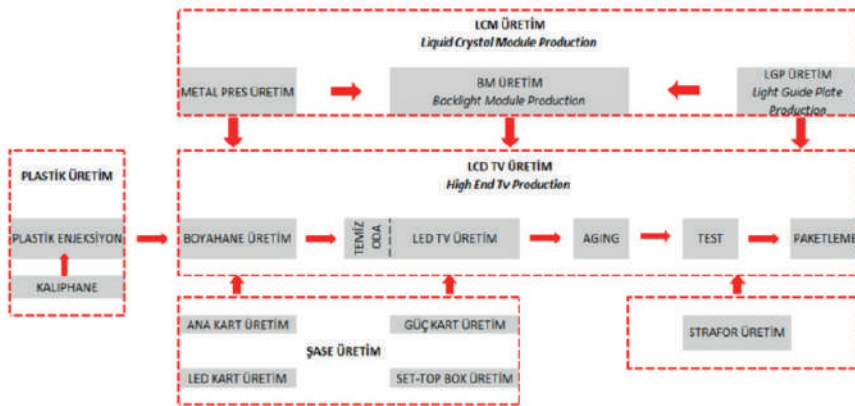


Figure 1. TV mainboard production

MATERIAL AND METHODS

Material

The mainboard production line of an electronics manufacturing facility operating in the Aegean Region was selected as the application area. The SMD materials (resistors, capacitors, etc.) used on this line are supplied in reels and automatically placed by the machines on the production line. Since the required number of reels is manually calculated based on the production quantity, over- or under-supply frequently occurs.

Theoretical Framework

In the past decade, numerous publications have been made on the concepts of Industry 4.0, MES, and ERP. The diversity of methods found in the literature stems from various approaches ranging from forecasting algorithms to system integration.

In the literature review conducted, 32 studies were examined. Some of these studies focused solely on production data forecasting and did not include ERP integration. For example, El Madany et al. (2022) proposed a hybrid time series model for supply forecasting; however, ERP integration was not included. Similarly, the ARIMA-based forecasting model developed by IFS Applications was used for demand prediction but did not support real-time MES integration.

The unique aspect of this study is the combination of ANN-based production forecasting with ERP-MES system integration. While MES systems monitor the production process, ERP systems handle corporate planning. The coordinated operation of these two systems will enhance production efficiency and the quality of decision-making.

Methods

Material tracking problems were analyzed using a fishbone diagram under the categories of material, method, human, and process:

- **Material:** Small-sized components → confusion and loss
- **Method:** Total quantity → excess reels → stock inflation
- **Human:** Inconsistency between the system and the production floor
- **Process:** Too many feeding points → irregular distribution

Consequently, this structure results in unnecessary inventory and production disruptions.

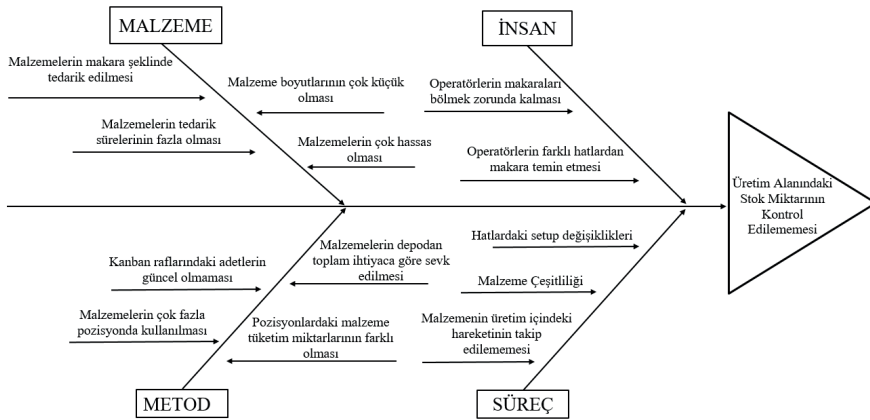


Figure 2. Fishbone Diagram

Data Collection Process

Production data, material usage records, and order information were collected through the factory's Manufacturing Execution System (MES). A Pareto analysis was applied to identify critical materials. According to the analysis, 24 out of 72 materials accounted for 80% of the total inventory. These 24 materials were selected as critical input data for the model.

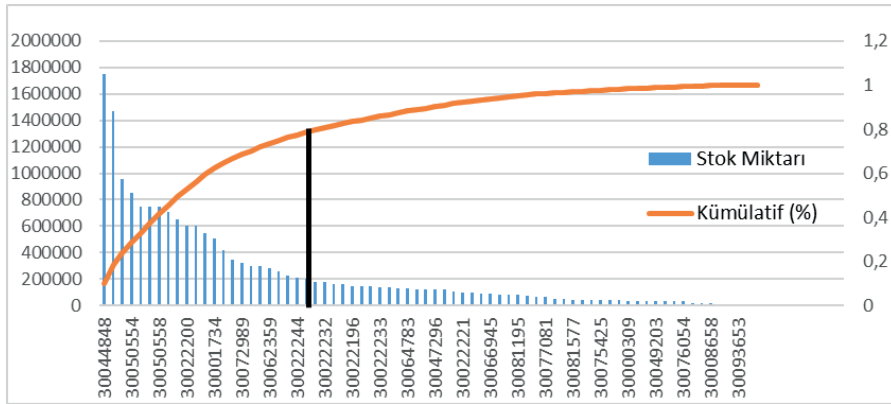


Figure 3. Distribution of the 24 most used materials according to the Pareto analysis

Artificial Neural Network Model

The modeling process was carried out in the MATLAB environment. A multilayer feedforward artificial neural network model was chosen, and historical production quantities were used as input variables. The output

was the predicted production quantity for the next 10 days. 70% of the data was used for training and 30% for testing. The performance of the model was evaluated based on MSE (Mean Squared Error) and R^2 criteria.

ERP Integration

The forecasting results were transferred to the ERP system, which was developed on a PostgreSQL basis, via a REST API. The production, inventory, and order modules within the ERP system read this data to provide real-time suggestions to managers. The system provides the purchasing team with future production forecasts based on historical data.

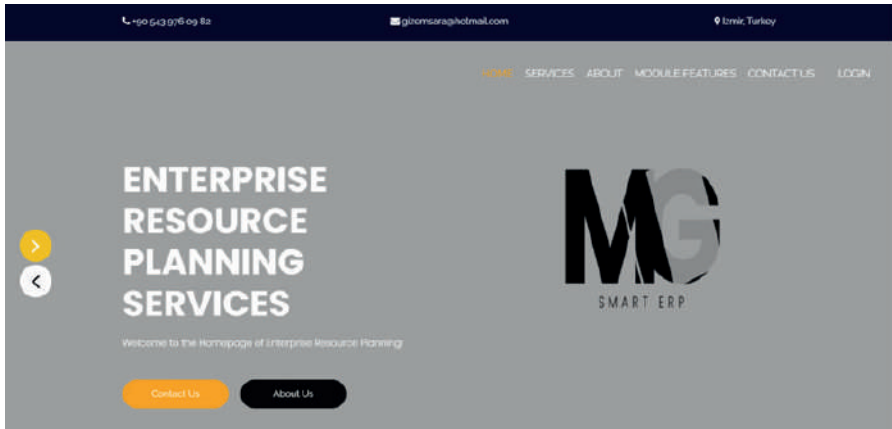


Figure 4. ERP Login Screen

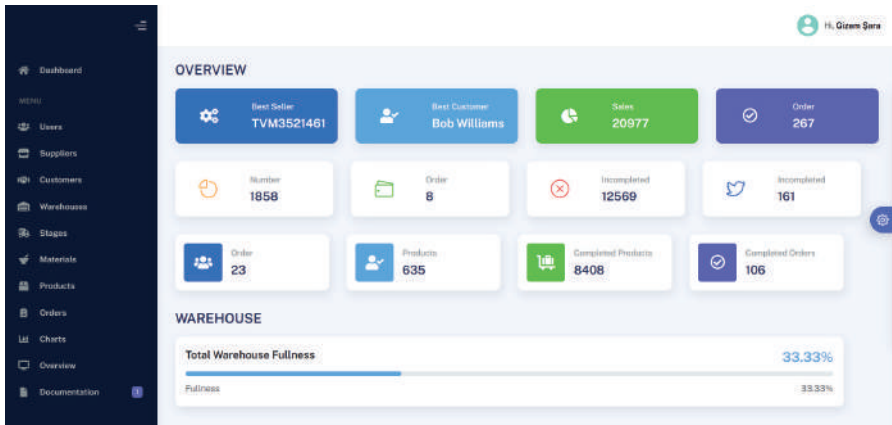


Figure 5. ERP Screen

RESULTS

This study aimed to address the problems encountered in the most critical areas of production processes such as production planning and inventory management through artificial intelligence and digital system integration. The artificial neural network model developed based on field observations and system data made highly accurate production forecasts using historical production data; these forecasts were integrated with the PostgreSQL-based ERP system, contributing to the digitalization of decision support processes.

As a result of the analyses, it was found that a large portion of production inventory consisted of a small number of critical materials, and it was understood that planning processes should be optimized around these materials. In this context, Pareto analysis and fishbone diagrams were effective both in identifying the problem and in generating solution strategies.

A high correlation coefficient was obtained for the artificial neural network model. This indicates that the model has a strong forecasting capacity and can make reliable predictions regarding production quantities. Through integration with the ERP system, these predictions were systematically transferred to the production, procurement, and inventory modules; thus, manual planning errors, time losses, and waste were significantly reduced.

Moreover, the developed system's ability to provide real-time data flow between ERP and MES not only enhances planning but also improves operational efficiency. This situation can provide transparency in the production process and allow decision-makers to intervene quickly in the field. Problems such as material surplus, misdirection, and operator confusion on the production line will be significantly reduced through this system.

This study also demonstrated that:

- Artificial intelligence models integrated into the production process provide not only efficiency but also strategic flexibility.
- The harmonious operation of ERP and MES systems is a cornerstone in the journey of digital transformation.
- With data-driven forecasting, businesses can optimize their resources not only based on the past but also for the future.

As a result, the proposed method of this study is a low-cost, sustainable digitalization example that can be easily applied in small and medium-sized manufacturing enterprises. The method is not merely a forecasting

algorithm but also a digital transformation strategy that shapes decision-making processes.

DISCUSSION AND CONCLUSION

The implementation of production forecasting using the artificial neural network model developed in this study, and the integration of the forecast results into the ERP system, provided accuracy, speed, and efficiency in production planning. The high R^2 value of the model represents a significant achievement compared to similar studies. For instance, El Madany et al. (2022) used time series analysis methods for supply chain forecasting but did not integrate model accuracy with ERP systems. From this perspective, the proposed model is innovative not only in terms of forecasting performance but also in system integration.

Studies based on deep learning models, such as those by Bengio (2009), also emphasize that neural networks perform excellently in nonlinear and complex systems. In this study, the dynamic structure of the production system was successfully modeled using an artificial neural network. However, the number of data points and the sampling period used for training the system were limited. This may affect the model's generalization capacity.

The findings supported by Pareto analysis and fishbone diagrams enabled the basic problems on the production floor to be identified within a broader framework. It was understood that the core problems encountered in material planning stem not only from technical issues but also from organizational and procedural deficiencies. In this regard, the socio-technical integration requirement frequently emphasized in literature supports the findings of this study.

In addition, the integration of predictions into the ERP system provided a significant advantage to managers in real-time decision-making processes. As emphasized by Ghadge et al. (2022) in the context of the green supply chain, digitalization positively affects not only production but also sustainability and resource efficiency.

In conclusion, this study offers a practical example of digital transformation through AI-assisted planning and ERP integration, while also contributing original insights to fill existing gaps in literature.

Future Work and Practical Implications

This study demonstrates the applicability of real-time data-driven forecasting systems integrated with ERP and MES in small and medium-sized enterprises (SMEs) in the manufacturing sector. Expanding the model

to different production lines or multi-product systems may further optimize resource planning. In addition, a fully integrated MES-ERP infrastructure can be extended to include other functional modules such as supply chain management, maintenance planning, and workforce allocation, thereby evolving into a holistic digital transformation system.

References

- Abacı SH, Tahtalı Y, Şekeroğlu A, 2020. Comparison of some different clustering methods in double dendrogram heat maps. 1st International Applied Statistics Conference, 1-4 October 2020, Page: 270, Tokat, Turkey.
- Aires, S., et al., "Requirements Elicitation in ERP Implementation Process," 2022.
- Alaskari, O., et al. "Framework for Implementation of Enterprise Resource Planning (ERP) Systems in Small and Medium Enterprises (SMEs): A Case Study." 2021.
- Arı A. ve Önder, H. (2013). Farklı Veri Yapılarında Kullanılabilecek Regresyon Yöntemler, *Anadolu Tarım Bilim. Derg.*,28(3):168-174
- Aydın, S. E. ve Küçükyaşar, M. (2017). "Bir Elektronik Fabrikasında Malzeme Sipariş Miktarlarının Optimizasyonu ve Kanban Uygulaması" (Lisans Tezi) Dokuz Eylül Üniversitesi.
- Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends® in Machine Learning*, 2(1), 1-127.
- Bircan, H. (2004). Lojistik Regresyon Analizi: Tıp Verileri Üzerine Bir Uygulama. *Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 2004 / 2: 185-208.
- Brandes P, Das D, 2006. Locating behaviour cynicism at work: Construct issues and performance implications. *Employee Health, Coping and Methodologies*. (Editors: Pamela L. Perrewe, Daniel C. Ganster), JAI Press, pp.233-266, New York.
- Brecher, C., et al. "Viable System Model for Manufacturing Execution Systems." 2013. MES, Endüstri 4.0.
- Can, M.B., Eren, Ç., Kuru, M., Özkan, Ö. ve Rzaeva, Z. (2012). "Veri Kümelerinden Bilgi Keşfi: Veri Madenciliği", *Başkent Üniversitesi Tıp Fakültesi XIV. Öğrenci Sempozyumu*, Ankara.
- Coşlu, E. (2013). "Veri madenciliği." Akademik bilişim.
- D'Antonio, G., et al. "A Novel Methodology to Integrate Manufacturing Execution Systems with the Lean Manufacturing Approach." 2017.
- Deniz, Ö. (2005). Poisson Regresyon Analizi. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi* Yıl:4 Sayı:7 Bahar 2005/1 S. 59-72.
- E. Kabalcı, Esnek Hesaplama Yöntemleri- II: Yapay Sinir Ağları. Jeoloji Mühendisliği ABD, ders notları.
- EDUCBA, Pyria Pedamkar (2018). "Introduction To Data Mining" <https://www.educba.com/data-mining-techniques-for-business/>
- Fayyad, U. M. (1996). Piatetsky-Shapiro, G., Smyth, P. and Uthurusamy, R. "Advances in Knowledge Discovery and Data Mining", USA: MIT Press.
- Freese, J. ve Long, J. S. (2006). *Regression Models for Categorical Dependent Variables Using Stata*. College Station: Stata Press.

- Ghadge, A., et al., “Link Between Industry 4.0 and Green Supply Chain Management: Evidence from the Automotive Industry,” 2022.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.
- Guliyev, N. J. ve Ismailov, V. E. (2016). A single hidden layer feedforward network with only one neuron in the hidden layer can approximate any univariate function. Ithaca: Cornell University Library, arXiv.org.
- Güçle, G. (2010). Veri ambarı ve veri madenciliği teknikleri kullanılarak öğrenci karar destek sistemi oluşturma. Yayınlanmamış yüksek lisans tezi. Denizli, Pamukkale Üniversitesi Fen Bilimleri Enstitüsü.
- Hinton, G. (2012). Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580.
- Holsheimer, M. ve Siebes, A. (1994). “Data Mining: The Search for Knowledge in Databases”, CWI Technical Report, Amsterdam, s. 2.
- Hung, S., Yen, D. C. ve Wang, H. (2005). “Applying Data Mining to Telecom Churn Management”, Expert Systems with Applications, s. 1-10.
- Hyndman, R. J. (2004). ‘The interaction between trend and seasonality’, International Journal of Forecasting, 20(4), 561–563.
- Jaskó, M., et al. “Development of Manufacturing Execution Systems in Accordance with Industry 4.0 Requirements: A Review of Standard- and Ontology-Based Methodologies and Tools.” 2020.

Acknowledgment

I would like to thank Prof. Dr. Mehmet Çakmakçı for his academic support and guidance throughout the execution and direction of this study. I am also grateful to the officials of the manufacturing company for facilitating field observations and data access.

Conflict of Interest

The authors declare that there is no conflict of interest regarding this study.

Author Contributions

Gizem Şara Onay:

Conceptual framework of the study, data collection, execution of analyses, development of the artificial neural network model, ERP integration, and manuscript writing.

Mehmet Çakmakçı:

Academic supervision, methodological guidance, evaluation of the modeling process, interpretation of results, and revision of the manuscript.

Using Statistical Moments in Hierarchical Machine Learning for Estimation of Birefringence in Mode-Locked Fiber Laser Systems

Hasan Arda Solak¹

Şeyma Koltuklu²

Sueda Turgut³

Mahmut Bağcı⁴

Abstract

Adaptive control and self-tuning of mode-locked fiber laser systems is an interesting topic in applied optics. Rapid and accurate detection of the cavity birefringence value is critical for the adaptive control and self-tuning of fiber laser systems. The birefringence varies randomly and significantly affects the mode-locking performance of fiber laser. In addition, the birefringence value in the laser cavity cannot be measured directly. In this study, from a new perspective, the birefringence value is determined (estimated) by hierarchical implementation of supervised machine learning algorithms. Unlike previous studies, instead of using the laser pulse energy directly, the energy evolution is recorded and a separate time series is obtained for each case of birefringence. The four statistical moments (mean, variance, skewness and kurtosis) of these time series are used as input variables (features) in the machine learning algorithm. When the findings of this study are compared

- 1 Marmara University, Faculty of Business Administration, Department of Management Information Systems, Türkiye
- 2 Marmara University, Faculty of Business Administration, Department of Management Information Systems, Türkiye
- 3 Marmara University, Faculty of Business Administration, Department of Management Information Systems, Türkiye
- 4 Marmara University, Faculty of Business Administration, Department of Management Information Systems, Türkiye

with the results of previous studies, it is seen that the birefringence value can be estimated with higher accuracy in a short time with the hierarchical approach. More accurate classification of birefringence increases efficiency of algorithms that enable adaptive control and self-tuning of mode-locked fiber laser systems. Consequently, the study contributes to the advancement of mode-locked fiber laser technology by enhancing performance in various industrial and scientific applications, enabling broader and more efficient use of laser systems.

INTRODUCTION

Mode locking is a phenomenon frequently observed in optical resonator cavities, where nonlinear interactions in the cavity synchronize different cavity modes to generate localized and stable light pulses (Bağcı and Kutz, 2020; Bağcı and Kutz, 2022). Mode-locked fiber lasers play a crucial role in many scientific and industrial applications due to their ability to generate ultra-short pulses with high peak powers. However, the performance and stability of these systems are highly sensitive to variations in cavity birefringence, a parameter that fluctuates randomly due to environmental factors such as temperature changes or mechanical vibrations (Brunton et al., 2014). Since the birefringence value cannot be directly measured during operation, maintaining optimal performance through adaptive control and self-tuning remains a significant challenge.

Since there is the possibility of making femtosecond laser pulses with gaining media in solids, liquids, and gases, most of the time, solid-state crystal, semiconductor, or fiber materials are preferred in the construction of lasers because of their practical use. For their optical gain, these systems generally work with ion-doped insulating crystals or glasses and pump in an optical way. As soon as the ions of the gain medium absorb energy from an optical pump, these ions excite to their higher energy level and spontaneously return to the lower energy level, photon emission occurs. These photons reflect countless times between the two cavity mirrors of the laser, during which amplification of the light wave occurs. When the optical gain becomes greater than the losses such as scattering or absorption, laser light results (Sennaroğlu, 2010; Sennaroğlu, 2007).

The shortest achievable pulse duration in a laser system depends on multiple factors, including the refractive index of the material used (via the Kerr effect), nonlinear physical interactions (e.g., polarization effects, absorption, scattering), the cavity length, and the reflectivity/transmission properties of the mirrors. Achieving a stable and robust mode-locking state is particularly difficult, as minor perturbations and environmental fluctuations

can easily disrupt the cavity dispersion and destabilize the mode-locking regime (Bağcı and Kutz, 2024). The randomness of fiber setups and the environmental sensitivity of laser parameters demand specialized techniques to ensure stable operation. These systems are often modeled using nonlinear Schrödinger-type equations to capture the complex interplay of dispersion and nonlinearity in fiber cavities (Kutz, 2006).

In this context, one of the most critical parameters affecting mode-locking stability is the birefringence in the laser cavity, which effectively represents cavity-induced loss. This birefringence changes based on environmental conditions such as position, temperature, and humidity. Variations in birefringence disrupt the stability of the laser system, necessitating the continuous adjustment of filter settings to ensure consistent performance. Hence, accurate detection of birefringence changes and automated tuning of the laser cavity parameters become essential tasks for maintaining stable mode-locked operation. In existing studies, the kurtosis of the laser output signal has been used to estimate birefringence (Fu et al., 2014; Kutz and Brunton, 2015; Bağcı and Kutz, 2024).

This study presents a novel method that uses supervised machine learning algorithms to improve adaptive control and self-tuning of mode-locked fiber lasers. The suggested approach uses mean, variance, and skewness as input features in addition to kurtosis to increase the accuracy of birefringence estimation, in contrast to earlier approaches that only use kurtosis. A rich dataset is created for machine learning model training by examining the pulse energy evolution over time and identifying these four statistical moments. The classification performance and estimation speed are greatly enhanced by this method, which enables high-precision, real-time control of the laser system.

To this end, a numerical model of the fiber laser cavity is first developed, and the resulting laser output is analyzed statistically. Before applying machine learning, data preprocessing—such as normalization and cleaning—is applied to enhance model performance. Then, supervised learning algorithms including Logistic Regression, Support Vector Machines, k-Nearest Neighbors, Decision Trees, Random Forests, and Naive Bayes, are then employed and compared in terms of accuracy, precision, recall and F1-score.

NUMERICAL METHODS

Theoretical Model

In this study, the propagation of the optical field in a fiber is modeled using a coupled nonlinear Schrödinger (CNLS) equation framework (Menyuk, 1987; Menyuk, 1989).

$$i \frac{\partial u}{\partial z} + \frac{D}{2} \frac{\partial^2 u}{\partial t^2} - Ku + \left(|u|^2 + A|v|^2 \right) u + Bv^2 u^* = iRu,$$

$$i \frac{\partial v}{\partial z} + \frac{D}{2} \frac{\partial^2 v}{\partial t^2} - Kv + \left(A|u|^2 + |v|^2 \right) v + Bu^2 v^* = iRv.$$

The model describes two orthogonally polarized electric field envelopes $u(z, t)$ and $v(z, t)$ in an optical fiber. Here, t is the normalized time scaled to the pulse's full-width half-maximum, and z is the normalized propagation distance with respect to the cavity length. The components u and v represent the fast and slow polarization components, respectively (Brunton et al., 2014).

The parameter K denotes the birefringence of the cavity, while D represents the average group velocity dispersion. The parameters A and B are nonlinear coupling coefficients corresponding to cross-phase modulation and four-wave mixing, respectively. These values are determined by the physical properties of the fiber, and satisfy the condition $A + B = 1$. For a silica-based optical fiber, typical values are $A = 2/3$ and $B = 1/3$.

The terms Ru and Rv represent the gain-loss operator, incorporating the effects of Ytterbium-doped amplification, saturable gain, and bandwidth-limited amplification. The operator R is defined as:

$$R = \frac{2g_0(1 + \tau\partial_t^2)}{1 + (1/e_0) \int_{-\infty}^{\infty} (|u|^2 + |v|^2) dt} - \Gamma$$

where g_0 is the dimensionless pumping strength, and e_0 is the dimensionless saturation energy of the gain medium. The parameter τ represents the gain bandwidth, while Γ accounts for losses due to output coupling and fiber attenuation.

After each cavity round-trip, the application of wave plates and a passive polarizer is modeled via the discrete application of Jones matrices (Ding and Kutz, 2009; Komarov et al., 2005; Jones, 1941). The optical elements are defined as follows:

Quarter-wave plate:

$$W_{\lambda/4} = \begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix},$$

Half-wave plate:

$$W_{\lambda/2} = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix},$$

Polarizer:

$$W_p = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Here, $W_{\lambda/4}$ represents the quarter-wave plate (with orientation angles α_1 and α_2), $W_{\lambda/2}$ represents the half-wave plate (with angle α_3), and W_p is the polarizer (with angle α_p). If the principal axes of these elements are not aligned with the fast axis of the cavity, a rotation matrix is applied:

$$J_j = R(\alpha_j)W_jR(-\alpha_j), R(\alpha_j) = \begin{bmatrix} \cos(\alpha_j) & -\sin(\alpha_j) \\ \sin(\alpha_j) & \cos(\alpha_j) \end{bmatrix}$$

where α_j represents the orientation angle of each wave plate or polarizer ($j = 1, 2, 3, p$). These rotation angles are considered as control variables in the model and are essential for achieving mode-locking solutions. Recent experimental results have shown that these control variables can be manipulated electronically with ease, making them highly suitable for adaptive laser control systems (Shen et al., 2012).

Feature Descriptions

Input Variables

The first moment, mean, represents the average value of the laser's output power (or energy) over multiple cycles. It is used to measure the overall performance of the laser. To calculate the mean, all output energy values are summed and this total is divided by the number of measurements. Mathematically, it is expressed as:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Here, \bar{x} represents the mean output energy, x_i denotes each measurement value, and n is the number of measurements.

The variance shows the spread of output energy values around the mean. A higher variance indicates that the measurements deviate more from the mean, and thus the laser output is more unstable. The variance (s^2) is calculated as:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{(N-1)}$$

Skewness indicates whether the distribution is symmetric and whether it leans to the right or left. Positive skewness suggests a longer tail on the right side, while negative skewness indicates a longer tail on the left side. Skewness (Sk) is calculated as:

$$Sk = \frac{\sum_{i=1}^N (x_i - \bar{x})^3}{(N-1)s^3}$$

Kurtosis measures the extremity or the tailedness of the distribution. High kurtosis indicates more outliers in the distribution, showing greater deviations from the mean. Kurtosis () is calculated as:

$$= \frac{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4 \right)}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2}$$

The use of these statistical moments summarizes the dynamic behavior of the laser, offering strong and distinctive features for estimating the birefringence value (Fu et al., 2014; Kutz and Brunton, 2015).

Exploratory Data Analysis of Input Features

For generation of the sample dataset the parameters in the CNLS model are specified as follows:

D	A	B	g_0	e_0	τ	Γ
-0.4	2/3	1/3	1.73	4.23	0.1	0.1

Also, the birefringence K parameter is iterated between -0.32 and 0.32 with 0.01 step size, while the polarizer angle (α_p) is varying from -90° to 90° with 1° of step size. Thus, the number of instances of training sample dataset is $65 \times 181 = 11765$. To test the machine learning models, we generated a separate mis-aligned (noised) dataset in which we have 1000 instances with random K and α_p values.

To gain a better understanding of the dataset's underlying structure, the distributions of all input features were examined before the model was trained. Every input variable displays a unique statistical pattern, as shown in **Figure 1**.

The intricacy and non-linearity of the birefringence estimation task are revealed by these visualizations' features. Additionally, they defend the use of tree-based models, like Random Forests, which don't require rigid parametric assumptions and are ideal for handling skewed and non-normally distributed data.

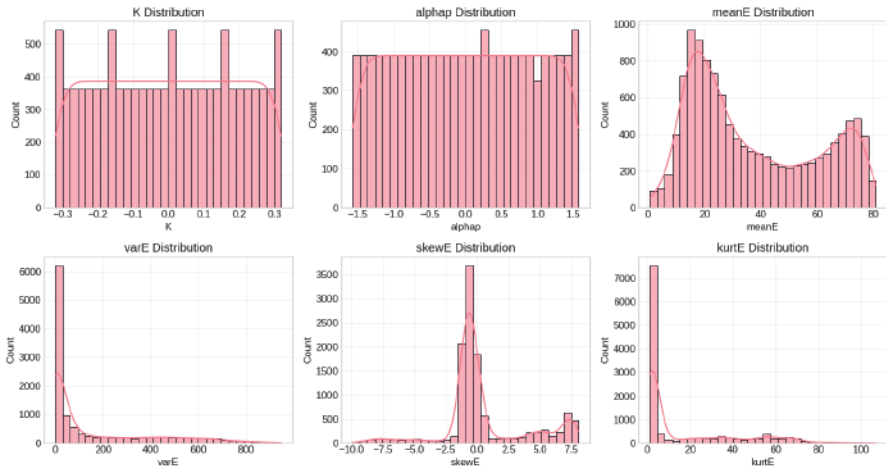


Figure 1. Histogram of all input features used in the model. The figure highlights the statistical characteristics and potential skewness of each feature distribution.

Supervised Machine Learning Methods

In recent years, machine learning algorithms and artificial intelligence techniques have been extensively employed to enhance the functionality and performance of mode-locked fiber lasers (Bağcı and Kutz, 2024). In particular, Artificial Neural Networks (ANNs) have been utilized to improve system architecture, operational efficiency, and control mechanisms. These

methods have also demonstrated notable success in areas such as signal processing, system modeling, channel equalization, and the efficient control of optical systems (Freire et al., 2023; Ma et al., 2022; Mezzi et al., 2023; Boscolo and Finot, 2020). Additionally, Genetic Algorithms (GAs) have been adopted to fully automate the startup procedures of laser systems, thereby contributing to enhanced reliability and operational efficiency (Ma et al., 2022; Han et al., 2024; Woodward and Kelleher, 2016).

In the present study, six supervised machine learning algorithms are employed, including Logistic Regression (LR), Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Decision Trees (DT), Random Forests (RF), and Naive Bayes (NB).

Logistic Regression (LR) is a widely used method for binary classification tasks. This supervised learning technique models a binary outcome variable using an S-shaped logistic (sigmoid) function (Tolles, 2016). Through this function, LR defines decision boundaries between classes, which may range from simple linear to complex nonlinear structures depending on the nature of the classification problem (Gudivada et al., 2016; DeMaris, 1995).

Support Vector Machines (SVM) are supervised algorithms applicable to both classification and regression problems. This method separates data points into different classes using a hyperplane that maximizes the margin between classes (Vapnik and Cortes, 1995; Burges, 1998). Data points closest to the hyperplane, known as support vectors, are critical to defining the decision boundary. In the context of mode-locked fiber lasers, SVMs have been effectively used to analyze cavity parameter spaces essential for achieving stable mode-locking.

K-Nearest Neighbors (KNN) is a non-parametric, instance-based learning algorithm that can be applied to both classification and regression problems. One of the key advantages of KNN is the absence of a training phase; instead, it assigns class labels based on similarity metrics—such as Euclidean or Manhattan distance—to the nearest neighbors in the feature space (Cover, 1967; Hall et al., 2008; Keller et al., 1985). Although computationally intensive, KNN is particularly effective in scenarios where data are balanced and class boundaries are not well defined.

Decision Trees (DT) represent a non-parametric supervised learning approach that uses a hierarchical tree structure to model decisions based on input features (Twa et al., 2005; Rokach, 2014; Brodley and Utgoff, 1995). The algorithm sequentially partitions the input space and assigns class labels based on impurity measures such as Gini index or entropy (Quinlan, 1986).

DTs are especially effective in large and complex datasets for uncovering meaningful patterns and identifying influential features (Myles et al., 2004).

Random Forest (RF) is an ensemble-based supervised learning algorithm designed for both classification and regression tasks. It consists of multiple decision trees constructed with random subsets of the training data and features. The ensemble nature of RF helps mitigate overfitting and improves generalization performance by averaging predictions from multiple models (Breiman, 2001; Schonlau and Zou, 2020; Ho, 1995). RF is known for its robustness and scalability, making it well-suited for high-dimensional and large-scale datasets (Paul et al., 2018).

Naive Bayes (NB) is a probabilistic classification algorithm grounded in Bayes' theorem. It operates under the assumption of conditional independence among input features (Bernardo and Smith, 2001). NB computes the posterior probabilities of each class and assigns the class label with the highest probability using maximum likelihood estimation (Hastie et al., 2001). Despite its simplicity and strong independence assumption, Naive Bayes remains a highly effective model, particularly in large-scale and high-dimensional problems due to its computational efficiency.

Model Training and Validation

Model Training

The selected supervised machine learning algorithms were trained on the training dataset to estimate the birefringence value. The training process involves the model learning the relationships between the input variables (statistical moments) and the output variable (birefringence value). During this stage, the model's parameters are optimized to minimize the error on the training data.

Grid Search-Based Hyperparameter Optimization

To ensure optimal model performance in birefringence estimation, a systematic hyperparameter tuning process was applied to the Random Forest algorithm. A grid search strategy combined with five-fold cross-validation was employed to explore a wide range of parameter combinations. This procedure involved evaluating 432 distinct configurations across multiple decision trees, resulting in a total of 2160 model fits.

The hyperparameters considered in this optimization included the number of trees in the forest, maximum tree depth, the minimum number of samples required to split an internal node, the minimum number of samples required to be at a leaf node, the method used to select the number

of features at each split, and class weight adjustments to handle potential class imbalances.

Following this exhaustive search, the configuration that achieved the highest cross-validated accuracy was selected as the optimal model setup. This final set of parameters enhanced the model's ability to generalize across different birefringence levels and contributed significantly to the overall robustness and precision of the classification results.

Metrics for Performance Measurement

The performance of the trained model was evaluated on the validation dataset using various metrics. These metrics provide comprehensive information about the accuracy and reliability of the model's predictions.

Accuracy: The ratio of correctly predicted instances to the total number of instances. It is a general indicator of overall performance and can be sufficient on its own, especially for balanced datasets.

Precision: Indicates how many of the instances predicted as positive by the model are actually positive. It is important for minimizing false positives. For instance, in critical laser systems, the cost of a false positive alarm can be high.

Recall: Shows how many of the actual positive instances were correctly identified as positive by the model. It is important for minimizing false negatives. This metric is crucial because the failure to correctly identify the birefringence value can lead to unstable laser system operation.

F1-score: The harmonic mean of precision and recall. It provides a more informative performance measure than accuracy alone, especially in unbalanced datasets.

RESULTS

Running multiple classification models on the birefringence dataset yielded insightful comparative results regarding their segmentation capability. As illustrated in **Figure 2**, all models performed above a moderate threshold; however, their success varied significantly. Among the tested algorithms, Random Forest and Decision Tree achieved the highest accuracy—approaching 0.99 across all evaluated metrics, including precision, recall, and F1-score. In contrast, traditional methods such as Naive Bayes and Logistic Regression showed noticeably lower performance, with accuracies around 0.55 and 0.53, respectively. This clear distinction highlights the superior ability of tree-based ensemble methods in capturing the nonlinear and statistical characteristics of birefringence dynamics.

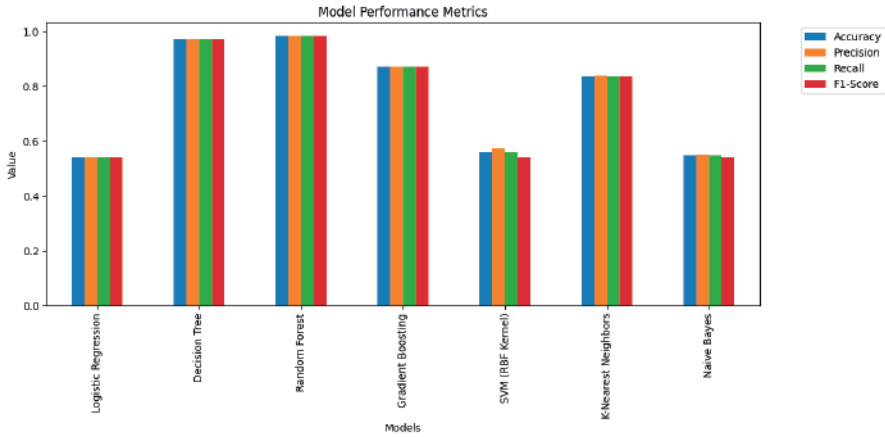


Figure 2. Comparison of classification performance metrics (Accuracy, Precision, Recall, F1-Score) across supervised machine learning models used for birefringence level prediction.

Following the comparative evaluation, the Random Forest classifier was selected for hierarchical classification due to its balance of interpretability, scalability, and strong empirical performance. To assess its robustness under different iterations, the continuous K values are discretized (classified) into 2 to 32 levels using quantile-based binning.

Figure 3 presents the accuracy trends across the number of iterations (hierarchically). The model achieved 98.9% accuracy for (2^1)-class (or two-class) classification, and despite a gradual decline in performance with increasing iteration number, the model retained strong accuracy (88.2%) even at (2^5)-class (or 32-class) classification. The decline is expected due to the increased complexity and reduced separation between classes in higher-resolution binning.

Furthermore, a summary of performance metrics is shown in **Table 1**, where precision, recall, and F1-score values remain highly consistent with overall accuracy. All scores remained above 0.88, validating the model's reliability across both coarse and fine classifications.

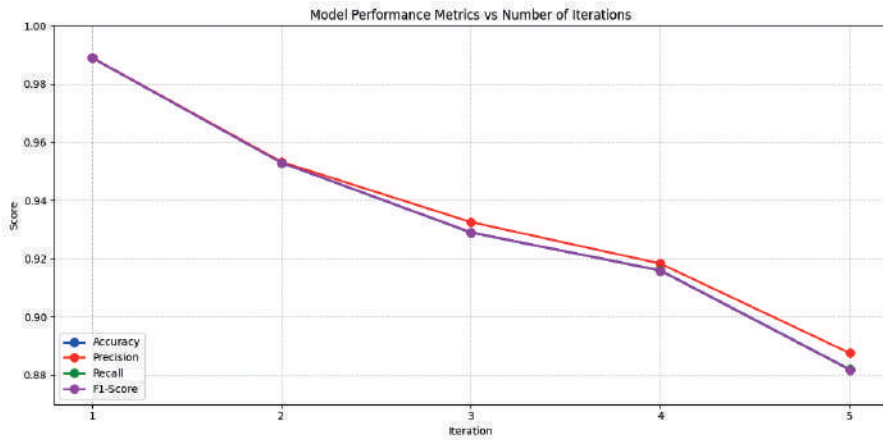


Figure 3. Accuracy, precision, recall, and F1-score trends of the Random Forest model across different birefringence classification levels (number of iterations).

Table 1. Performance metrics (accuracy, precision, recall, and F1-score) of the Random Forest classifier for different iterations.

Iteration	Performance Metrics			
	Accuracy	Precision	Recall	F1-Score
1	0.9890	0.9890	0.9890	0.9890
2	0.9530	0.9532	0.9530	0.9530
3	0.9290	0.9326	0.9290	0.9290
4	0.9160	0.9183	0.9160	0.9159
5	0.8820	0.8876	0.8820	0.8819

To better understand the internal decision-making process of the Random Forest classifier, a feature importance analysis is performed. As illustrated in **Figure 4**, the polarizer angle (α_p) emerges as the most significant contributor to model performance, followed by higher-order statistical moments such as skewness and kurtosis.

DISCUSSION AND CONCLUSION

Using polarizer angle and statistical moments of the energy evolution as inputs, the study’s findings showed how well a hierarchical machine learning technique estimates birefringence in mode-locked fiber laser systems. With accuracy values above 92% in low-resolution class definitions, Random Forest performed noticeably better than the other classifiers that were

evaluated. This lends credence to the claim that complex and nonlinear optical dynamics were best handled by ensemble-based decision models.

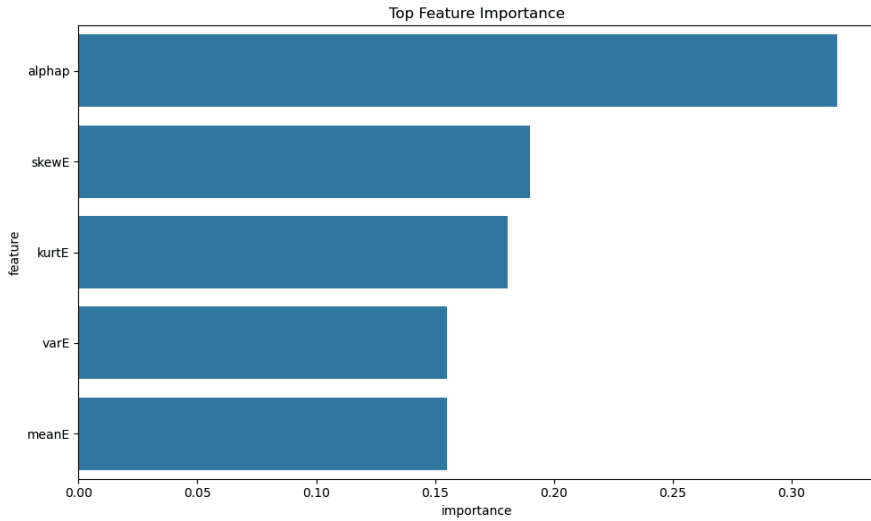


Figure 4. Relative importance of input features in the Random Forest model for birefringence classification. The polarizer angle (α_p) dominates in importance, followed by skewness (skewE), kurtosis (kurtE), variance (varE), and mean (meanE).

While a previous study employing sparse representation techniques under misalignment reported an accuracy of 88% (Fu et al., 2014), our method achieved a comparable level of accuracy while offering greater robustness in multi-level classification scenarios. Moreover, the present study eliminates the need for alignment-dependent spectrogram inputs by directly utilizing time-domain statistical features. This methodological distinction provides advantages in terms of both computational efficiency and implementation flexibility for real-world laser control applications.

The consistent performance across all levels of classification indicates the reliability of moment-based features in capturing essential variations in the laser cavity. In particular, skewness and kurtosis—higher-order statistical measures—proved highly discriminative, confirming that birefringence states influence the asymmetry and tailedness of the energy evolution data. The use of the polarizer angle (α_p) as an input also adds to the model's predictive power while preserving its independence from physical modeling constraints.

In practical terms, this work aids in the creation of adaptive laser control real-time birefringence recognition systems. Accurate and quick birefringence detection can increase laser systems' self-tuning capabilities, shorten maintenance schedules, and boost long-term operational effectiveness. Furthermore, the model's ability to scale to higher resolution levels raises the possibility of its incorporation into more intricate fiber laser networks or multi-NPR (nonlinear polarization rotation) setups.

In conclusion, a very precise, understandable, and ready-to-implement solution for birefringence classification in fiber laser systems was provided by the hierarchical Random Forest classifier trained on moment-based features. Future research might examine hybrid models that combine spectral and statistical features for even greater precision, real-time deployment, or transfer learning across various laser architectures.

References

- Bağcı M, Kutz JN, 2020. Spatiotemporal mode locking in quadratic nonlinear media. *Phys. Rev. E* 102(9), 022205.
- Bağcı M, Kutz JN, 2022. Mode-locking in quadratically nonlinear waveguide arrays. *Opt. Express* 30, 28454-28469.
- Bağcı M, Kutz N, 2024. Machine learning for self-tuning mode-locked lasers with multiple transmission filters. *Optical Society*, 79–89.
- Bernardo JM, Smith AFM, 2001. *Bayesian theory*. Measurement Science and Technology, vol. 12, pp. 221–222.
- Boscolo S, Finot C, 2020. Artificial neural networks for nonlinear pulse shaping in optical fibers. *Optics & Laser Technology*, 131: 106446.
- Breiman L, 2001. Random forests. *Machine Learning*, 4: 5–32.
- Brodley C, Utgoff P, 1995. Multivariate decision trees. *Machine Learning*, 19: 45–77.
- Burges CJ, 1998. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2: 121–167.
- Cover PHT, 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(11): 21–27.
- DeMaris A, 1995. A Tutorial in Logistic Regression. *Journal of Marriage and Family*, 57(14): 956–968.
- Ding E, Kutz JN, 2009. Operating regimes, split-step modeling, and the Haus master mode-locking model. *Journal of the Optical Society of America B*, 26(11): 2290–2300.
- Freire P, Manuylovich E, Prilepsky JE, Turitsyn SK, 2023. Artificial neural networks for photonic applications—from algorithms to implementation: tutorial. *Advances in Optics and Photonics*, 15: 739–834.
- Fu X, Brunton SL, Kutz JN, 2014. Classification of birefringence in mode-locked fiber lasers using machine learning and sparse representation. *Optics Express*, 22(7): 8585–8597.
- Gudivada VN, Irfan MT, Fathi E, Rao DL, 2016. Cognitive analytics: Going beyond big data analytics and machine learning. In: Bhatnagar R, Srinivasan K (Editors), *Cognitive Computing: Theory and Applications*. Elsevier, pp. 169–205, Amsterdam – Netherlands.
- Han D, Guo R, Li G, Chen Y, Zhang B, Ren K, Zheng Y, Zhu L, Li T, Hui Z, 2024. Automatic mode-locked fiber laser based on adaptive genetic algorithm. *Optical Fiber Technology*, 83.
- Hall P, Park BU, Samworth RJ, 2008. Choice of neighbor order in nearest-neighbor classification. *The Annals of Statistics*, 36(15): 2135–2152.
- Hastie T, Tibshirani R, Friedman JH, 2001. *The Elements of Statistical Learning: Data Mining Inference and Prediction*. Springer, New York – USA.

- Jones RC, 1941. A new calculus for the treatment of optical systems. I. Description and discussion of the calculus. *Journal of the Optical Society of America*, 31(7): 488–493.
- Keller JM, Gray MR, Givens JA, 1985. A fuzzy k-nearest neighbor algorithm. *IEEE Transactions on Systems, Man, and Cybernetics*, 14: 580–585.
- Komarov A, Leblond H, Sanchez F, 2005. Multistability and hysteresis phenomena in passively mode-locked fiber lasers. *Physical Review A*, 71: 053809.
- Kutz JN, 2006. Mode-locked soliton lasers. *SIAM Review*, 48(4): 629–678.
- Ma X, Lin J, Dai C, Lv J, Yao P, Xu L, Gu C, 2022. Machine learning method for calculating mode-locking performance of linear cavity fiber lasers. *Optics & Laser Technology*, 149.
- Ma X, Lv J, Luo J, Liu X, Yao P, Xu L, 2023. Pulse convergence analysis and pulse information calculation of NOLM fiber mode-locked lasers based on machine learning method. *Optics & Laser Technology*, 163.
- Menyuk CR, 1987. Nonlinear pulse propagation in birefringent optical fibers. *IEEE Journal of Quantum Electronics*, 23(2): 174–176.
- Menyuk CR, 1989. Pulse propagation in an elliptically birefringent Kerr medium. *IEEE Journal of Quantum Electronics*, 25(12): 2674–2682.
- Mezzi R, Bahloul F, Karar AS, Ghandour R, Salhi M, 2023. Predicting behavior of photonic crystal fiber lasers using artificial neural networks. *Optics Communications*, 542.
- Myles AJ, Feudale RN, Liu Y, Woody NA, Brown SD, 2004. An introduction to decision tree modeling. *Journal of Chemometrics*, 18(16): 275–285.
- Paul A, Mukherjee DP, Das P, Gangopadhyay A, Chintla AR, Kundu S, 2018. Improved random forest for classification. *IEEE Transactions on Image Processing*, 27(18): 4012–4024.
- Rokach L, 2014. *Data Mining with Decision Trees*. World Scientific Pub. Co. Inc., Singapore.
- Quinlan R, 1986. Induction of decision trees. *Machine Learning*, 1: 81–106.
- Schonlau M, Zou RY, 2020. The random forest algorithm for statistical learning. *The Stata Journal*, 20(11): 3–29.
- Sennaroğlu A, 2007. Fotonik ve katıhal lazerleri. *Bilim ve Teknik*, 40–46.
- Sennaroğlu A, 2010. Katıhal femtosaniye lazerleri. *Bilim ve Teknik*, 48–53.
- Shen X, Li W, Yan M, Zeng H, 2012. Electronic control of nonlinear-polarization rotation mode locking in Yb-doped fiber lasers. *Optics Letters*, 37: 3426–3428.
- Tolles WJM, 2016. Logistic regression: Relating patient characteristics to outcomes. *Journal of the American Medical Association*, 316(15): 533–534.

- Twa MD, Parthasarathy S, Roberts CJ, Mahmoud AM, Raasch TW, Bullimore MA, 2005. Automated decision tree classification of corneal shape. *Optometry and Vision Science*, 82: 1038–1046.
- Vapnik V, Cortes C, 1995. Support-vector networks. *Machine Learning*, 20: 273–297.
- Woodward RI, Kelleher EJR, 2016. Towards ‘smart lasers’: Self-optimisation of an ultrafast pulse source using a genetic algorithm. *Scientific Reports*, 6: 37616.

Acknowledgment

This research was supported by the Scientific and Technological Research Council of Türkiye (TÜBİTAK) under the 2209-A University Students Research Projects Support Program, with the grant number 1919B012414330.

Conflict of Interest

The authors have declared that there is no conflict of interest.

Examination of Supervised Machine Learning Algorithms in Employee Turnover Prediction

Melisa Dikici¹

Gökçe Sabriye Hörük²

Deniz Efendioğlu³

Abstract

This study classifies employee layoffs into two types: voluntary and involuntary. The financial effects of voluntary turnover are highlighted in organizations. The ability to predict redundancies is a very important aspect of employee retention strategies for any organization. Supervised machine learning algorithms, which are more accurate, were analyzed in the context of employee turnover prediction. The literature review identified some of the key parameters that influence turnover, which include working conditions, job satisfaction, management support, pay fairness, and career opportunities. The algorithms to be studied will include Logistic Regression, Decision Trees, Random Forests, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Naive Bayes. Random Forest showed the best accuracy; hence, it is recommended for complicated datasets. Logistic Regression, though less accurate, is simple and interpretable, hence useful for strategic decision-making. This study henceforth highlights that the choice of algorithms should be fully aligned with data structure and organizational priorities for better human resource management and reduction in turnover ratio.

- 1 Ankara Yıldırım Beyazıt University, Faculty of Engineering & Natural Science, Industrial Engineering Department, 06010, Ankara, Turkey
- 2 Ankara Yıldırım Beyazıt University, Faculty of Engineering & Natural Science, Industrial Engineering Department, 06010, Ankara, Turkey
- 3 Ankara Yıldırım Beyazıt University, Faculty of Engineering & Natural Science, Industrial Engineering Department, 06010, Ankara, Turkey

INTRODUCTION

The turnover of employees happens in two ways: through voluntary and involuntary turnover. Voluntary turnover takes place when the decision to leave the job comes from the employee himself, that is, resignation. Involuntary turnover takes place when the employer decides on termination of the employment (Hong et al., 2007).

The factors that constitute voluntary turnover are employees looking for better or more lucrative positions for their own career advancement. Employees dissatisfied with their job or unable to get along with managers or colleagues, and/or in stressful working conditions, lead the decisions to leave voluntarily. Furthermore, personal reasons such as relocating, health issues, and responsibilities towards family may drive a person to leave employment. Long working hours, or at times inability or difficulty in maintaining the so-called ‘work-life’ balance, may also be some reasons for voluntary quits (Russell et al., 2013).

On the other hand, the causes of involuntary turnover are mainly associated with the decisions of employers. Factors such as failure to meet expectations of performance or other disciplinary actions commonly lead to an employer’s decision for employee termination. Furthermore, organizational changes like downsizing, mergers, and cost-cutting policies are some of the major reasons for involuntary turnover. Economic fluctuations or the development of technologies that make certain jobs redundant also fall into this category of turnover (Leana et al., 1987).

It might lead to unforeseen consequences for companies, especially when skilled employees leave the company. It could result in high costs because of workforce disruptions, recruitment efforts, and training processes. Since involuntary turnover is initiated by employers themselves, predictions are not required. However, in the case of voluntary turnover, which occurs at the discretion of employees, companies try to estimate which of their employees are likely to quit. These estimates are important for workforce planning and in devising strategies for better employee engagement and retention (Setiawan et al., 2020).

Some of the machine learning algorithms are used for predicting employee turnover. Whether unsupervised or supervised, machine learning techniques are increasingly being used to predict turnover and help inform retention strategies. Which of these algorithms is used will depend upon the size and nature of the data set.

Supervised learning methods have been widely used for the prediction of employee turnover, either because they can handle structured data or due to their better performance in terms of predictive accuracy. Decision trees, random forests, gradient boosting, logistic regression, support vector machines, neural networks, and many other techniques have been greatly tested for this purpose (Zhao et al., 2018) (Punnoose et al., 2016) (Jhaver et al., 2019). Some of the benefits are that several supervised methods have demonstrated high accuracy in predicting turnover even with noisy HR data, hence finding them quite effective for practical applications (Punnoose et al., 2016) (Jhaver et al., 2019). These methods have also been tested across different organizational sizes and complexities, hence providing guidelines for their effective use under diverse HR scenarios (Zhao et al., 2018).

Unsupervised methods, like clustering, are not applied very often to turnover predictions but may provide specific knowledge regarding patterns and groupings of data on employees. An example could be K-Means clustering combined with PCA: this was also tried without giving results that were notably good compared to supervised methods (Birzniece et al., 2022). The unsupervised nature of K-Means has shown limited success regarding the accuracy of the forecast, and it usually needs complementary techniques to enhance their utility (Birzniece et al., 2022).

The following part is about the supervised machine learning algorithms and the role of each of the supervised machine learning algorithms in predicting employee turnover. The article delves into the factors influencing employees' decisions to leave their jobs, offering detailed explanations of these factors. A comprehensive literature review was conducted, analyzing the machine learning methods employed in prior studies. The article provides in-depth insights and analyses of these techniques. The findings aim to guide organizations by highlighting the advantages of supervised machine learning techniques applicable to predicting employee turnover and their impact on the accuracy of such predictions.

In this prediction, some employee turnover parameters and reasons were considered. Many reasons can be present to cause employees to leave their jobs. Some of them are as follows:

Working conditions significantly impact employees' well-being, including their physical and psychological environment. Factors like office ergonomics, safety, and hygiene deficiencies can lead to dissatisfaction and eventual turnover (Salleh, Nair, & Harun, 2012). Job satisfaction, or the lack thereof, is another critical factor; dissatisfaction in one's role is often a precursor to an employee leaving the organization (Arshad & Puteh, 2015). Inadequate

training, which hinders employees from developing necessary skills for their roles, further exacerbates this issue (Nketsiah & Nkansah, 2024). A lack of management support negatively impacts employee motivation and engagement, reducing their commitment to the organization (Ribes, Touahri, & Perthame, 2017). Pay fairness, or the perception of receiving lower-than-market-standard salaries, also plays a crucial role in turnover decisions (Effendi, 2024). Work-life imbalance, characterized by excessive workload or inflexible hours, often leads to burnout and dissatisfaction among employees (Ongori, 2007). Similarly, inequitable workload distribution can exacerbate burnout and stress, pushing employees toward resignation (Zhang, Cao, Qu, & Wang, 2024). Organizational commitment is another significant determinant; when employees lose alignment with the organization's values and goals, turnover becomes more likely (Nketsiah & Nkansah, 2024). A lack of career advancement opportunities within the organization further fuels the intention to leave, especially when external opportunities offer better prospects (Hong, Wei, & Chen, 2007). Leadership style mismatches and workplace conflicts, including poor team dynamics, further contribute to dissatisfaction and disengagement (Brown et al., 2017). The infectious effect, wherein employees are influenced by others leaving, can also create a ripple effect, especially in closely-knit teams (Teng et al., 2019). Technological changes that overwhelm employees or the lack of personal development opportunities can make them feel unsupported, further affecting their retention (Yin, Hu, & Chen, 2024). Job security concerns and feelings of social isolation at the workplace are also critical contributors to employee dissatisfaction and eventual turnover (Recilla et al., 2024). Anomalies in project management and organizational processes, along with mismatched business culture, can cause employees to feel disconnected from the organization (Ajit, 2016). Finally, perceptions of injustice regarding organizational decisions and practices can undermine employee trust and commitment, prompting them to seek opportunities elsewhere (Nketsiah & Nkansah, 2024).

MATERIAL AND METHODS

Machine Learning-Supervised Learning

1. Logistic Regression

Since logistic regression can provide firms with an understanding of the factors that affect employee retention, it is a very crucial research area in predicting employee turnover. Logistic regression finds its fit in modeling binary outcomes-whether an employee will stay or leave. Some key elements of this predictive strategy:

Salary: A business should differentiate itself from the competitors with regards to pay. According to studies, pay plays a significant role in influencing the dismissal rate, with those workers who earn better income showing a lower propensity to leave their jobs. (Effendi, 2024). **Promotion:** career development is made visible to employees of various business units who have opportunities for promotion; this displays different tendencies from the point of dismissal. (Effendi, 2024)(Chen, 2023).

Demographic Factors: These factors have a bearing on commitment and job satisfaction; therefore, age, marital status, and length of service are the key predictors of turnover. (Chen, 2023). Logistic regression was proven to be very powerful for predicting turnover and reaching F1-scores and accuracies that were very high in several works. (Taner et al., 2024). Contrary, it often falls back behind the top models on accuracy, namely Random Forest and Support Vector Machines. (Zhang et al., 2024) (Liao, 2023). Despite the advantages of logistic regression, there are some of its disadvantages that should be considered, such as possible class imbalance in turnover data, which may affect the performance of the model. Future studies should investigate hybrid models incorporating logistic regression with other machine learning methods to improve predicted accuracy. Logistic regression, though simple to predict turnover, usually performs worse in comparison to other more sophisticated machine learning models. Accordingly, the studies indicate that such models as Random Forests and Support Vector Machines result in a higher degree of accuracy in predictions of turnover. Nevertheless, logistic regression is still applicable because it is easy to work with and can clearly show which factors are driving certain phenomena. (Effendi, 2024) (Chen, 2023) (Krishna et al., 2023) In the end, logistic regression is helpful in predicting employee turnover, whereby one gets to understand those factors that influence an employee to leave. While it may not always match the predictive power of more complex models, the simplicity and interpretability make it a very valuable tool for HR analytics.

2. Decision Trees

Decision trees, combined with a variety of algorithms and datasets, have produced some promising results in predicting employee turnover. Certain studies were able to provide evidence that decision tree models, like the RandomTree algorithm, have reached high accuracy in the classification of turnover intention, which includes job satisfaction and organizational commitment as important predictors. (Živković et al., 2024) Decision trees are also useful in classifying employees into different categories of attrition,

using demographic data and employment history, hence enabling predictions about the future cases of turnover. (Ahmed et al., 2023)

Various studies have replicated the identification of critical factors contributing to employee turnover using the Decision Tree analysis. Some commonly identified predictors are job satisfaction, opportunities for personal growth, affective organizational commitment, salary satisfaction, and interpersonal relationships. . (Živković et al., 2024) Other important factors were monthly income, overtime, age, distance from home, and years in the company. (Gao et al., 2019) Such information has helped the organization understand the root cause for turnover and come up with effective retention policies. It has been contrasted with other well-known machine learning ensembles, including Support Vector Machines, Random Forests, and Gradient Boosting Trees. Despite being straightforward and easy to understand, the model that a decision tree produces performs somewhat worse than ensemble approaches that use gradient-based decision trees, which are more effective with structured and unbalanced data and offer scalable solutions for strategic business analytics (Zhang et al., 2024). Nevertheless, because of their ease of use and ability to highlight the primary causes of turnover, decision trees continue to be very valuable. (Kanuto, 2024)

Decision Trees have practical applications in predicting employee turnover. In this regard, identification of at-risk employee groups will enable organizations to take necessary steps to reduce the rate of turnover. Decision Trees also allow companies to make immediate assessments of turnover risk using existing HR data and provide valuable insights for decision-making. Also, the integration of Decision Trees with other techniques, such as K-means clustering, enhances the prediction and analysis of turnover, hence bringing comprehensive solutions for workforce management. Decision Trees are basically one of the powerhouse tools for employee turnover predictions that provide very valuable insight into the vital factors while allowing organizations to formulate an effective retention policy. The application of the same in human resource management results in informed decisions and ensures a stable workforce. (Yunmeng et al., 2019)

3. Random Forests

The findings demonstrate that Random Forests have so far outperformed other machine learning algorithms in forecasting employee turnover. Several studies have demonstrated that Random Forest models achieve high accuracy rates, typically outperforming alternative techniques like logistic regression, decision trees, and support vector machines (Gao et al.,

2019) (Liao, 2023) (Zhang et al., 2023) (Nayak et al., 2023). According to one such study, Random Forests have a prediction accuracy of 98.8%, making them more effective than logistic regression and KNN (Zhang et al., 2023). A number of these research determined the primary factors that have a significant impact on employee turnover. These often include years of employment, age, distance from home, monthly salary, and overtime (Gao et al., 2019) (Atef et al., 2022). Also, work satisfaction has been identified as a significant predictor of turnover, and Random Forest models support this finding (Chang et al., 2022). These factors are essential for creating an accurate prediction model and could offer guidance on HR tactics to lower attrition.

The literature has proposed innovations in algorithms for the enhancement of the predictive capabilities of Random Forests. An example includes an improved weighted quadratic Random Forest algorithm to handle high-dimensional and unbalanced datasets, which proves to remarkably improve recall and F-measure compared to traditional Random Forest and other algorithms (Gao et al., 2019). Such efforts are important in refining prediction models and improving their applicability in real-world scenarios. Therefore, due to the intricacies and complexities of human behaviors along with dynamic work-place ecologies, employee churn is usually a hard task despite all the success of Random Forests. It is expected that further integration of data sources from more diversified platforms along with more hybrid models that would use ensembles of Random Forests will be explored for making accurate and reliable predictions in further research. (Zhao et al., 2018) (Islam et al., 2018) Moreover, continuous feature engineering for feature selection and embedding/dimensionality reduction contributes a lot to fine-tuning these models (Islam et al., 2018). Random Forests are a strong competitor in employee turnover prediction and achieved high accuracy with complexities of datasets. Innovation and research on this part of the area will promote the effectiveness of these works more and provide many valuable insights into making decisions for organizations targeting reducing turnover and improving employees' retention.

4. Support Vector Machine (SVM)

It is one of the preferred machine learning methods in regression analysis as well as classification problems. Its working principle is basically to divide a certain data set into groups or classes. Thus, when a new data is to be classified, it is easily determined which data set or group it belongs to. If the boundary line is drawn non-linearly rather than straight when determining the groups of the data set, this method is expressed as a non-linear support

vector machine (SVM). SVM contributes to the prediction of employee turnover in turnover estimation. In addition, it helps determine strategies. It has been observed that the results obtained from balanced data sets are highly consistent in turnover estimation (Kumar et al., 2023). It provides great support to the decision maker in the decision process thanks to its special functions to model complex problems in linear and non-linear data sets. Since its grouping feature is strong, it has been observed that it is successful in dividing the turnover estimation of employees into groups (Kumar et al., 2023, Teng et al., 2019).

5. K-Nearest Neighbors (KNN)

A machine learning algorithm used for classification and regression analysis. KNN makes predictions based on input data and the associated data to reach a result. It is preferred in turnover prediction to model the complex decisions regarding employee resignation. The steps of the KNN algorithm are as follows: In the first step, the value of K (the number of neighbors) is determined. Then, the Euclidean distance is calculated. (Euclidean distance is a method for measuring the distance between two data points.) (Kabak et al., 2016). In the third step, the K nearest neighbors are identified. For this, the positions of the elements in the dataset are checked and the nearest ones are identified. Then, the number of points in each group is determined, and the new point is assigned to the group with the higher count, completing the algorithm's process (Kumar et al., 2023).

6. Naive Bayes

One of the machine learning methods, Naive Bayes is used in problems with large datasets and for solving classification problems. Although it is frequently used in many fields, it is observed that there is relatively less research in the literature on predicting employee turnover (Valle and Ruz, 2015). Naive Bayes helps in calculating class probabilities using Bayes' theorem. Bayes' theorem is stated as follows.

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)}$$

LITERATURE REVIEW

Teng et al. (2019) proposed the Contagious Effect Heterogeneous Neural Network (CEHNN), which is based on the idea that an employee's departure could influence their colleagues to exhibit similar turnover tendencies. They created a random training and test set to evaluate the model's performance.

Their approach highlights the importance of considering social contagion effects in turnover prediction models.

Hong et al. (2007) discussed two types of turnover: voluntary and involuntary. They emphasized that predicting turnover in advance could benefit organizations by enabling preventive measures to retain employees. High turnover rates pose financial disadvantages for businesses. The study attempted to predict turnover using logit and probit models, which aimed to determine the likelihood of employees leaving an organization based on historical data. The success rate for both models reached up to 95.2%. Hong's work compared these two models for predicting employee turnover, examining their accuracy, effectiveness, and practical applications. Accurate turnover predictions can support organizations in improving strategic planning and human resource management.

Yin et al. (2024) emphasized the importance of minimizing employee turnover in the finance sector, as customer relationships are heavily dependent on employees. The study utilized features such as employee satisfaction, performance, tenure, last evaluation scores, the number of projects participated in, average monthly working hours, years spent at the company, work accident history, whether the employee left the company, promotion history within the last five years, department, and salary level (low, medium, high) to analyze turnover.

CatBoost was found to be more effective for categorical data, while XGBoost was faster and more efficient. The research focused on the differences between these two algorithms. Regression algorithms proved effective for imbalanced datasets. When using CatBoost, long processing times for categorical data were unnecessary. It provided quick predictions and avoided overfitting. XGBoost, on the other hand, utilized regularization to prevent overfitting and worked well with datasets containing missing values, offering fast predictions.

Recilla et al. (2024) emphasize the importance of maintaining low employee turnover rates regardless of the industry. Frequent job changes among employees pose an ongoing challenge for companies, making it crucial to control turnover to maintain a competitive edge. Accurate predictions, therefore, become essential in addressing this issue.

The study analyzed 30 variables using data from 2,035 respondents and employed the Genetic Algorithm approach. Four operators—FMSAX, CAX, IBAX, and AX—were applied to compare optimization processes. The results revealed that the FMSAX operator consistently achieved the highest

and most reliable performance. For instance, the “Co-workers” variable was optimized from an initial fitness value of 33,489 to 24,360.97. In contrast, IBAX and AX showed less consistent outcomes, while CAX demonstrated moderate performance. Overall, FMSAX emerged as the most effective operator.

Zhang et al. (2024) conducted statistical analyses to explore the relationship between employee turnover and various indicators. Pearson correlation analysis was used to assess the relationships among these indicators, and machine learning models such as Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naive Bayes, and Logistic Regression were applied to predict employee turnover.

The Random Forest model achieved the highest accuracy at 98.8%, followed by the SVM and KNN models, which also demonstrated strong performance. However, the Naive Bayes and Logistic Regression models showed lower accuracy. The findings indicate that the data used are effective in predicting employee turnover, and machine learning models significantly contribute to improving prediction accuracy. As the accuracy of these predictions increases, businesses can benefit by taking appropriate measures to address potential turnover risks.

Ajit et al. (2016) addressed voluntary turnover, a costly and undesirable issue for companies. Voluntary employee departures often involve significant expenses, including onboarding and training for replacements, particularly when highly skilled employees leave. The study focused on analyzing voluntary turnover and considered key parameters such as age, tenure, pay, overall job satisfaction, and employees’ perceptions of fairness.

For prediction, the study utilized the Extreme Gradient Boosting (XGBoost) technique and tested seven different models: Logistic Regression, Naïve Bayes, Random Forest, K-Nearest Neighbor (KNN), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), and XGBoost. The importance of XGBoost was highlighted, with the authors emphasizing that future research should shift focus from identifying who might leave to exploring what can be done to retain employees.

This study focuses on measuring employee commitment and determining how satisfied workers are with their current conditions in a retail company in Malaysia. It explores the relationship between job satisfaction and turnover intention. To investigate this relationship, a questionnaire was developed. The developed questionnaire was based on the Job Descriptive Index, Organizational Commitment Questionnaire, and Lee and Mowday’s

turnover intention items, and was answered by 62 participants. The results of the survey revealed that employees in the retail sector in Malaysia are not satisfied with their salary. Additionally, it was found that employees have a moderate level of commitment to their jobs. There is a negative and significant relationship between job satisfaction factors and turnover intention. In response to the findings, the article suggests that increasing salaries and managers supporting employees could have a positive effect in reducing turnover rates.

One of the most important factors for the success of firms is the employees' commitment to the company. A decrease in employee engagement and situations such as employees leaving the company create obstacles to the success of firms. In their study, Arshad and Puteh (2015) aim to determine the turnover tendencies of employees. A total of 106 employees from four different branches of the company were surveyed, as mentioned in the article. The factors affecting employee turnover in firms are listed as Perceived Organizational Support, Job Stress, Work-Life Balance, and Available Job Alternatives/Opportunities. SPSS was used to analyze the results. The findings show that work-life balance and other job alternatives are key factors leading employees to leave their organizations.

In his literature-based study, Ongori (2007) examined the reasons for employee turnover, its effects on organizations, and the measures companies can take to address employee turnover. The study categorized the reasons for employee turnover into organizational factors, environmental factors, and personal factors. Criteria such as dissatisfaction with management, career goals, insufficient salary, and unfavorable working conditions were discussed. The study also highlighted both the negative and positive impacts of employee turnover on organizations and proposed strategies to address this issue.

In their study, Alla and Rajaa (2019) classified employee turnover into Functional/Dysfunctional turnover and Voluntary/Involuntary turnover. The study listed and analyzed the factors associated with each group, arguing that organizations should not only aim to reduce turnover rates but also analyze them according to these classifications to take appropriate actions. The article also included strategies to address employee turnover.

They have worked on predicting employee turnover and developing strategies to retain employees within the organization. In their research, Ribes et al. (2017) utilized mathematical and statistical models. By applying machine learning methods, they were able to more effectively identify the reasons behind employee turnover. Additionally, they compared various

machine learning techniques, such as Random Forest, Linear Discriminant Analysis, and Support Vector Machines (SVMs). They found that more complex models were more successful in prediction compared to simpler models. Through the developed mathematical and statistical models, they were able to predict employees with a high likelihood of turnover in advance, allowing for the implementation of strategies to retain these employees.

In their study, Brown et al. (2024) proposed the use of deep learning techniques to predict employee turnover. They compared deep learning methods with other machine learning techniques, such as Random Forest, Decision Trees, and Support Vector Machines. They argued that artificial intelligence-based deep learning algorithms are more effective than traditional machine learning methods. They suggested that better results are obtained when working with very large datasets. The authors highlighted that deep learning methods provide a significantly higher accuracy in predicting employee turnover. They stated that this could enable organizations to develop more accurate strategies.

In Effendi's (2024) study, the factors influencing employee turnover were examined, and Logistic Regression analysis was used as the method. The study identified factors affecting employee commitment, such as salary, career advancement, teamwork, and working hours. The method revealed how these factors influenced employees' decisions to leave their jobs. According to the model's results, salary was found to be a strong determinant among the factors influencing employee turnover. Furthermore, the article reported that the model could predict employee turnover with an accuracy of 82%.

In their study, Nketsiah and Nkansah (2024) examined the turnover intentions of employees in the banking sector in Ghana based on certain criteria. The research explored how factors such as trust and commitment influence employees' decisions to stay in their current jobs. Data was collected from the sample, and it was found that there is an inverse relationship between employees' level of organizational commitment and their turnover intentions. The authors noted that the turnover rates of employees in local and foreign banking firms may differ from each other. They suggested that Human Resources activities should be restructured to support and enhance employees' commitment to the organization.

The explanation of the numbers corresponding to the turnover prediction parameters in the table is stated below.

Table 1. Parameters of Employee Turnover.

Parameters of Employee Turnover			
1	Working Conditions	13	Team Dynamics
2	Job Satisfaction	14	Infectious Effect
3	Lack of Training	15	Technological Changes
4	Management Support	16	Personal Development
5	Pay Fairness	17	Unsuitability for Work
6	Work-Life Balance	18	Lack of Recognition
7	Workload Balance	19	Job Security
8	Organizational	20	Lack of Social Ties
9	Career Opportunities	21	Project Dynamics
10	External Opportunities	22	Business Culture
11	Leadership Style	23	Perception of Justice
12	Workplace Conflicts		

Table 2. Employee Turnover Parameters in the Studies

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Salleh, R., et al. (2012)		✓						✓	✓														
Arshad, H., & Puteh, F.	✓			✓			✓																
Ongori, H. (2007)			✓	✓		✓																	
Alla, A. A., & Rajaa, O.		✓								✓		✓											
Teng, M., et al. (2019)				✓									✓	✓									
Hong, W. C., et al. (2007)					✓	✓					✓												
Yin, Z., et al. (2024)	✓	✓													✓								
Recilla, V. J., et al. (2024)				✓												✓							
Zhang, J., et al. (2024)			✓														✓						
Ajit, P. (2016)		✓				✓												✓					
Ribes, E., et al. (2017)		✓																	✓	✓			
Brown, A., et al.										✓													
Effendi, S. R. F. (2024)					✓		✓														✓		
Nketsiah, T. A., & Chang, Y., et al. (2024)								✓											✓			✓	
	✓							✓															✓

According to the data in the table, “job satisfaction” and “organizational commitment” are among the most frequently mentioned factors. This illustrates that the employees’ level of job satisfaction and organizational commitment is directly related to their decision to leave the job. Further, factors like “management support” and “work-life balance” are also high on mentions, showing how important it is for workers to receive expectations from leaders in the organization and their personal life balance.

Table 3. Machine Learning Methods Used in the Studies.

Study	Random Forest	Support Vector Machine	CatBoost	XGBoost	Artificial Intelligent	Logistic Regression	Statistical Analysis Methods	KNN	Naive Bayes
Salleh, R., Nair,							✓		
Arshad, H., &							✓		
Teng, M., Zhu, H.,	✓	✓							
Yin, Z., Hu, B., &			✓	✓					
Recilla, V. J.,					✓				
Zhang, J., Cao, Z.,		✓							
Ajit, P. (2016)								✓	✓
Ribes, E., Touahri,	✓	✓							
Brown, A., Davis,					✓				
Effendi, S. R. F.						✓			
Nketsiah, T. A., &							✓		

Studies in the literature have been examined and a table of machine learning methods used in the studies has been stated. In the table, it is concluded that Statistical Analysis Methods and Support Vector Machine methods are used more frequently than other methods. It has been observed that the combination of Support Vector Machine algorithm with Random Forest method increases the accuracy in the estimation of leaving the job and is useful in evaluating the effects of the previously mentioned effective factors in leaving the job.

The figure presents some variations of machine learning and data analytics techniques adopted for the predictions of employee turnover. The most frequent occurrence is the ‘supervised’ technique, with 14 papers in total. All the well-known algorithms, such as Support Vector Machine, Random Forest, XGBoost, Catboost, Logistic Regression, and Deep Learning, form a part of this category. Techniques like Deep Learning and Logistic Regression occur in two papers. More precisely, the RF and XGBoost algorithms were compared in two papers published in the year 2024. K-Means Clustering and Hierarchical Clustering were the unsupervised techniques applied in only 2 papers. Besides, 2 papers have applied the Hybrid methods. These techniques combined supervised and unsupervised techniques. In one of these, they used GA for prediction, while others used the techniques that combined Big Data Analytics with Machine Learning techniques. In view of

this distribution, in most cases, employee churn could be predicted with the application of the supervised machine learning techniques, while only some concrete cases could have unsupervised and hybrid approaches applied.

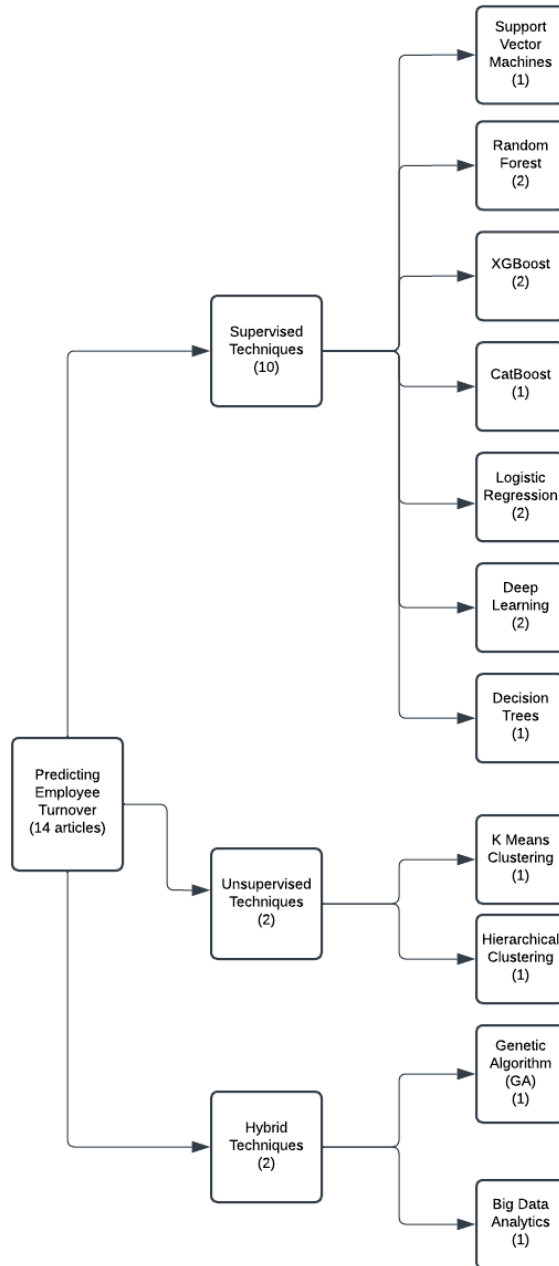


Figure. 1. Machine Learning Techniques for Employee Turnover Prediction in the Literature

CONCLUSION

In this study, employee layoffs were categorized into two types: voluntary and involuntary. voluntary layoffs could result in a financial loss to the company. Predicting the redundancies earlier is a situation that most of the companies want to have so that measures can be taken depending on the situation and retention strategies can be developed for employees who are at risk of quitting their jobs. In order to predict the exit from work, machine learning algorithms are generally used. Supervised machine learning algorithms generally provide more accurate results than machine learning algorithms. For this reason, the study of supervised machine learning algorithms was carried out in the study. While the data were being examined within the literature research in order to be able to use machine learning algorithms, some parameters came to the fore in employee data. The most used ones among these parameters are: working conditions, job satisfaction, lack of training, management support, pay fairness, work-life balance, workload balance, organizational commitment, career opportunities, external opportunities, technological changes, personal development. It has, in fact studied the role of algorithms operating according to these parameters in forecasting. Logistic Regression, Decision Trees, Random Forests, Support Vectos Machines(SVM), K-Nearest Neighbors (KNN), Naive Bayes from Supervised machine learning algorithms have been examined in the literature and their role in forecasting.

Each of the machine learning algorithms used for predicting employee turnover rate has various advantages and disadvantages. Logistic Regression is very powerful, especially in terms of simplicity and interpretability. The transparency provided by employees in understanding the reasons for leaving their jobs makes it easier for human resources teams to make strategic decisions. However, it usually stands behind more complex algorithms, such as Random Forest or Support Vector Machine, in terms of accuracy. On the other hand, Decision Trees, due to their simplicity and visualization, are useful in showing major reasons for the employee turnover rate. However, this algorithm is not as effective as Random Forest or Gradient Boosting Trees, especially on unstable and complex datasets. The Random Forest attracts attention as one of the algorithms showing the highest accuracy rates in employee turnover rate forecasting. Strong performance allows, in particular for large and complex data sets, the elaboration of effective strategies in human resource management. However, these high accuracy rates have to be offset against increased computational costs. SVM is an excellent choice because of its accuracy and success in classifying complex datasets. Other algorithms, like KNN and Naive Bayes, work well in certain

situations but are less preferred for large data sets. Therefore, Random Forest yielded very good performance with high accuracy rates, while its application areas are really wide in predicting the rate of employee turnover and recommending that the organization make a decision by taking into consideration the structure of datasets, their balance, and business needs. Simpler models could offer effective alternatives, at least when Logistic Regression, speed of implementation, and interpretability are a priority. The same approach may offer organizations a strategic benefit in the engagement of workers and in limiting turnover.

References

- Ahmed, S., Ahmed, A., & Jwmaa, S. (2023). Analyzing The Employee Turnover by Using Decision Tree Algorithm. *2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, 1-4. <https://doi.org/10.1109/HORA58378.2023.10156709>.
- Alaskar, L., Crane, M., & Alduailij, M. (2019). Employee Turnover Prediction Using Machine Learning. *Communications in Computer and Information Science*. https://doi.org/10.1007/978-3-030-36365-9_25.
- Atef, M., Elzanfaly, D., & Ouf, S. (2022). Early Prediction of Employee Turnover Using Machine Learning Algorithms. *International Journal of Electrical and Computer Engineering Systems*. <https://doi.org/10.32985/ijeccs.13.2.6>.
- Birzniece, I., Andersone, I., Nikitenko, A., & Zvirbule, L. (2022). Predictive Modeling of HR Dynamics Using Machine Learning. *Proceedings of the 2022 7th International Conference on Machine Learning Technologies*. <https://doi.org/10.1145/3529399.3529403>.
- Chang, V., Mou, Y., Xu, Q., & Xu, Y. (2022). Job satisfaction and turnover decision of employees in the Internet sector in the US. *Enterprise Information Systems*, 17. <https://doi.org/10.1080/17517575.2022.2130013>.
- Chen, B. (2023). Factors of Employee Attrition: A Logistic Regression Approach. *Advances in Economics, Management and Political Sciences*, 20, 214-225.
- Effendi, S. R. F. (2024). The Effect of Job Division, Promotion, Project, Salary, Length of Work With Employee Turnover Using Logistic Regression Analysis. *Jurnal Indonesia Sosial Teknologi*, 5(3), 1204-1210.
- Gao, X., Wen, J., & Zhang, C. (2019). An Improved Random Forest Algorithm for Predicting Employee Turnover. *Mathematical Problems in Engineering*. <https://doi.org/10.1155/2019/4140707>.
- Hong, W. C., Wei, S. Y., & Chen, Y. F. (2007). A comparative test of two employee turnover prediction models. *International Journal of Management*, 24(4), 808.
- Islam, M., Alam, M., Islam, M., Mohiuddin, K., Das, A., & Kaonain, M. (2018). An Adaptive Feature Dimensionality Reduction Technique Based on Random Forest on Employee Turnover Prediction Model. *Advances in Intelligent Systems and Computing*, 269-278. https://doi.org/10.1007/978-981-13-1813-9_27.
- Jhaver, M., Gupta, Y., & Mishra, A. (2019). Employee Turnover Prediction System. *2019 4th International Conference on Information Systems and Computer Networks (ISCON)*, 391-394. <https://doi.org/10.1109/ISCON47742.2019.9036180>.

- Kanuto, A. (2024). Identifying Patterns and Predicting Employee Turnover Using Machine Learning Approaches. *International Journal of Science and Business*. <https://doi.org/10.58970/ijsb.2373>.
- Kabak, M., Saglam, F., & Aktaş, A. (2017). Usability analysis of different distance measures on TOPSIS. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 32(1).
- Krishna, S., S., & Borah, D. (2023). Machine Learning for Ensuring Sustainable Development: Predicting Employee Attrition in the Workplace. *2023 International Conference on Advanced Computing Technologies and Applications (ICACTA)*, 1-7.
- Kumar, P., Gaikwad, S. B., Ramya, S. T., Tiwari, T., Tiwari, M., & Kumar, B. (2023). Predicting Employee Turnover: A Systematic Machine Learning Approach for Resource Conservation and Workforce Stability. *Engineering Proceedings*, 59, 117.
- Leana, C. R., & Ivancevich, J. M. (1987). Involuntary job loss: Institutional interventions and a research agenda. *Academy of Management Review*, 12(2), 301-312.
- Liao, C. (2023). Employee turnover prediction using machine learning models. *Proceedings of SPIE*, 12596, 125960X-5. <https://doi.org/10.1117/12.2672733>.
- Nayak, S., & Palai, P. (2023). Employee Attrition System Prediction using Random Forest Classifier. *International Journal of Computer and Communication Technology*. <https://doi.org/10.47893/ijcct.2023.1445>.
- Punnoose, R., & Ajit, P. (2016). Prediction of Employee Turnover in Organizations using Machine Learning Algorithms: A case for Extreme Gradient Boosting. *International Journal of Advanced Research in Artificial Intelligence*. <https://doi.org/10.14569/IJARAI.2016.050904>.
- Russell, C. J. (2013). Is it time to voluntarily turn over theories of voluntary turnover? *Industrial and Organizational Psychology*, 6(2), 156-173.
- Setiawan, I., Suprihanto, S., Nugraha, A. C., & Hutahaean, J. (2020). HR analytics: Employee attrition analysis using logistic regression. *IOP Conference Series: Materials Science and Engineering*, 830(3). <https://doi.org/10.1088/1757-899X/830/3/032001>.
- Taner, Z., Hızıroğlu, O. A., & Hızıroğlu, K. (2024). Leveraging Machine Learning Methods for Predicting Employee Turnover Within the Framework of Human Resources Analytics. *Journal of Intelligent Systems: Theory and Applications*, 7(2), 145-158.
- Valle, M. A., & Ruz, G. A. (2015). Turnover prediction in a call center: behavioral evidence of loss aversion using random forest and naïve bayes algorithms. *Applied Artificial Intelligence*, 29(9), 923-942.

- Zhang, C., & Han, W. (2024). Ensembles of decision trees and gradient-based learning for employee turnover rate prediction. *PeerJ Computer Science*. <https://doi.org/10.7717/peerj-cs.2387>.
- Zhang, J., Cao, Z., Qu, H., & Wang, M. (2024). Management and prediction of employee turnover in enterprises based on big data analytics and machine learning. *Applied and Computational Engineering*, 76, 103-108.
- Živković, A., Šebalj, D., & Franjkovic, J. (2024). Prediction of the Employee Turnover Intention Using Decision Trees. *Proceedings of the 2024 International Conference*, 325-336. <https://doi.org/10.5220/0012538400003690>.
- Zhao, Y., Hryniewicki, M., Cheng, F., Fu, B., & Zhu, X. (2018). Employee Turnover Prediction with Machine Learning: A Reliable Approach. *Lecture Notes in Computer Science*, 737-758. https://doi.org/10.1007/978-3-030-01057-7_56.
- Yunmeng, Z., & Chengyi, Z. (2019). The Application of The Decision Tree Algorithm Based on K-means in Employee Turnover Prediction. *Journal of Physics: Conference Series*, 1325. <https://doi.org/10.1088/1742-6596/1325/1/012123>

Conflict of Interest

The authors have declared that there is no conflict of interest.

Author Contributions

Melisa Dikici and Gökçe Sabriye Hörük write the Article. Deniz Efendioğlu supervised the article.

A Macroeconomic Perspective on Türkiye's Climate Crisis Risk: Decision Tree Analysis

Mervenur Ünver¹

Şahika Gökmen^{2,3}

Abstract

Green Swan is a risk concept used to describe unpredictable, sudden, and systemic economic crises caused by climate change. Awareness of Green Swan risks began with the 2015 Paris Climate Agreement and was conceptually framed by a 2020 report from the Bank for International Settlements (BIS) and the Banque de France. Green Swan risks are not only environmental but also have systemic impacts that trigger economic and social vulnerability. A review of the literature reveals that no empirical studies specific to Türkiye, which analyze such risks on an annual basis and model them using machine learning, have been encountered. This study aims to model Türkiye's climate-related risk levels using environmental, climatic, and economic indicators, based on ND-GAIN and CO₂ emissions data from 1995 to 2022. A total of 13 variables were used to classify annual risk levels through the Classification and Regression Tree (CART) algorithm. Risks were evaluated across four levels: uncertain, low, medium, and high. The model's accuracy was measured at 82.14%. The findings indicate that GDP growth rate, labor force participation rate, and temperature anomalies serve as key determinants in determining risk levels. While offering a novel national-level contribution to the Green Swan literature, this study proposes a feasible and policy-relevant analytical method for policymakers to detect climate-induced economic vulnerabilities at an early stage.

- 1 Master Student, Department of Econometrics, Faculty of Economics and Administrative Sciences, Ankara Hacı Bayram Veli University, 06500 Ankara, Türkiye
- 2 Assoc. Prof. Dr., Department of Econometrics, Faculty of Economics and Administrative Sciences, Ankara Hacı Bayram Veli University, 06500, Ankara, Türkiye
- 3 Researcher, Statistics Department, Uppsala University, Uppsala, Sweden

INTRODUCTION

The concept of Green Swan is used to define climate-induced economic risks that may result in sudden, unpredictable, and systemic impacts. This concept was introduced to the literature through a report published in 2020 by the Bank for International Settlements (BIS) and the Banque de France. These risks, which can have devastating effects not only on the environment but also on economic and financial systems, point to climate shocks that traditional financial risk analyses are inadequate to foresee. The first global awareness of Green Swan risks began with the 2015 Paris Climate Agreement, which initiated international policy-level discussions on the macroeconomic vulnerabilities associated with climate change.

Green Swan risks emerge in diverse forms not only on a global scale but also at the national level and therefore require interdisciplinary and data-driven methodologies for proper assessment by decision-makers. The macroeconomic effects of climate change are felt more acutely in countries with fragile economies, posing a significant threat to sustainable development policies. Due to its geographical location and economic structure, Türkiye is among the countries vulnerable to such risks.

However, a review of the existing literature reveals that no studies specifically analyzing Green Swan risks in Türkiye on an annual basis and examining them through machine learning-based models have been encountered. Existing international studies are largely confined to scenario-based assessments. Most international studies on climate-related financial risks focus predominantly on scenario-based assessments of transition and physical climate risks, while quantitative, machine-learning-based approaches remain relatively limited (Bolton et al., 2020; McKinsey Global Institute, 2020; IMF, 2024). One of the few studies addressing Green Swan risks in the context of Türkiye is that of Açıkgöz and Gökmen (2022). This study discusses the potential for climate-induced financial instability in Türkiye and Annex II countries within a conceptual scenario analysis framework, based on transition and physical risk scenarios developed by international organizations such as the NGFS.

In this context, the main objective of the study is to determine Türkiye's climate-related risk levels on an annual basis and to analyze the environmental, economic, and climatic indicators influencing these risks using the machine learning-based Classification and Regression Tree (CART) method, thus generating findings that can serve as early warnings against Green Swan risks and contribute to the literature both methodologically and practically.

In the study, a machine learning-based classification model was developed using ND-GAIN and CO₂ emissions data from 1995 to 2022, along with 13 climatic, environmental, and economic variables. Accordingly, the subsequent sections of the study present the materials and methods, findings, and finally, the conclusion and policy recommendations.

METHODS AND FINDINGS

Methods

CART Algorithm

The Classification and Regression Tree (CART) algorithm is a supervised machine learning method utilized for both classification and regression analyses. CART operates by partitioning the dataset into increasingly homogeneous subgroups with respect to the target variable. At each splitting step, the algorithm selects the variable and threshold that result in the greatest reduction of impurity—measured by the Gini index in classification problems. This process continues until a predefined stopping criterion, such as the minimum number of observations at a node or the maximum tree depth, is met.

In this study, the CART algorithm was employed to classify Türkiye's annual climate-related risk levels based on climatic, environmental, and economic indicators. The reasons for preferring the algorithm include its high interpretability, ability to model nonlinear relationships, and suitability for small to medium-sized datasets. Furthermore, the decision tree outputs generate transparent and intuitive decision rules that can be easily understood by policymakers, making it a particularly valuable tool for climate-based early warning systems.

Findings

In this study, annual climate-related risk levels were initially determined by evaluating the ND-GAIN Index and CO₂ emission data for the period 1995–2022. The risk levels were classified into four-category risk level classification for each year: uncertain (0), low (1), medium (2), and high (3). The reason of using the ND-GAIN Index is that it is a global indicator that jointly assesses countries' levels of climate vulnerability and adaptive capacity. CO₂ emissions, along with other environmental and economic indicators, were included in the model as variables directly associated with climate risk. All indicators used in the study were obtained from the databases of the Notre Dame Global Adaptation Initiative (ND-GAIN), Our World in Data, and the World Bank. The common years across all variables were

identified, and the dataset was temporally harmonized to cover the 1995–2022 period.

The model incorporates 13 variables representing environmental, economic, and climatic factors which are Winter Temperature Anomaly, Summer Temperature Anomaly, Spring Temperature Anomaly, Autumn Temperature Anomaly, Ratio of Foreign Trade to Gross Domestic Product, Renewable Energy Consumption, Labor Force Participation Rate, Gross Domestic Product Growth Rate, Forest Land Ratio, Agricultural Land Ratio, Per Capita Energy Consumption, Per Capita Electricity Consumption, and Access to Clean Fuels.

Subsequently, a machine learning-based decision tree model was developed to identify the main indicators influencing these risk levels. In this context, the Classification and Regression Trees (CART) algorithm was employed. The model was constructed using 13 environmental, climatic, and economic indicators.

All data processing, modeling, and analysis procedures were conducted in the R Studio environment. The classification was implemented using the *rpart* package, and the model achieved an accuracy of 82.14%. Additionally, the relative importance of each variable in determining climate-related risk levels was evaluated.

RESULTS

In line with the classification model developed in the study, Türkiye's climate-related risk levels for the period 1995–2022 were visualized over time, and the structure of the variables influencing the risk levels was revealed through the decision tree. Additionally, variable importance ranking was analyzed to identify which indicators contributed most significantly to the model. The graphs presented in *Figure 1* provides a comprehensive representation of both the temporal distribution and the model's decision rules and variable weights.

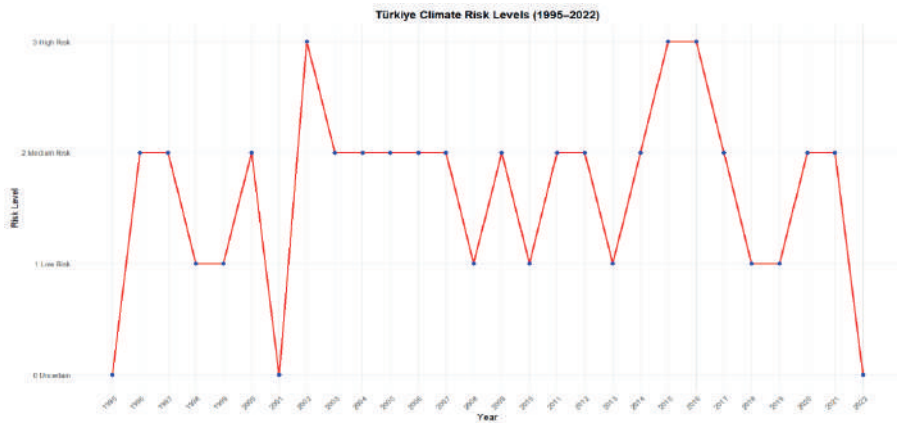


Figure 1. Changes in Turkey's Climatic Risk Levels by Year (1995–2022)

Based on the findings of this study, Türkiye’s climate-related risk levels between 1995 and 2022 were classified and visualized in *Figure 1*. The annual data were categorized into four classes based on the ND-GAIN Index and per capita CO₂ emissions: 0 – Uncertain, 1 – Low Risk, 2 – Medium Risk, and 3 – High Risk.

As it can be seen in *Figure 1*, the most frequent category was “medium risk,” representing 46% of the total observations. This result suggests that, during the analyzed period, Türkiye’s climate vulnerability was predominantly at a moderate level.

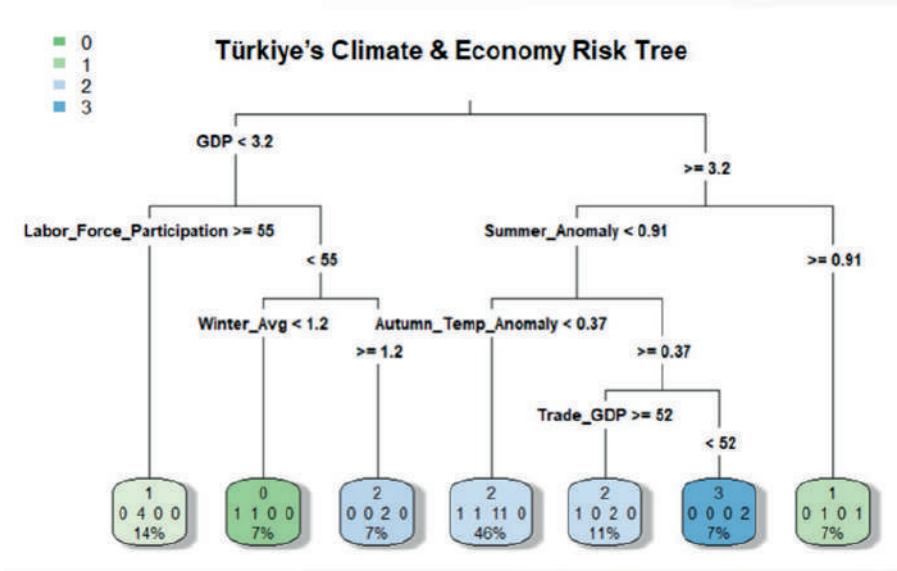


Figure 2. Classification of risk levels using the decision tree model

The decision tree model created for the classification process visualizes the variables and threshold values that are effective in determining the climate-related risk level for each year. This model clearly identifies the key threshold values and determining variables that influenced Türkiye's climate-related risk levels between 1995 and 2022. In the decision tree presented in *Figure 2*, the variable that performs the initial split at the root node is the Gross Domestic Product (GDP) growth rate. Years with a GDP growth rate below 3.2 and those with a rate of 3.2 or above are directed to different branches. This indicates that macroeconomic performance plays a leading role in determining climate-related risk levels.

For years in which the Gross Domestic Product (GDP) growth rate was below 3.2, the second variable influencing the split was the labor force participation rate (Labor_Force_Participation). When this rate exceeded 55, those years were mostly classified as low risk (1).

For years with a labor force participation rate below 55, the determining variable in the decision mechanism was the winter temperature anomaly. For example:

- If the winter temperature anomaly (Winter_Avg) was below 1.2, the year was classified as uncertain risk (0).
- If the winter temperature anomaly (Winter_Avg) exceeded 1.2, the year was classified as medium risk (2).

For years in which the GDP growth rate was ≥ 3.2 , the split was made based on the summer temperature anomaly (Summer_Anomaly).

- When the summer temperature anomaly was below 0.91, further splits were influenced by the autumn temperature anomaly and the trade-to-GDP ratio (Trade_GDP). In particular, when the trade-to-GDP ratio was below 52, the year was classified as high risk (3). If the Summer_Anomaly was 0.91 or above, the year was directly classified as low risk (1).

This model clearly demonstrates which variables are dominant under different socioeconomic and climatic conditions, as well as the thresholds at which the risk level increases. The largest leaf of the decision tree represents the medium risk level, encompassing 46% of the dataset.

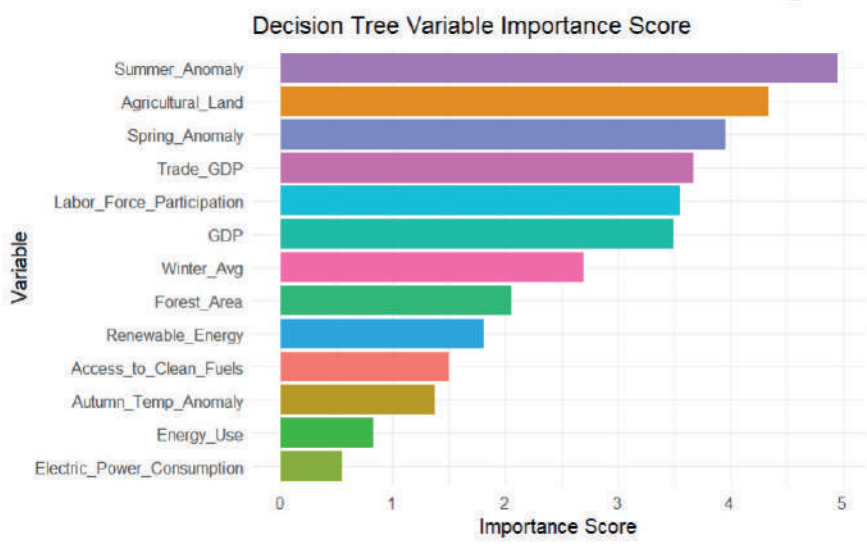


Figure 3. Importance level of variables according to the decision tree model

Figure 3 presents the relative contributions of the variables used in the decision tree model to the classification process. The values displayed on the “Importance Score” axis indicate which variables were more influential in the model’s decision-making mechanism.

The highest importance score belongs to the Summer Temperature Anomaly (Summer_Anomaly) variable, indicating that climate-related risk levels in Türkiye are significantly affected by temperature deviations during the summer season.

The second most important variable is Agricultural Land Ratio (Agricultural_Land). Its high score suggests that land use and agricultural activities play a determining role in climate vulnerability.

These are followed by Spring Temperature Anomaly (Spring_Anomaly), Trade-to-GDP Ratio (Trade_GDP), and Labor Force Participation Rate (Labor_Force_Participation). In particular, the labor force participation rate served as a critical splitting variable, as it appeared at an early branching point in the decision tree structure.

On the other hand, variables such as Electric Power Consumption, Energy Use, and Autumn Temperature Anomaly (Autumn_Temp_Anomaly) had relatively limited contributions to the model. The low importance scores of these variables indicate that they were less influential in the classification decisions.

This analysis is significant in demonstrating how environmental and economic variables can have differing levels of influence when jointly evaluated in determining climate-related risk levels. The findings provide a data-driven perspective for policymakers on which areas should be prioritized.

DISCUSSION AND CONCLUSION

In this study, Türkiye's annual climate-related risk levels for the period 1995–2022 were categorized into four groups and classified using the Classification and Regression Tree (CART) model, based on environmental, climatic, and economic indicators. The findings indicate that variables such as summer temperature anomaly, agricultural land ratio, and spring temperature anomaly stand out in determining risk levels. The model's classification accuracy of 82.14% supports both the explanatory power of the selected variables and the overall reliability of the model.

The analysis of the risk classes can be summarized as follows: The low-risk level was observed during periods characterized by a combination of low GDP growth and high labor force participation rates. This finding aligns with the study by Achuo, Nchofoung, Zanfack, and Epoge (2023), which highlights the positive impact of labor force participation on environmental sustainability.

In the medium-risk category, periods characterized by low summer and autumn temperature anomalies, but high GDP growth rates stand out. This finding supports the McKinsey Global Institute (2020) report, which emphasizes the non-linear effects of climate change and the vulnerabilities that arise when systemic thresholds are exceeded.

In the high-risk category, a low share of foreign trade in GDP corresponds with the observation made in the IMF (2024) report, which highlights the limited capacity of trade volume to offset production and consumption losses in the face of climate shocks.

The model results reveal that Green Swan risks should be evaluated not only through environmental indicators but also in conjunction with economic structures and socioeconomic factors. In this regard, the decision tree model developed in the study emerges as a transparent, interpretable, and data-driven early warning system for forecasting climate-related risks specific to Türkiye.

In conclusion, this study aims to contribute to Green Swan literature at the national level while also providing a valuable foundation for policymakers

to develop foresight-based intervention strategies. Future research is recommended to compare different machine learning algorithms, conduct regional-level analyses, and expand the model through long-term scenario assessments. This model can be utilized as a practical early warning tool by national agencies involved in environmental and economic planning.

References

- Achuo E. D., Nchofoung T. N., Zanfack L. J. T., Epoge C. E., 2023. The nexus between labour force participation and environmental sustainability: Global comparative evidence. *Heliyon*, 9(11). Access address: <https://doi.org/10.1016/j.heliyon.2023.e21434>; Date of access: May 2025.
- Açıkgöz Ş, Gökmen Ş, 2022. Understanding the Effects of Green Swan Events on Financial Stability: Annex II Countries and Turkey. In: Aladağ ÇH, Potas N (Editors), *Modeling and Advanced Techniques in Modern Economics*, World Scientific Publishing, pp. 127–161. https://doi.org/10.1142/9781800611757_0006; Date of access: May 2025.
- Bolton P., Després M., Pereira da Silva L. A., Samama F., Svartzman R., 2020. *The Green Swan: Central Banking and Financial Stability in the Age of Climate Change*. Bank for International Settlements and Banque de France. <https://www.bis.org/publ/othp31.htm>; Date of access: May 2025.
- Chen C., Kirabaeva K., Zhao D., 2024. Investing in climate adaptation under trade and financing constraints: Balanced strategies for food security. IMF Working Paper No. WP/24/184. Access address: <https://doi.org/10.5089/9798400289132.001>; Date of access: May 2025.
- Woetzel L., Pinner D., Samandari H., Engel H., Krishnan M., Boland B., Powis C., 2020. *Climate risk and response: Physical hazards and socioeconomic impacts*. McKinsey & Company. Access address: <https://www.mckinsey.com/capabilities/sustainability/our-insights/climate-risk-and-response-physical-hazards-and-socioeconomic-impacts>; Date of access: May 2025.
- Notre Dame Global Adaptation Initiative, 2024. *Country Index Scores*. Access address: <https://gain.nd.edu/our-work/country-index/>; Date of access: May 2025.
- Our World in Data, 2024. Access address: <https://ourworldindata.org/climate-change>; Date of access: May 2025.
- World Bank, 2024. *World development indicators (WDI)*. Access address: <https://data.worldbank.org>; Date of access: May 2025.

Acknowledgment

I would like to express my sincere gratitude to my advisor, Assoc. Prof. Dr. Şahika Gökmen, whose guidance, academic insights, and valuable contributions have greatly enriched every stage of this study.

Mplus Jamovi Amos and Spss Programs in Partial Mediation Model: A Simulation Study

Murat Yildirim¹

Abstract

This study aims to investigate the performance of Mplus, Jamovi, Amos and Spss programs, which are widely used for partial mediation model, with a certain sample size and categorical data. In line with the purpose of the study, R program was used to generate data for partial mediation model. The three path coefficients of the model are pre-assigned as real values with a 0.6, b 0.7 and c path 0.3 and there is no other pre-assigned value. Sample size selection is 100, 200, 300, 400 and 500 units. In the established model, X latent variable has 4, M latent variable has 3 and Y latent variable has 5 observed variables. All observed variables have 5-categorical structure. The model was analyzed in all programs; direct effects, indirect effects and fit values were obtained. The estimation techniques of the programs are WLSMV, DWLS, ML and OLS, respectively. The closeness of the programs to the previously assigned values as performance criteria was intended to be examined. For this situation, Mean Absolute Bias- MAB and Relative Bias- RB values were used. The findings obtained were as follows. In all programs, the direct effect from the latent variable X to the latent variable Y was found to be insignificant in sample sizes of 100 and 200 units, and all other path coefficients were found to be significant. In this case, the existing partial mediation model was transformed into a full mediation model. This result indicates that the sample size of 100 and 200 units is insufficient for model studies on partial mediation relationship. While all programs had overestimation in the sample size of 100 units, underestimation was obtained in sample sizes increasing with 200 units. All programs had quite good fit values in all volumes. According to MAB and RB values, Spss indicates that its use for categorical data is problematic with its quite high deviation results. In other programs, MAB and RB values of 300 units sample were found to be good and the best performance was provided by Mplus program with WLSMV estimation technique. The next performance ranking after Mplus program was Jamovi and Amos.

1 Tokat Gaziosmanpaşa University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Türkiye

INTRODUCTION

Structural equation modeling (SEM) has a historical background with the assumption of linearity (Bollen 1989) that addresses the relationships between latent variables through observed variables. This assumption of linearity has been included in many packaged programs. However, this limitation is not sufficiently appropriate in the branches of science where SEM is frequently used (Lee and Zhu, 2000; Wall and Amemiya, 2007; Wall, 2009). Because researchers in the fields of psychometrics, education and marketing generally address the quadratic variables of latent variables (Arminger and Muthen, 1998).

When the hypotheses of solid theories are examined, we see that latent variables point to interaction and quadratic variables. For example, in psychology, the theory of reasoned action, expectancy-value attitude theory or an extension of the theory of planned behavior can be given. We see another example in educational sciences. Ganzach (1997) stated that children's expectations from education depend on the nonlinear relationships between their parents' level of education (Lee and Zhu, 2000; Moosbrugger et al., 2009). Many theories in social and behavioral sciences have nonlinear relationships of latent variables (Kelava and Brandt, 2009; Kelava and Nagengast, 2012).

Therefore, expanding the linearity restriction by adding nonlinear relationships of latent variables to SEM studies will allow the creation of strong theories by estimating robust hypotheses (Wall and Amemiya, 2007).

Recent studies have determined that adding nonlinear relationships of latent variables to the model provides better causality (Lee and Song, 2003). More than one model type has been defined for nonlinear structure types and more than one estimation technique has been developed for them. In this way, a rapidly growing literature has emerged in this field (Wall and Amemiya, 2007).

Model studies addressing the interaction effects of latent variables have become important statistical tools in many branches of science. Studies in psychology and behavioral sciences increasingly address the mediation and moderation effects of independent variables in order to better reveal causality on independent variables (Çağlayan and Özenç, 2024; Hayes, 2018).

In a content study conducted in educational sciences in the national literature, it was determined that 74% of the theses conducted between 2004 and 2023 were single mediation models. This difference seen in mediation analysis studies was effective in determining the model in the current study.

Mostly ready-made scales or partially developed scales were used as data collection tools for mediation analysis (Çağlayan and Özenç, 2024). The data obtained through scales are in the type of ordinal categorical data and there are multiple estimation techniques for their evaluation in frequentist statistics. These techniques are recommended as Unweighted Least Square (ULS) and Diagonally Weighted Least Square (DWLS). When there is continuous data and the data assumes normality, the most commonly used method is the Maximum Likelihood (ML) method (Jöreskog, 1969; Maydeu-Olivares, 2017; Shi et al., 2018; Shi and Maydeu-Olivares, 2020).

It is observed that in model studies in social and behavioral sciences, ordinal categorical data is frequently used and researchers also frequently use the ML method as a estimation technique. However, it is quite difficult to ensure normality in data belonging to ordinal variables. Different techniques developed under these conditions need to be used. Since the model will only calculate the indirect effect with the mediation model, the performance of these techniques should be examined even if the data is normally distributed (Yang-Wallentin et al., 2010).

The present study aims to measure the performance of Mplus, Jamovi, Amos and Spss package programs, which are frequently used in simple mediation models. Considering that researchers are not conscious about choosing the estimation technique, the default estimation techniques given directly for each program are WLSMV (Weighted least squares mean and variance-corrected weighted least squares), DWLS, ML and OLS (Ordinary least squares), respectively. The given estimation techniques were evaluated by considering biases from the perspective of sample size.

MATERIAL AND METHODS

Material

The model established in the study is a simple/partial mediation model, and the number of categories, the selection of the number of observed variables, the sample size, and the prior assignment of values for the path coefficients were determined by using similar studies in the literature (e.g., Forero and Maydeu-Olivares, 2009; Li, 2021; Yang-Wallentin et al., 2010). The model of the study is shown in Figure 1.

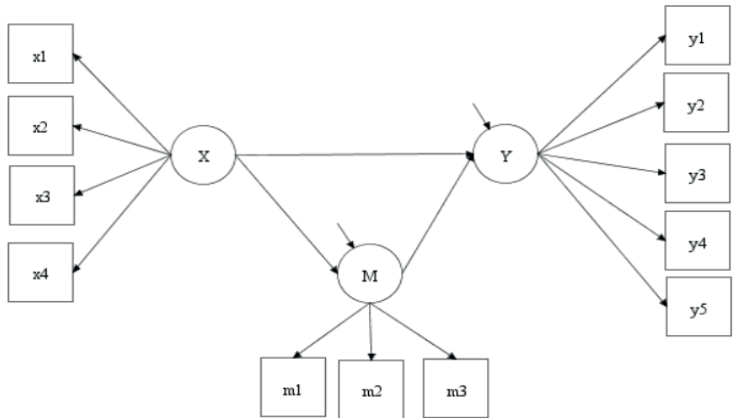


Figure 1. Path Diagram of the model

In the model, the path coefficients were previously assigned from X to Y as 0.3, from X to M as 0.6, and from M to Y as 0.7. With this assignment, the indirect effect was obtained as 0.42. No assignment was made for factor loadings. The number of categories was determined as five. The sample size was determined as 100, 200, 300, 400, and 500 units.

Methods

The Collection of the Data

The data generation process was carried out by taking into account the entire sample volume and the previously assigned values of the model via the R program. After the data were generated continuously in the R program, they were converted into five-category data.

Statistical Analysis

The techniques used by the estimation methods determined for mediation analysis in calculating the indirect effect are as follows. While WLSMV and DWLS estimation techniques use the Delta method, ML and OLS use the Bootstrapping method. This selection process is again the method that the existing package programs assign the estimation technique by default.

In measuring the performances of WLSMV, DWLS, ML and OLS estimation techniques, the closeness of the path coefficient values to the previously assigned values was examined from the perspective of the sample size. In measuring the closeness, mean absolute bias (MAB) and relative bias (RB) values were examined. For the evaluation of information where biases are considered, based on previous simulation studies (Curran et al.,

1996; Bandalos, 2002; Flora and Curran, 2004; Kaplan, 1989; Yang-Wallentin et al., 2010), it was considered that a relative bias of less than 5% is insignificant, a bias between 5-10% indicates a moderate level of bias, and a bias greater than 10% indicates a significant bias. Since the true value and the estimated value will be compared in all program findings, the unstandardized coefficients are taken into account and included in the tables. However, the standardized coefficients are included in the path diagram according to the sample size of the model.

RESULTS

The data generation process for the simulation study was successfully achieved for all sample sizes in the R program. In the current order, analyses were performed for all sample sizes for the model in all programs and the findings were determined. The outputs of Mplus, Jamovi, Amos and Spss, as well as the outputs of MAB and RB, will be given in this section. First, the outputs of the Mplus program were obtained as follows.

Outputs of the Mplus Program

The analysis findings starting from a sample size of 100 units up to 500 units are as follows.

Table 1. Mplus sample size 100

Sample	Estimator	Path Coefficient : True Value	Estimated Value (S.E)	p Value (%2.5 Low- %2.5Upper Limit)	Difference
100	WLSMV	X -> M: 0.6	0.759 (0.159)	0.000 (0.495- 1.125)	0.159
100	WLSMV	M -> Y: 0.7	0.761 (0.191)	0.000 (0.467- 0.196)	0.061
100	WLSMV	X -> Y: 0.3	0.033 (0.222)	0.882 (-0.452- 0.389)	-0.267
100	WLSMV	X-> M -> Y: 0.42	0.578 (0.222)	0.009 (0.307- 1.131)	0.158
Model Fit	CFI	TLI	SRMR	RMSEA(%90 CI and RMSEA Probability<=.05)	χ^2 /sd (p value)
	1.000	1.016	0.037	0.000 (0.000- 0.041 and 0.975)	41.804/51 (0.8172)

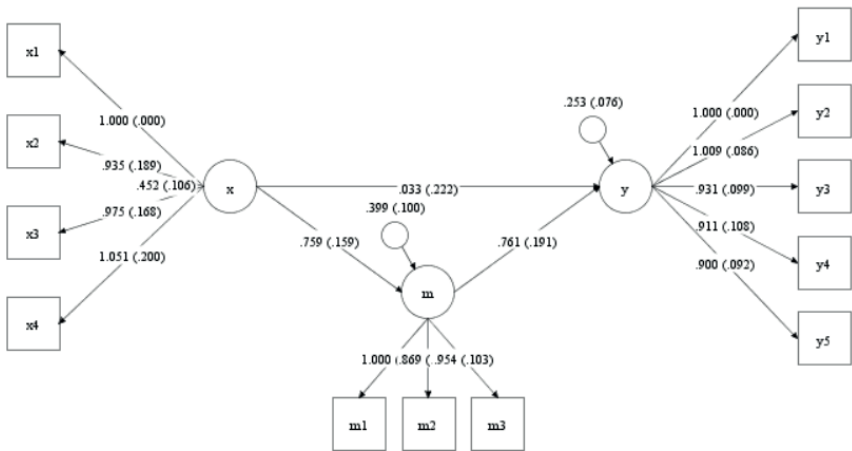


Figure 2. Mplus path diagram N= 100

It is seen that the model fit values are quite good (Schermelele-Engel et al., 2003). When the path coefficients considered in the study are examined, it is seen that all paths are significant except the direct effect of X on Y. This situation transforms the model from a partial mediation model to a full mediation model. In this case, the sample size of 100 units is not sufficient for the partial mediation model. In addition, overestimation is observed in all paths except the direct effect of X on Y.

Table 2. Mplus sample size 200

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (%2.5 Low- %2.5Upper Limit)	Difference
200	WLSMV	X -> M: 0.6	0.509 (0.085)	0.000 (0.351- 0.685)	-0.091
200	WLSMV	M -> Y: 0.7	0.776 (0.121)	0.000 (0.562- 1.040)	0.076
200	WLSMV	X -> Y: 0.3	0.045 (0.105)	0.666 (-0.168- 0.245)	-0.255
200	WLSMV	X-> M -> Y: 0.42	0.394 (0.087)	0.000 (0.252- 0.590)	-0.026
Model Fit	CFI	TLI	SRMR	RMSEA(%90 CI and RMSEA Probability<=.05)	χ^2 /sd (p value)
	1.000	1.004	0.030	0.000 (0.000- 0.038 and 0.992)	46.283/51 (0.6612)

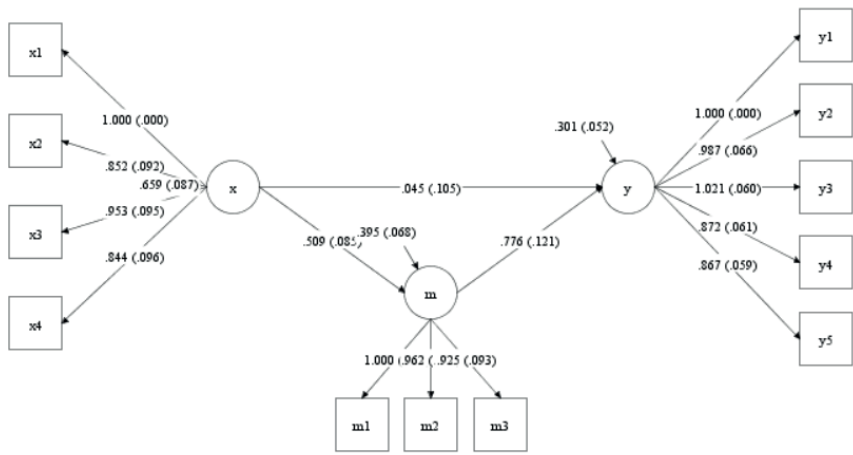
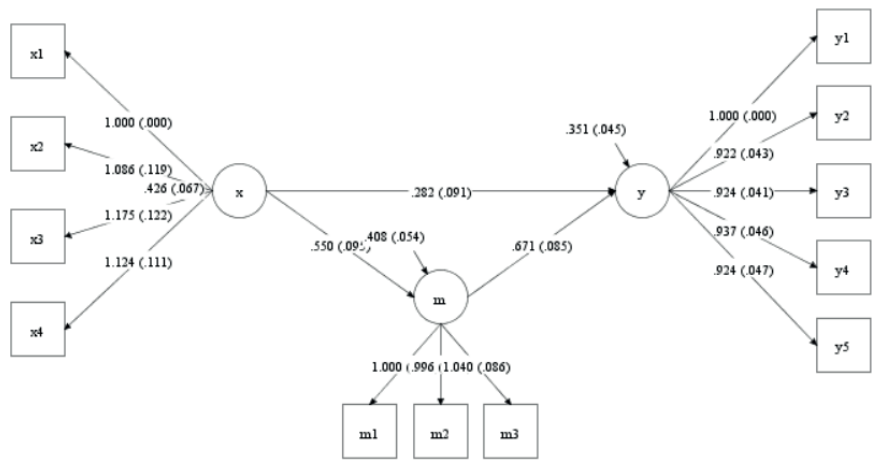


Figure 3. Mplus path diagram N= 200

At 200 units, model fit values were again obtained quite well, while the direct effect of X on Y was obtained as insignificant. In this case, as in the 100-unit sample size, the model has moved from a partial mediation model to a full mediation model. We can say that the partial mediation model is not sufficient at 200 units. In addition, instead of an overestimation as in 100 units, there is underestimation at this volume, except for the direct effect of M on Y.

Table 3. Mplus sample size 300

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (%2.5 Low- %2.5Upper Limit)	Difference
300	WLSMV	X -> M: 0.6	0.550 (0.095)	0.000 (0.378- 0.755)	-0.05
300	WLSMV	M -> Y: 0.7	0.671 (0.085)	0.000 (0.506- 0.841)	-0.029
300	WLSMV	X -> Y: 0.3	0.282 (0.091)	0.002 (0.107- 0.466)	-0.018
300	WLSMV	X-> M -> Y: 0.42	0.369 (0.074)	0.000 (0.244- 0.537)	-0.051
Model Fit	CFI	TLI	SRMR	RMSEA(%90 CI and RMSEA Probability<=.05)	χ^2 /sd (p value)
	1.000	1.007	0.021	0.000 (0.000- 0.016 and 1.000)	38.324/51 (0.9049)



It is seen that there is a significant difference except for the linear effect of X on Y at 300 units ($p>0.001$) and the model fit values are quite good. Underestimation was observed in all path coefficients. We can say that a sample size of 300 units shows a borderline adequacy for the partial mediation model.

Table 4. Mplus sample size 400

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (%2.5 Low- %2.5Upper Limit)	Difference
400	WLSMV	X -> M: 0.6	0.566 (0.086)	0.000 (0.411- 0.750)	-0.034
400	WLSMV	M -> Y: 0.7	0.534 (0.061)	0.000 (0.422- 0.665)	-0.166
400	WLSMV	X -> Y: 0.3	0.284 (0.069)	0.000 (0.148- 0.419)	-0.016
400	WLSMV	X-> M -> Y: 0.42	0.303 (0.055)	0.000 (0.209- 0.424)	-0.117
Model Fit	CFI	TLI	SRMR	RMSEA(%90 CI and RMSEA Probability<=.05)	χ^2 /sd (p value)
	0.998	0.998	0.023	0.017 (0.000- 0.037 and 0.999)	56.904/51 (0.2647)

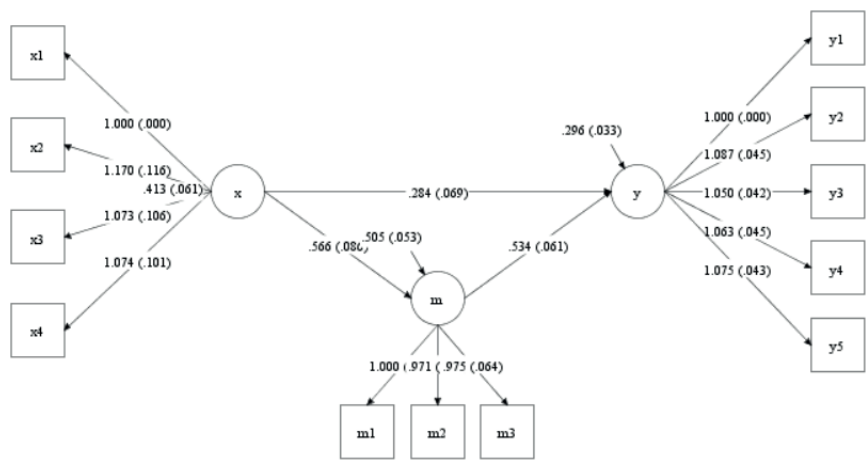


Figure 5. Mplus path diagram N= 400

It is seen that all path coefficients of the model are significant at 400 units. The fit values are at a level that can be seen as a result of using real data sets. In addition, there is underestimation.

Table 5. Mplus sample size 500

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (%2.5 Low- %2.5Upper Limit)	Difference
500	WLSMV	X -> M: 0.6	0.565 (0.059)	0.000 (0.457- 0.685)	-0.035
500	WLSMV	M -> Y: 0.7	0.576 (0.078)	0.000 (0.427- 0.737)	-0.124
500	WLSMV	X -> Y: 0.3	0.265 (0.072)	0.000 (0.121- 0.410)	-0.035
500	WLSMV	X-> M -> Y: 0.42	0.325 (0.050)	0.000 (0.236- 0.432)	-0.095
Model Fit	CFI	TLI	SRMR	RMSEA(%90 CI and RMSEA Probability<=.05)	χ^2 /sd (p value)
	0.996	0.994	0.022	0.028 (0.008- 0.042 and 0.996)	70.625/51 (0.0357)

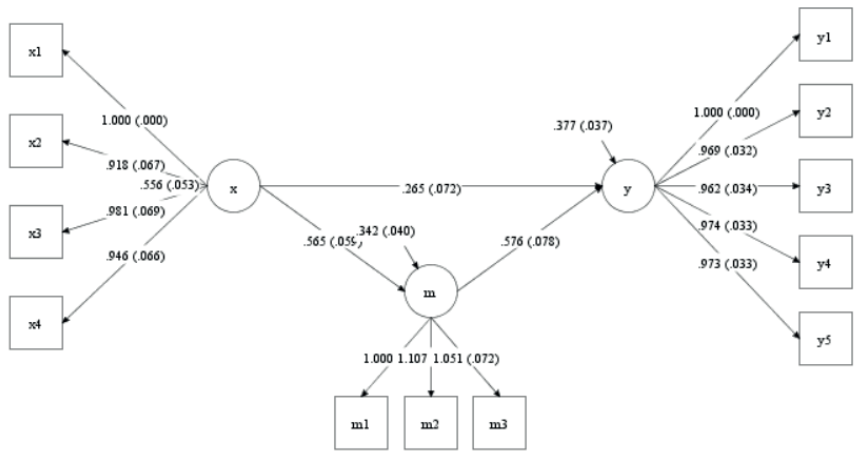


Figure 6. Mplus path diagram N= 500

At 500 units, the path coefficients were again found to be significant, but while the underestimation continued, a decrease in this value was observed with the increase in the sample size. In addition, an increase in the χ^2 value from the fit values was clearly observed with the increase in the sample size and had the highest value at 500 units. The R program was used for the visuals of the MAB and RB values to determine the closeness between the estimated values obtained from the program and the real values. The findings regarding the sample sizes are given in Figures 7 and 8.

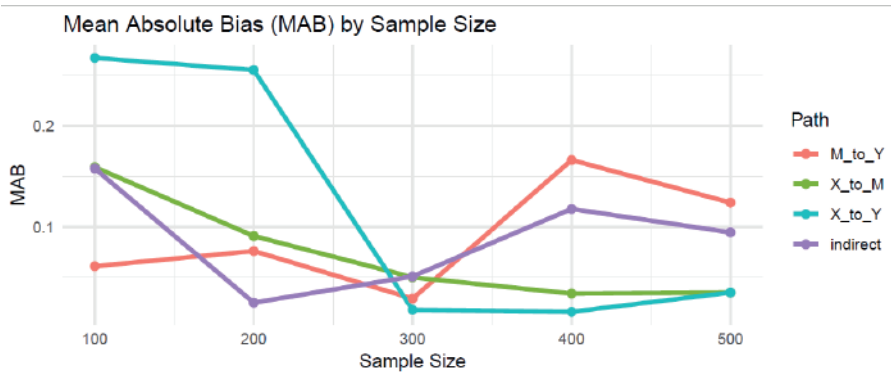


Figure 7. Mplus MAB value

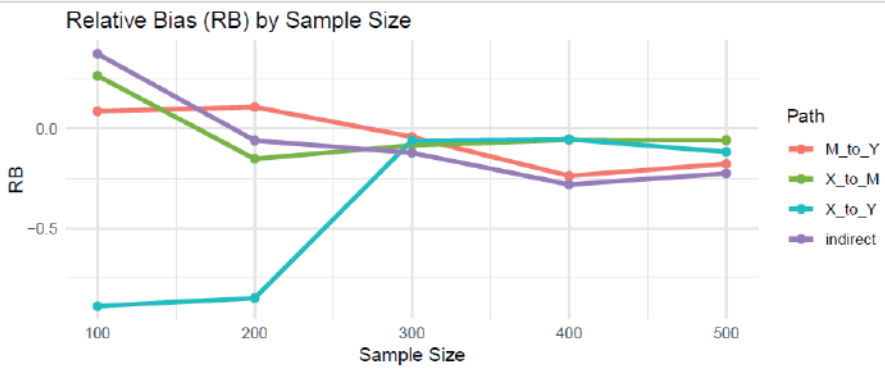


Figure 8. Mplus RB value

When the MAB and RB graphs are examined, while there is a very good result in the 300-unit sample size, this situation did not improve as the sample size increased. We can state that the lowest bias values were observed in 300 units. While the RB values in 400 and 500 units again indicate low, insignificant bias, this situation resulted in a greater deviation in the MAB values.

Outputs of the Jamovi Program

The Jamovi program uses the DWLS estimation technique for the model with the semLJ module and the Delta method for calculating indirect effects. The estimation technique selection process is generally specified in this program as in the Mplus program with the use of ordered categorical data in programs. In addition, although the Jamovi program allows for the selection of more than one estimation technique, for example, instead of the WLSMV method that was tried to be selected for this model, the DWLS technique was used. The findings obtained for all sample sizes were as follows.

Table 6. Jamovi sample size 100

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (Low-Upper Limit)	Difference
100	DWLS	X -> M: 0.6	0.7594 (0.1311)	<.001 (0.503-1.016)	0.1594
100	DWLS	M -> Y: 0.7	0.7609 (0.159)	<.001 (0.450-1.072)	0.0609
100	DWLS	X -> Y: 0.3	0.0328 (0.174)	0.850 (-0.308-0.373)	-0.2672
100	DWLS	X-> M -> Y: 0.42	0.578 (0.164)	<.001 (0.257-0.899)	0.158
Model Fit	Scaled CFI	Scaled TLI	Scaled SRMR	Scaled RMSEA(p value)	χ^2 /sd (p value)
	1.000	1.025	0.037	0.000 (0.976)	41.5/51 (0.825)

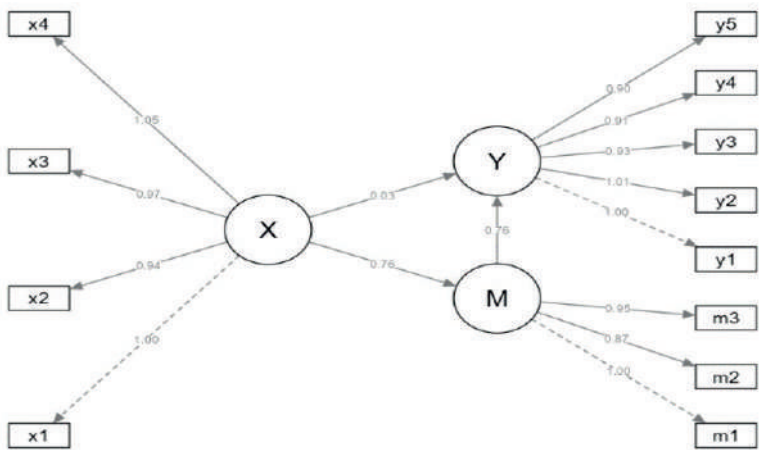


Figure 9. Jamovi path diagram N=100

Except for the path coefficient of X on Y in 100 units, all other paths are both significant and overestimated. Model fit values are quite good (Schermelleh-Engel et al., 2003). With the direct effect being insignificant, the partial mediation model turns into a full mediation model. In this case, it is said that a sample size of 100 units will not be sufficient for a partial mediation model. In addition, the outputs of the Mplus program in 100 units and the outputs of Jamovi gave very close values.

Table 7. Jamovi sample size 200

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (Low-Upper Limit)	Difference
200	DWLS	X -> M: 0.6	0.5085(0.0798)	< .001 (0.352-0.665)	-0.0915
200	DWLS	M -> Y: 0.7	0.7755 (0.1140)	<.001 (0.552-0.999)	0.07555
200	DWLS	X -> Y: 0.3	0.0452 (0.1002)	0.652 (-0.151-0.242)	-0.2548
200	DWLS	X-> M -> Y: 0.42	0.394 (0.084)	<.001 (0.230-0.558)	-0.026
Model Fit	Scaled CFI	Scaled TLI	Scaled SRMR	Scaled RMSEA(p value)	χ^2 /sd (p value)
	1.000	1.004	0.030	0.000 (0.992)	46.1/51 (0.668)

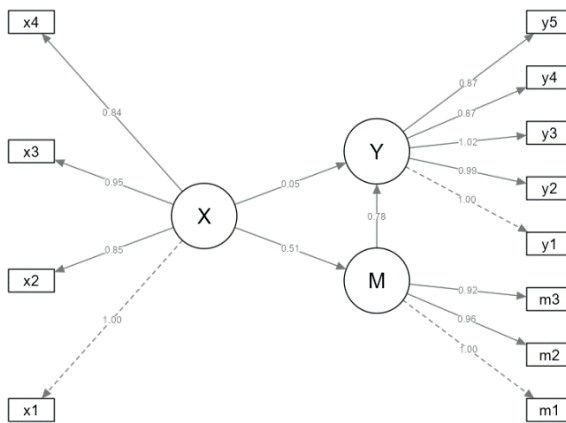


Figure 10. Jamovi path diagram N=200

The direct effect of X on Y in 200 units is also obtained with insignificant and incomplete estimation. This situation transforms the partial mediation model into full mediation. It is seen that the other path coefficients are significant. There is an increase in the χ^2 value from the model fit values. It cannot be said that 200 units are sufficient for the partial mediation model.

Table 8. Jamovi sample size 300

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (Low- Upper Limit)	Difference
300	DWLS	X -> M: 0.6	0.550 (0.0899)	<.001 (0.374- 0.727)	-0.05
300	DWLS	M -> Y: 0.7	0.671 (0.0844)	<.001 (0.506- 0.837)	-0.029
300	DWLS	X -> Y: 0.3	0.282 (0.0877)	0.001 (0.110- 0.454)	-0.018
300	DWLS	X-> M -> Y: 0.42	0.369 (0.070)	<.001 (0.231- 0.507)	-0.051
Model Fit	Scaled CFI	Scaled TLI	Scaled SRMR	Scaled RMSEA(p value)	χ^2 /sd (p value)
	1.000	1.007	0.021	0.000 (1.000)	38.2/51 (0.907)

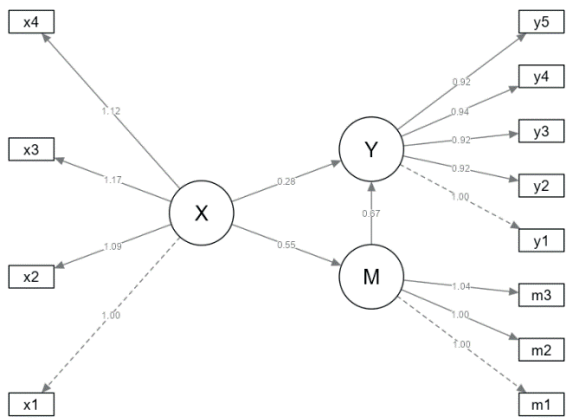


Figure 11. Jamovi path diagram N=300

In the 300-unit sample size, all path coefficients of the model were obtained with significant and incomplete estimation. A decrease is observed in the χ^2 value of the model fit values. When all data are evaluated, it can be said that the model's prediction and fit values are good, together with the closeness to the previously assigned values. This situation suggests that it will be improved even more as the sample size increases, due to the asymptotic theory.

Table 9. Jamovi sample size 400

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (Low-Upper Limit)	Difference
400	DWLS	X -> M: 0.6	0.566 (0.0803)	<.001 (0.409-0.724)	-0.034
400	DWLS	M -> Y: 0.7	0.534 (0.0604)	<.001 (0.416-0.652)	-0.166
400	DWLS	X -> Y: 0.3	0.284 (0.0691)	<.001 (0.149-0.420)	-0.016
400	DWLS	X-> M -> Y: 0.42	0.303 (0.053)	<.001 (0.199-0.406)	-0.117
Model Fit	Scaled CFI	Scaled TLI	Scaled SRMR	Scaled RMSEA(p value)	χ^2 /sd (p value)
	1.000	1.004	0.023	0.017 (0.999)	56.8/51 (0.268)

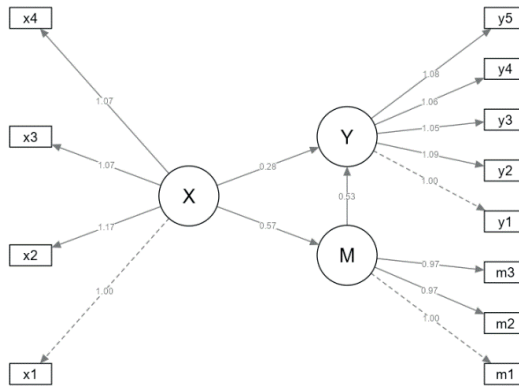


Figure 12. Jamovi path diagram N=400

All path coefficients are significant and underestimation is observed in all of them in 400 units. In addition, there is a decrease in model fit values. This situation may actually be due to the extremely good fit values in 300 units. Because the fit values in 400 units are actually seen as the highest fit values that can be made with real data sets. In this case, the decrease in model fit values actually seems reasonable.

Table 10. Jamovi sample size 500

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (Low- Upper Limit)	Difference
500	DWLS	X -> M: 0.6	0.565 (0.0581)	<.001 (0.451- 0.679)	-0.035
500	DWLS	M -> Y: 0.7	0.576 (0.0744)	<.001 (0.430- 0.722)	-0.124
500	DWLS	X -> Y: 0.3	0.265 (0.0699)	<.001 (0.128- 0.402)	-0.035
500	DWLS	X-> M -> Y: 0.42	0.325 (0.049)	<.001 (0.230- 0.421)	-0.095
Model Fit	Scaled CFI	Scaled TLI	Scaled SRMR	Scaled RMSEA(p value)	χ^2 /sd (p value)
	1.000	1.002	0.022	0.028 (0.996)	70.5/51 (0.037)

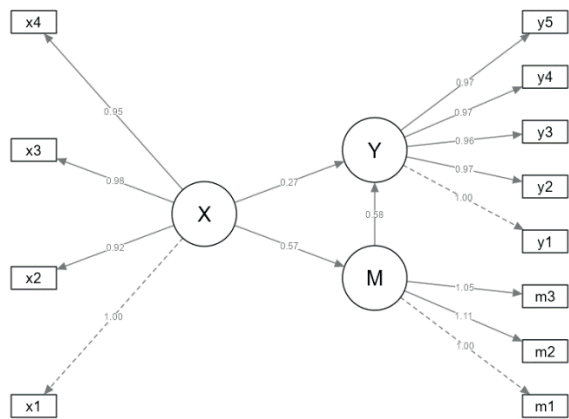


Figure 13. Jamovi path diagram N=500

All paths are significant and all have underestimation at 500 units. An increase in the χ^2 value from the model fit values occurred at this volume with 300 units. Model improvements do not occur after 300 units with the use of simulation data. This situation may be due to randomly obtained data. This situation can actually be clarified with real data sets. However, the Jamovi program, like the Mplus program, made its best estimates at 300 units.

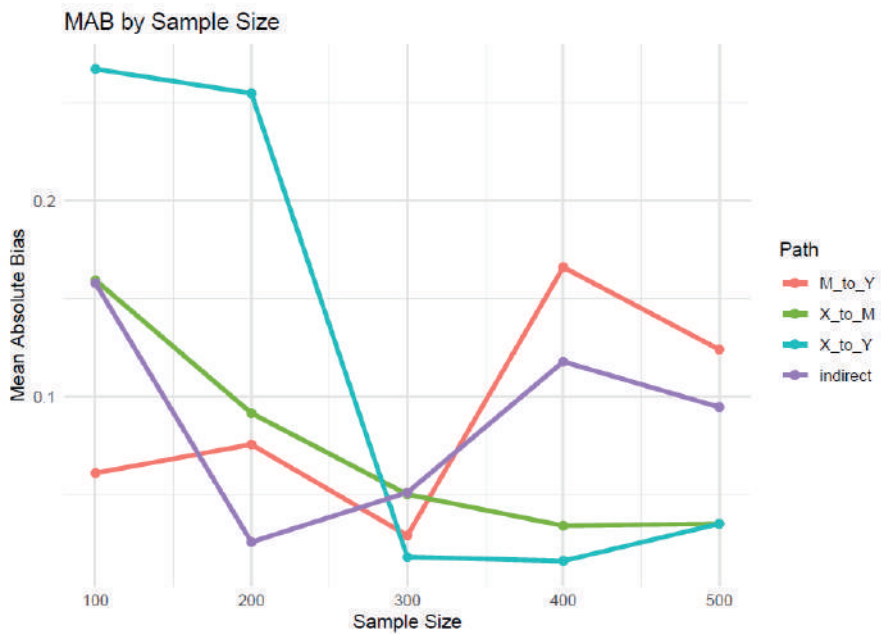


Figure 14. Jamovi MAB value

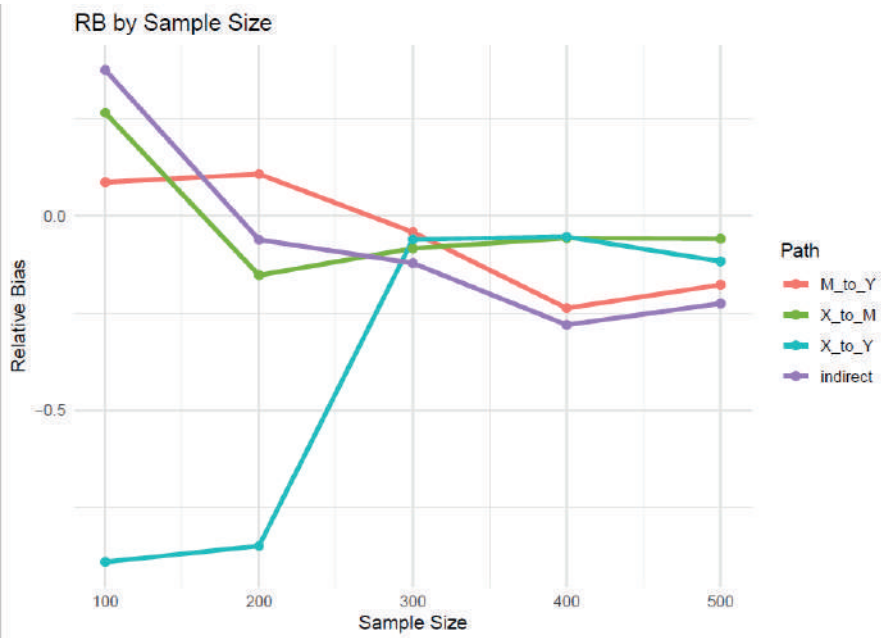


Figure 15. Jamovi RB value

As in the Mplus program, it is seen that the least bias values are obtained in the Jamovi program at 300 units. There are deteriorations at 400 and 500 units, but there is still a relative bias of less than 5%. This is an insignificant bias according to the threshold values specified in the literature. However, due to the asymptotic theory, further improvements were expected, but this did not happen. Although the MAB values also showed an increase at 400 and 500 units, even this increase did not approach the maximum value of 2%. In fact, everything seems to be fine at 300 units and above. However, it does not happen exactly as expected.

Outputs of the Amos Program

This program, which is frequently used by many application researchers in SEM studies, uses the ML estimation technique by default. This technique, which works with continuous data and multivariate normality assumption, is unfortunately also frequently used with categorical data. Although the Amos program allows selection in the case of categorical data, this situation is seen as a very technical issue for application researchers. Due to this situation, the partial mediation relationship was examined with the Amos program using the ML estimation technique. In fact, the use of the ULS technique is what should have been done, but since application researchers continued with the default estimation technique, ML was selected in the estimation of the model. This selection reflects the purpose of the study. Because with this research, it was aimed to carry out a study that could be an example for application researchers by not intensively addressing estimation techniques and theoretical knowledge. The findings obtained are as follows:

Table 11. Amos sample size 100

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (Low-Upper Limit)	Difference	CFI	TLI	RMSEA	χ^2 /sd
100	ML	X -> M: 0.6	0.720 (0.185)	***	0.120	1.000	1.008	0.000	0.952
100	ML	M -> Y: 0.7	0.770 (0.185)	***	0.07				
100	ML	X -> Y: 0.3	0.076 (0.185)	0.68	-0.224				
100	ML	X-> M -> Y: 0.42	0.555 (0.207)	(0.288-1.088)	0.135				

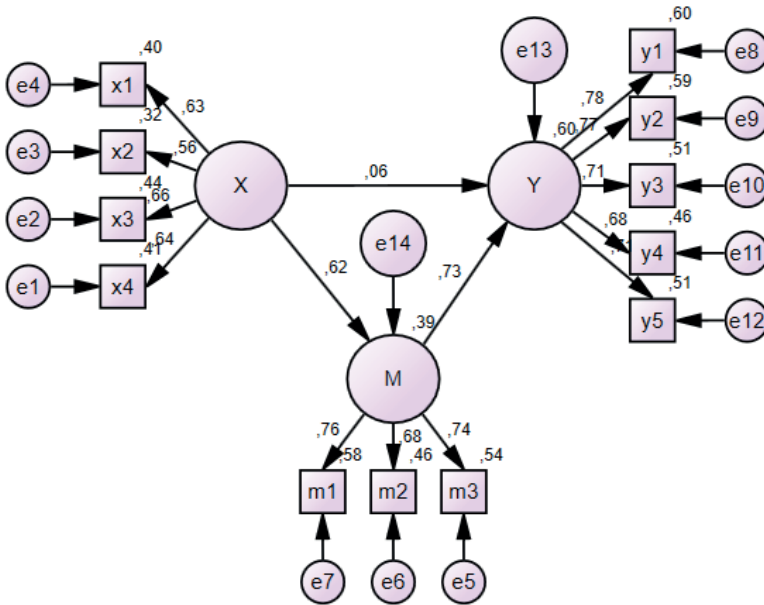


Figure 16. Amos path diagram N=100

The evaluation in 100 units is quite similar to those in the Mplus and Jamovi programs. In this program, the direct effect of X on Y was found insignificant and the model changed from partial mediation to full mediation. In this case, it actually says that many other issues should not be considered. Because the assumed model structure has changed and the interpretation and explanation of the relationships established in the literature have become quite difficult. However, the model fit values are quite good as in other programs (Schermelleh-Engel et al., 2003). The direct effect of X on Y was obtained as 0.033 (0.222), 0.0328 (0.174), 0.076 (0.185) in the models, respectively. An overestimation is observed in the ML and Amos programs. In the current situation, we can say that the WLSMV and DWLS techniques, which are suitable techniques for categorical data, make less estimates than the ML used for continuous variables.

Table 12. Amos sample size 200

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (Low-Upper Limit)	Difference	CFI	TLI	RMSEA	χ^2 /sd
200	ML	X -> M: 0.6	0.549 (0.108)	***	-0.051	1.000	1.011	0.000	0.856
200	ML	M -> Y: 0.7	0.840 (0.148)	***	0.140				
200	ML	X -> Y: 0.3	0.047 (0.112)	0.674	-0.253				
200	ML	X-> M -> Y: 0.42	0.461 (0.116)	(0.277-0.725)	0.041				

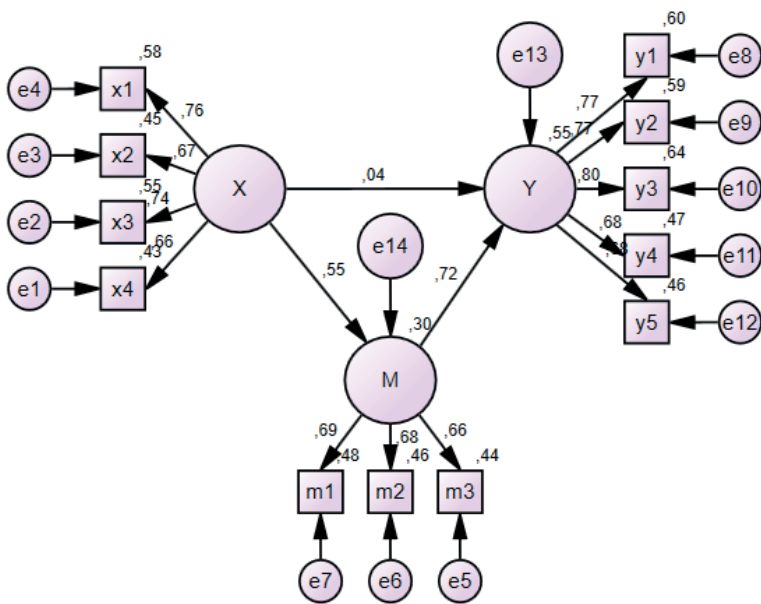


Figure 17. Amos path diagram N=200

At 200 units, it has the same results as the first two programs. The effect of X on Y was insignificant. Underestimation started again after 100 units. The model fit values are quite good.

Table 13. Amos sample size 300

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (Low-Upper Limit)	Difference	CFI	TLI	RMSEA	χ^2 /sd
300	ML	X -> M: 0.6	0.497 (0.085)	***	-0.103	1.000	1.003	0.000	0.934
300	ML	M -> Y: 0.7	0.663 (0.097)	***	-0.037				
300	ML	X -> Y: 0.3	0.241 (0.084)	0.004	-0.059				
300	ML	X-> M -> Y: 0.42	0.330 (0.062)	(0.222-0.466)	-0.09				

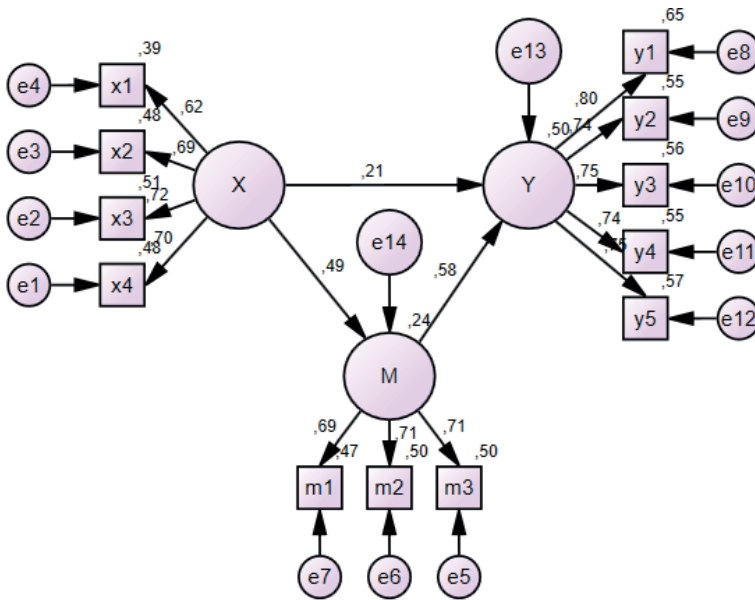


Figure 18. Amos path diagram N=300

The direct effect of X on Y is $p > 0.001$ and is significant for 5% and 0.1%. We see a result close to this situation in the best Mplus program and the next closest result in the Jamovi program outputs. All path coefficients are significant and underestimation is observed in all of them. The fit values are again quite good.

Table 14. Amos sample size 400

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (Low-Upper Limit)	Difference	CFI	TLI	RMSEA	χ^2 /sd
400	ML	X -> M: 0.6	0.522 (0.83)	***	-0.048	0.992	0.990	0.027	1.282
400	ML	M -> Y: 0.7	0.545 (0.069)	***	-0.155				
400	ML	X -> Y: 0.3	0.266 (0.072)	***	-0.034				
400	ML	X-> M -> Y: 0.42	0.282 (0.054)	(0.191-0.400)	-0.138				

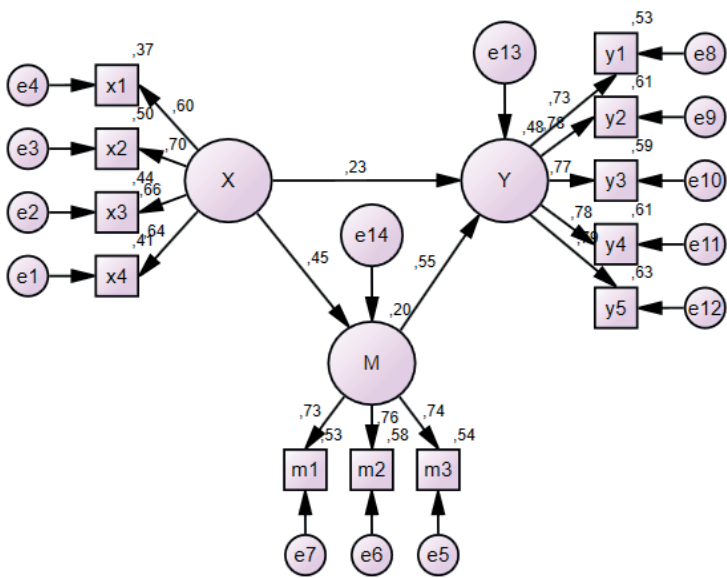


Figure 19. Amos path diagram N=400

Model fit values at 400 units are seen to have turned into the best fit values with real data. However, the proximity to real values is still not at the desired level. This situation is similar to other program outputs.

Table 15. Amos sample size 500

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (Low-Upper Limit)	Difference	CFI	TLI	RMSEA	χ^2 /sd
500	ML	X -> M: 0.6	0.616 (0.070)	***	0.016	0.991	0.989	0.029	1.415
500	ML	M -> Y: 0.7	0.539 (0.075)	***	-0.161				
500	ML	X -> Y: 0.3	0.265 (0.074)	***	-0.035				
500	ML	X-> M -> Y: 0.42	0.332 (0.054)	(0.232-0.448)	-0.088				

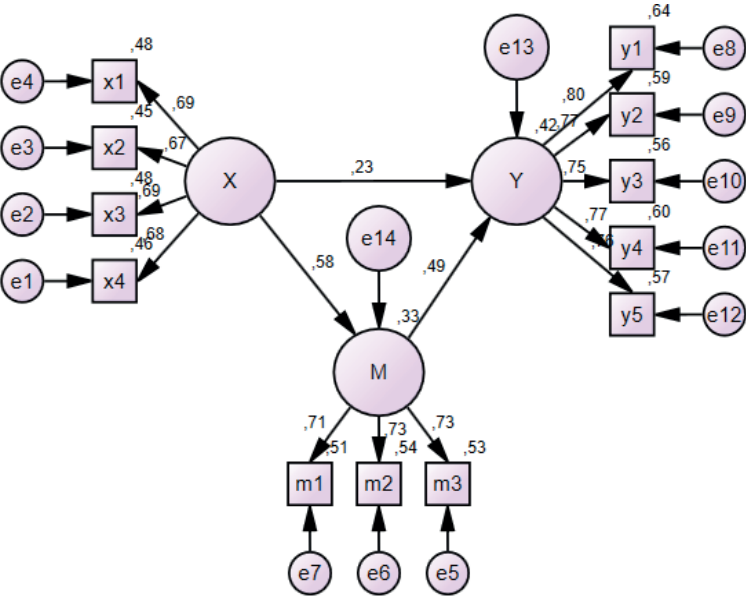


Figure 20. Amos path diagram N=500

In this volume, the closeness to the real values is a little better. This situation is accompanied by the model fit values. There is a similar result in the other first two programs. We can say that a single path coefficient is farther from the convergence point than the others. The path coefficient from M to Y was assigned as 0.7 as the real value and a good closeness to this value has not been achieved yet.

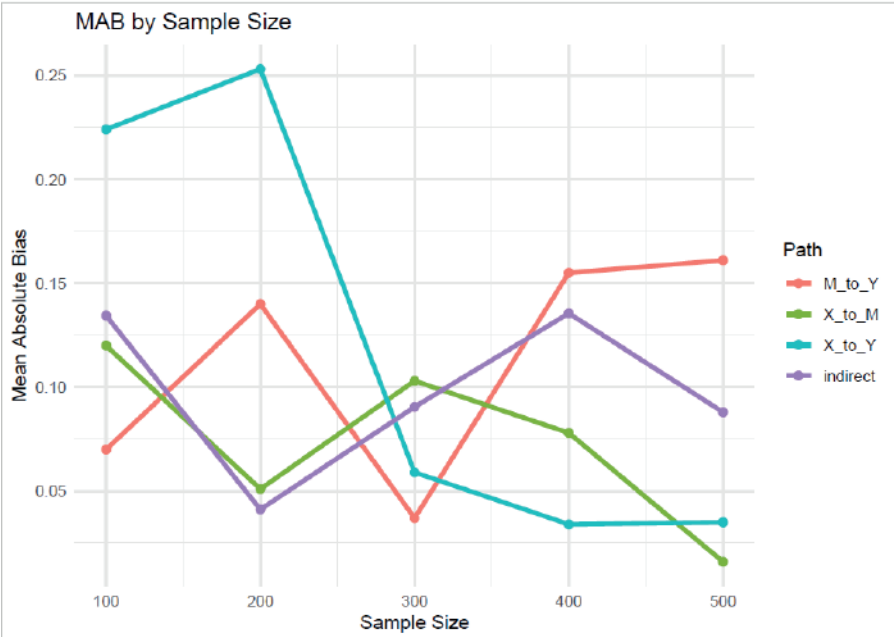


Figure 21. Amos MAB value

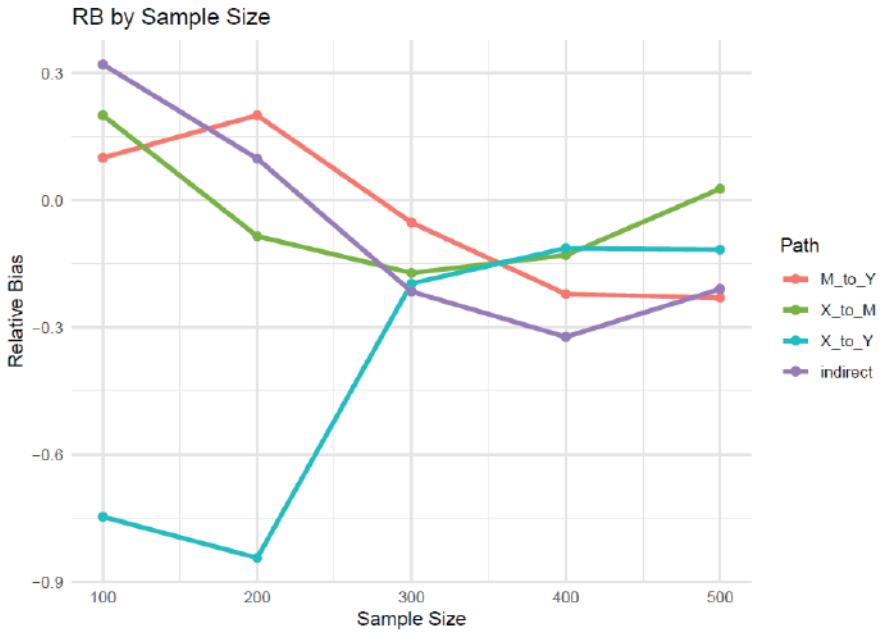


Figure 22. Amos RB value

We can say that the Amos program lags behind the other two programs in terms of bias with ML and Bootstrapping approaches. However, similarly, less bias was obtained in the Amos program at 300 units. However, when looking at the MAB values, a deviation of over 10% was recorded even at 300 units. This situation is over 15% at volumes of 400 and 500 units. Relative bias indicates that the excess seen in the MAB value is slightly less. When we consider all three programs, the least bias order is Mplus, Jamovi and Amos.

Outputs of Spss Program

Spss, Hayes' Process extension and OLS and Bootstrapping approaches were used to estimate the model. In this program, a selection process cannot be performed to determine the estimation techniques. However, unlike other programs, an input must be given according to the models enumerated by Hayes. The technique does not present a structure that can handle the measurement model while working with latent variables. Therefore, there is already a direct weakness. However, its frequent use by practitioners is present in many disciplines. Therefore, this program and technique that will reflect the purpose of the study were considered. Researchers working with latent variables need to specify each latent variable as a single variable when using this program. This eliminates the measurement model and calculates the average values of the latent variables and obtains a single variable. The outputs of the program are given below. Since path diagrams are given in other program outputs, they were also intended to be given here. For this, visuals were obtained through the R program.

Table 16. Spss sample size 100

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (LLCI- ULCI)	Difference
100	OLS	X -> M: 0.6	0.5224 (0.1008)	0.000 (0.3222- 0.7225)	-0.0776
100	OLS	M -> Y: 0.7	0.5208 (0.0847)	0.000 (0.3526- 0.6889)	-0.1792
100	OLS	X -> Y: 0.3	0.1678 (0.0955)	0.082 (-0.0217- 0.3573)	-0.1322
100	OLS	X-> M -> Y: 0.42	0.2720 (0.0631)	(0.1604- 0.4077)	-0.1480
Model Fit	R	R²	Mean Squared Error- MSE	F (df₁, df₂)	p Value
	0.6329	0.4005	0.7728	32.4037(2,97)	0.000

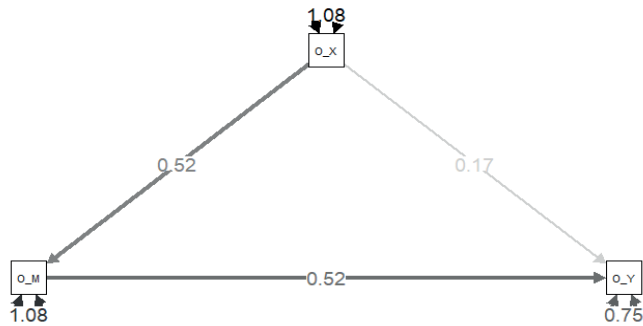


Figure 23. R path diagram N=100

Underestimation at 100 units and the direct effect of X on Y were obtained as insignificant in this program as in the other three programs. It is seen that the other path coefficients are significant. It is seen that the fit value is good. The 100-unit sample size also transformed the partial mediation model into a full mediation model in the Spss program. In addition, the explained variance for Y was obtained as 40.05%.

Table 17. Spss sample size 200

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (LLCI- ULCI)	Difference
200	OLS	X -> M: 0.6	0.4081 (0.0658)	0.000 (0.2785- 0.5378)	-0.1919
200	OLS	M -> Y: 0.7	0.5200 (0.0629)	0.000 (0.3960- 0.6441)	-0.1800
200	OLS	X -> Y: 0.3	0.1436 (0.0636)	0.0251 (0.0181- 0.2690)	-0.1564
200	OLS	X-> M -> Y: 0.42	0.2122 (0.0415)	(0.1373- 0.2983)	-0.2078
Model Fit	R	R²	Mean Squared Error- MSE	F (df₁, df₂)	p Value
	0.5910	0.3492	0.8473	52.8573(2,197)	0.000

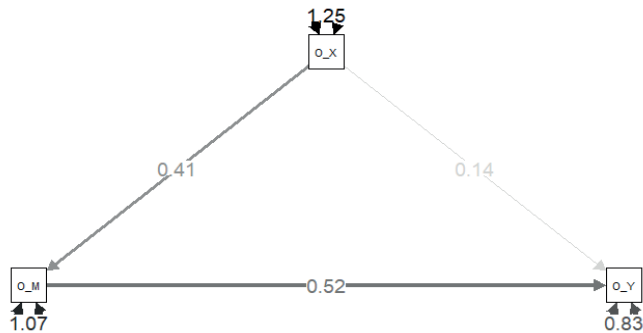


Figure 24. R path diagram N=200

In 200 units, underestimation and the direct effect of X on Y were obtained as insignificant. In this case, the model again turns into a full mediation model. In addition, the model fit value seems good. The explained variance rate for Y decreased to 34.92%.

Table 18. Spss sample size 300

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (LLCI-ULCI)	Difference
300	OLS	X -> M: 0.6	0.3869 (0.0566)	0.000 (0.2755- 0.4982)	-0.2131
300	OLS	M -> Y: 0.7	0.4621 (0.0503)	0.000 (0.3631- 0.5610)	-0.2379
300	OLS	X -> Y: 0.3	0.2474 (0.0528)	0.000 (0.1434- 0.3514)	-0.0526
300	OLS	X-> M -> Y: 0.42	0.1788 (0.0305)	(0.1208- 0.2443)	-0.2412
Model Fit	R	R ²	Mean Squared Error- MSE	F (df ₁ , df ₂)	p Value
	0.5913	0.3496	0.8677	79.8330(2,297)	0.000

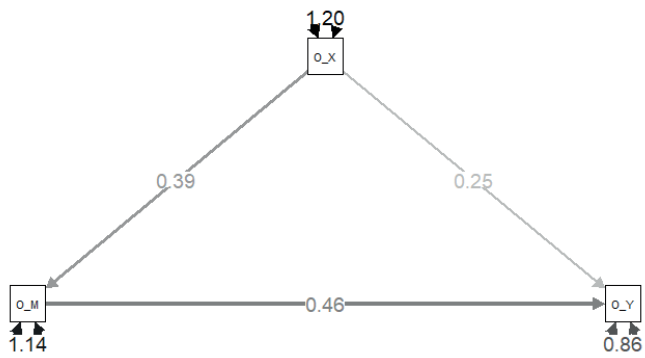


Figure 25. R path diagram N=300

It is seen that all path coefficients are significant in 300 units and underestimation occurs. The model fit value has been well-confirmed. The explained variance for Y is 34.96%, which is quite close to the previous sample size. Even though there are significant path coefficients in 300 units, it is obvious that they are quite far from the real values and also far from the estimated values obtained in 300 units in the other three programs.

Table 19. Spss sample size 400

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (LLCI- ULCI)	Difference
400	OLS	X -> M: 0.6	0.3791 (0.0522)	0.000 (0.2764- 0.4817)	-0.2209
400	OLS	M -> Y: 0.7	0.4626 (0.0423)	0.000 (0.3794- 0.5458)	-0.2374
400	OLS	X -> Y: 0.3	0.2493 (0.0469)	0.000 (0.1571- 0.3416)	-0.0507
400	OLS	X-> M -> Y: 0.42	0.1754 (0.0281)	(0.1225- 0.2323)	-0.2446
Model Fit	<i>R</i>	<i>R</i> ²	Mean Squared Error- MSE	<i>F</i> (<i>df</i> ₁ , <i>df</i> ₂)	p Value
	0.5901	0.3482	0.8860	106.0592(2,397)	0.000

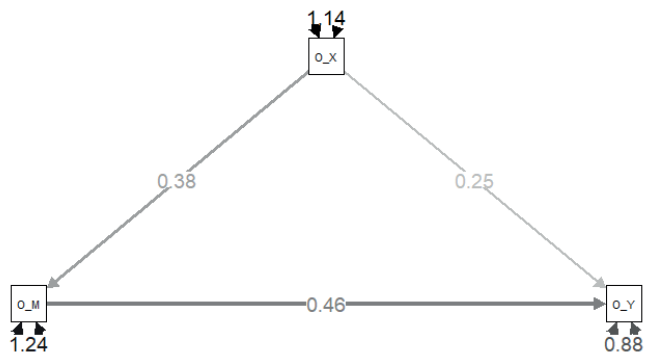


Figure 26. R path diagram N=400

The model fit is good at 400 units, the path coefficients are significant but far from the true values. Even in this state, 34.82% of the variance in Y is explained. When looking at other program outputs, it can be observed that the distance of the estimated values from the true values is not at this size.

Table 20. Spss sample size 500

Sample	Estimator	Path Coefficient :True Value	Estimated Value (S.E)	p Value (LLCI- ULCI)	Difference
500	OLS	X -> M: 0.6	0.4713 (0.0429)	0.000 (0.3889- 0.5573)	-0.1269
500	OLS	M -> Y: 0.7	0.4016 (0.0412)	0.000 (0.3206- 0.4826)	-0.2984
500	OLS	X -> Y: 0.3	0.2589 (0.0440)	0.000 (0.1725- 0.3454)	-0.0411
500	OLS	X-> M -> Y: 0.42	0.1900 (0.0254)	(0.1423- 0.2432)	-0.2300
Model Fit	<i>R</i>	<i>R</i> ²	Mean Squared Error- MSE	<i>F</i> (<i>df</i> ₁ , <i>df</i> ₂)	p Value
	0.5579	0.3112	0.9350	112.2813(2,497)	0.000

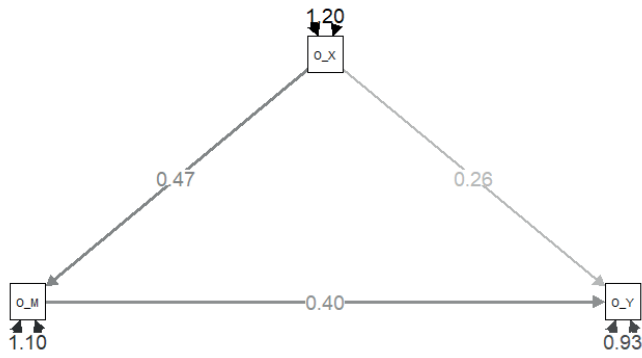


Figure 27. R path diagram N=500

The variance explained for Y in 500 units has a decrease of 31.12% and the model fit is good but the closeness to the real values is quite weak. When we look at the closeness to the real value in 500 units in other programs, we can say that Spss is quite behind at this point. In fact, this situation is present in the entire sample size.

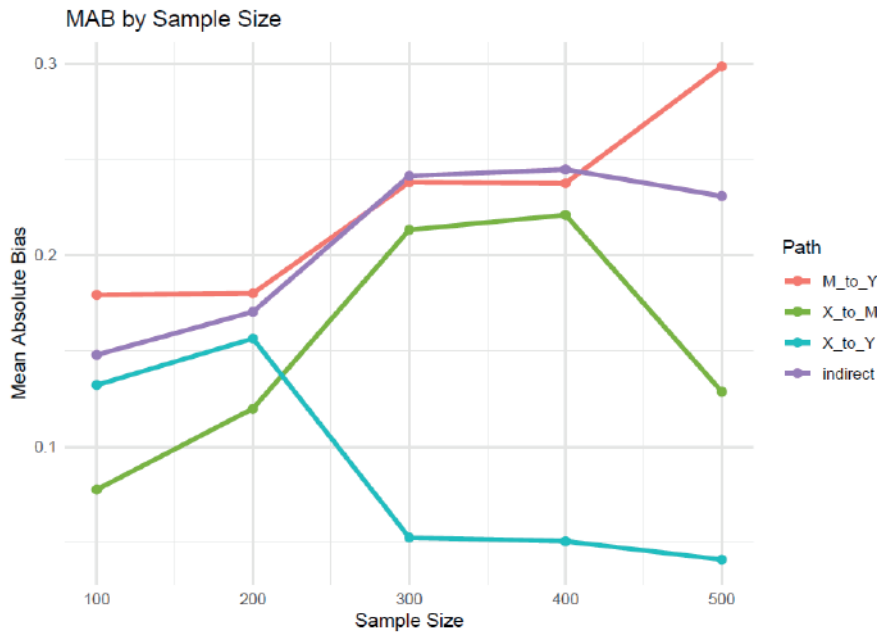


Figure 28. Spss MAB value

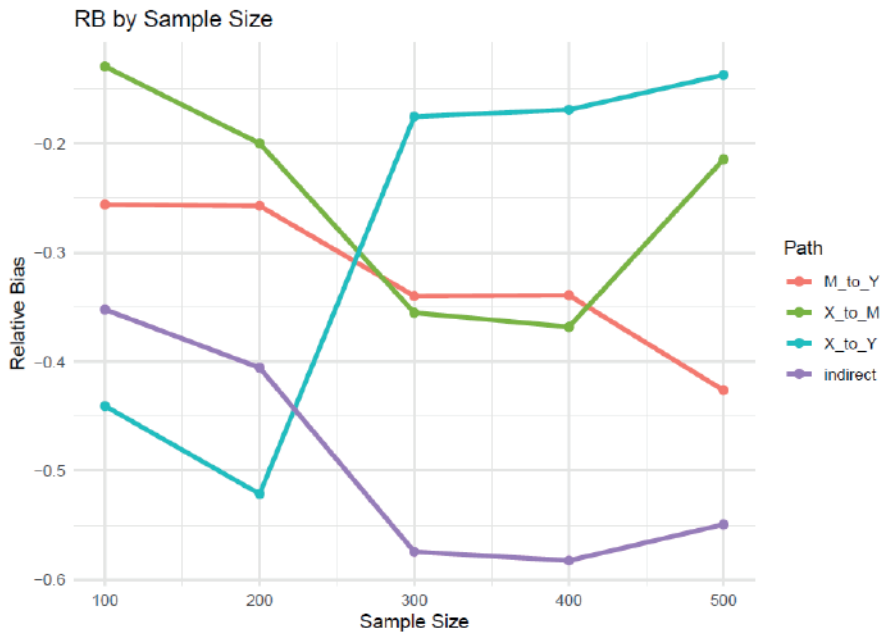


Figure 29. Spss RB value

In the simulation study conducted with the Hayes plugin of the Spss program, biases not seen in other programs were detected. While the other three programs gave the least bias in 300 units, they also showed low biases in the other 400 and 500 units. However, Spss produced estimation results that we can call quite weak together with OLS and Bootstrapping. This situation may be due to the model not addressing the measurement model. Relative biases have quite high deviations such as -60%. It is obvious that this result will lead to incorrect estimation of the model by being far from the threshold values given in the literature. The path coefficients did not follow a certain trend in any sample size and gave a rather scattered deviation image. The MAB values also indicate the same high deviations as the RB values, in other words, biases.

DISCUSSION AND CONCLUSION

This study aims to measure the performances of Mplus, Jamovi, Amos and Spss package programs, which are frequently used by applied researchers for the partial mediation model. The study aimed to attract the attention of applied researchers through the names of the package programs instead of an intensive process with theoretical information. The reason for this situation is that the examination of the performances of WLSMV, DWLS, ML and

OLS estimation techniques is likely to be seen as a more theoretical study by applied researchers. The simple partial mediation model (Çağlayan and Özenç, 2024; Hayes, 2018) was determined as a model frequently used by applied researchers. In assigning the real value for the model, it was thought that the model could transform into a different model in low sample sizes by assigning the $X \rightarrow Y$ low path coefficient value, which is the most important of the direct effects. Previous studies (Forero and Maydeu-Olivares, 2009; Li, 2021; Yang-Wallentin et al., 2010) were used for the previously assigned values in the simulation. In addition, categorical data were obtained for the model in the data generation process. Only path coefficients were examined as a performance measure in the study. The study has many limitations as such. For example, standard errors, model fit values, missing data, and skewed data that do not provide normality were not addressed.

The sample size used in the study is not like those used in similar simulation studies (Li, 2021). The reason for this situation is the desire to obtain performance findings in sample sizes frequently used by applied researchers. The deviations between the previously assigned path coefficients and the findings of the estimation techniques, in other words, biases, were addressed with MAB and RB values. In interpreting these values, studies in the literature (Curran et al., 1996; Bandalos, 2002; Flora & Curran, 2004; Kaplan, 1989; Yang-Wallentin et al., 2010) were used as a basis.

The analysis findings based on the sample size based on path coefficients indicate the following: In all programs at 100 and 200 units, the model changes from a partial mediation model to a full mediation model. The issue that causes this situation is that the direct effect of X on Y is assigned as low. This situation can also be observed in studies conducted with real data, so this issue is very important. Practical researchers need to pay attention to this critical path coefficient. Otherwise, the hypotheses turn into a dimension that will change from beginning to end. This situation actually points to an incorrectly specified model study. Studies in the literature regarding an incorrectly specified model (Lai, 2018) should be examined.

The analysis findings have quite good results in all sample sizes. We see that the results such as incorrect specification of the model in 100 and 200 units are eliminated in the sample size of 300 units. In all programs, the first sample size in which all path coefficients are significant is determined as 300. Path coefficients are also significant in 400 and 500 units. In the evaluation of biases of the programs, RB values and MAB values were examined based on previous studies (Curran et al., 1996; Bandalos, 2002; Flora & Curran, 2004; Kaplan, 1989; Yang-Wallentin et al., 2010). When the analysis

findings are examined, the biases in each sample size in the Hayes plugin of Spss are at a level that can be considered excessive. Therefore, we can say that this program, which does not take into account the measurement model, is quite weak in determining relationships supported by hypotheses and theories such as the mediation model. If a ranking is to be made between the other three programs, the outputs of the Mplus program showed the least bias. Thus, we can say that the WLSMV technique performs better than DWLS and ML.

The second place was taken by the Jamovi program when we look at it again within the framework of biases. With this result, we can say that the DWLS technique performs better than ML in categorical data. According to the available findings, the comments are not very clear when we consider the limitations of the study. However, despite the shortcomings of the study, a clear statement would be not to perform the mediation analysis in the Spss program. A clearer statement among other estimation techniques can only be achieved if issues such as missing data, standard errors, failure to provide multivariate normality, fit values and factor loadings are addressed in the simulation setup.

References

- Arminger, G., & Muthén, B. O. (1998). A Bayesian approach to nonlinear latent variable models using the Gibbs sampler and the Metropolis-Hastings algorithm. *Psychometrika*, 63, 271-300.
- Bandalos, D.L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling*, 9(1), 78-102.
- Bollen, K. A. (1989). Structural equations with latent variables. John Wiley & Sons.
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1, 16-29.
- Çağlayan, A. A., & Özenç, E. (2024). Eğitim Bilimlerinde Aracı Değişken Konusunda Yapılmış Lisansüstü Tezlerin İçerik Analizi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 20(2), 199-222.
- Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), 466-491.
- Forero, C. G., & Maydeu-Olivares, A. (2009). Estimation of IRT graded response models: limited versus full information methods. *Psychological methods*, 14(3), 275.
- Ganzach, Y. (1997). Misleading interaction and curvilinear terms. *Psychological methods*, 2(3), 235.
- Hayes, A. F. (2018). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach (second edition). The Guilford Press.
- Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34(2), 183-202.
- Kaplan, D. (1989). A study of the sampling variability and z-values of parameter estimates from misspecified structural equation models. *Multivariate Behavioral Research*, 24(1), 41-57.
- Kelava, A., & Brandt, H. (2009). Estimation of nonlinear latent structural equation models using the extended unconstrained approach. *Review of psychology*, 16(2), 123-132.
- Kelava, A., & Nagengast, B. (2012). A Bayesian model for the estimation of latent interaction and quadratic effects when latent variables are non-normally distributed. *Multivariate Behavioral Research*, 47(5), 717-742.
- Lai, K. (2018). Estimating standardized SEM parameters given nonnormal data and incorrect model: Methods and comparison. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(4), 600-620.

- Lee, S. Y., & Song, X. Y. (2003). Model comparison of nonlinear structural equation models with fixed covariates. *Psychometrika*, 68(1), 27-47.
- Lee, S. Y., & Zhu, H. T. (2000). Statistical analysis of nonlinear structural equation models with continuous and polytomous data. *British Journal of Mathematical and Statistical Psychology*, 53(2), 209-232.
- Li, C. H. (2021). Statistical estimation of structural equation models with a mixture of continuous and categorical observed variables. *Behavior research methods*, 53(5), 2191-2213.
- Maydeu-Olivares, A. (2017). Maximum likelihood estimation of structural equation models for continuous data: Standard errors and goodness of fit. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(3), 383-394.
- Moosbrugger, H., Schermelleh-Engel, K., Kelava, A., & Klein, A. G. (2009). Testing multiple nonlinear effects in structural equation modeling: A comparison of alternative estimation approaches. *Structural equation modeling in educational research: Concepts and applications*, (103-136): *Structural Equation Modeling in Educational Research: Concepts and Applications*. Teo & M. S. Khine (Ed.) Rotterdam, NL: Sense Publishers.
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online*, 8(2), 23-74.
- Shi, D., & Maydeu-Olivares, A. (2020). The effect of estimation methods on SEM fit indices. *Educational and psychological measurement*, 80(3), 421-445.
- Shi, D., DiStefano, C., McDaniel, H. L., & Jiang, Z. (2018). Examining chi-square test statistics under conditions of large model size and ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(6), 924-945.
- Wall M. M., & Amemiya Y. (2007). Nonlinear Structural Equation Modeling as a Statistical Method (321- 343): *Handbook of Latent Variable and Related Models*. Lee S.Y. (Ed.) North-Holland. Elsevier.
- Wall, M. M. (2009). Maximum likelihood and bayesian estimation for nonlinear structural equation models (540- 567): *The SAGE Handbook of Quantitative Methods in Psychology*. Millsap, R. E., & Maydeu-Olivares, A. (Ed.) Sage Publications Ltd.
- Yang-Wallentin, F., Jöreskog, K. G., & Luo, H. (2010). Confirmatory factor analysis of ordinal variables with misspecified models. *Structural Equation Modeling*, 17(3), 392-423.

Acknowledgment

There is no institution or person supporting the work.

Conflict of Interest

The study is a single author and has declared that there is no conflict of interest.

Author Contributions

The study is a single author and all contributions belong to the author.

Investigation of the Structure and Function of Acid-Sensing Ion Channels

Ziya Çakır¹

Abstract

Acid-sensing ion channels (ASICs) are members of the epithelial sodium channel/degenerin (ENaC/DEG) superfamily and are encoded by five distinct genes, giving rise to seven different subunits. These subunits predominantly assemble into trimeric ion channels that, upon activation by extracellular protons, generate a transient inward current, thereby enhancing cellular excitability. These ion channels, which are activated particularly by extracellular acidification (pH decrease), regulate intracellular ion balance and electrical activity. ASICs exhibit a broad range of tissue distributions and display diverse biophysical characteristics. Moreover, their capacity to form both homomeric and heteromeric trimers adds further complexity to their functional and pharmacological properties. Certain modulators have been identified that lower the proton concentration required for ASIC activation, thereby sensitizing these channels. The roles of ASICs in neurological diseases, pain mechanisms and psychiatric disorders are attracting increasing interest. Substantial evidence from transgenic mouse models and pharmacological investigations indicates that ASICs represent a promising target for therapeutic intervention in various pathological conditions. Further investigation of the molecular mechanisms of ASICs may enable the development of new therapeutic strategies in disease models. This review aims to summarize the current understanding of ASIC function, explore their physiological and pathological roles, discuss mechanisms of modulation, and identify critical gaps in knowledge that warrant further investigation.

1 Tokat Gaziosmanpaşa University, Faculty of Health Services Vocational School, Department of Oral and Dental Health, 60000, Tokat, TÜRKİYE

INTRODUCTION

Acid-Sensitive Ion Channels

Acid-sensitive ion channels were first discovered in rat sensory neurons in 1980 by Krishtal and Pidoplichko using the voltage clamp technique (Krishtal & Pidoplichko, 1980). Krishtal and Pidoplichko argued that these acid-evoked currents were mediated by a previously unidentified ion channel and thought that they might be acid-evoked Ca^{2+} channels. Lazdunski and colleagues later showed that these channels were sensitive to amiloride and permeable to Na^{+} ions, and described acid-sensitive ion channels for the first time (Waldmann, Champigny, Bassilana, Heurteaux, & Lazdunski, 1997). In 2007, the crystal structure of ASIC1a in chicken was described, and they showed that ASICs consist of trimers, i.e., three different subunits (Jasti, Furukawa, Gonzales, & Gouaux, 2007). Acid-sensitive ion channels, a subgroup of the voltage-insensitive proton-gated degenerin/epithelial sodium channel (DEG/ENaC) superfamily, were shown to be widespread in both the central and peripheral nervous systems (Grunder & Chen, 2010; Waldmann & Lazdunski, 1998). In subsequent studies, the basic properties of these channels were determined. These channels generally open with decreasing pH (Krishtal & Pidoplichko, 1980), are permeable to Na^{+} allow (Krishtal & Pidoplichko, 1981a), and are blocked by amiloride (Krishtal & Pidoplichko, 1981b). In normal cells, extracellular pH is between 7.3-7.4 (Calorini, Peppicelli, & Bianchini, 2012). ASICs are rapidly activated and desensitized when extracellular pH falls below the normal physiological value (approximately $\text{pH}=7.4$) (Hesslager, Timmermann, & Ahring, 2004). ASIC subunits have different pH sensitivities, activation kinetics and desensitization rates (Korkushko & Kryshal, 1984). ASIC members are named as “X-NaC”. The abbreviation “X” means that it is related to a basic property of the proteins (DEG, degenerin protein), while the abbreviation “C” means ‘channel’. For example; epithelial Na^{+} channel (ENaC), FMRF amide (Phe-Met-Arg-Phe-NH₂) gated Na^{+} channel (FaNaC), *Drosophila* gonad Na^{+} channel (dGNaC), human intestinal Na^{+} channel (hINaC), and brain, liver, and intestine Na^{+} channel (BLINaC) (Grunder & Chen, 2010). At least 6 ASIC subunits encoded by 4 different genes have been identified in rodents. There are 6 known ASIC subgroups: ASIC1a (Waldmann, Champigny, et al., 1997), ASIC1b (Chen, England, Akopian, & Wood, 1998), ASIC2a (Price, Snyder, & Welsh, 1996), ASIC2b (Lingueglia et al., 1997), ASIC3 (Waldmann, Bassilana, et al., 1997) and ASIC4 (Akopian, Chen, Ding, Cesare, & Wood, 2000). ASIC subtypes can exist as homomers or heteromers. However, ASIC2b and ASIC4 subtypes are not functional as homomers; they function by forming heteromers with other ASIC subtypes.

All ASIC subunits consist of two hydrophobic transmembrane domains (TM1 and TM2), a large extracellular loop and short intracellular N- and C-terminal regions. The typical structure of ASICs and their localization in the cell membrane are shown in Figure 1 (Osmakov, Andreev, & Kozlov, 2014).

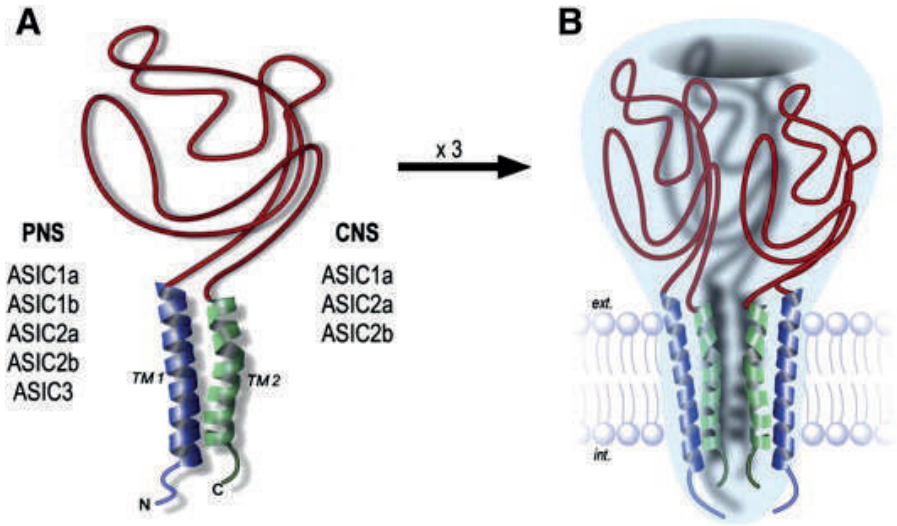


Figure 1. A. Schematic representation of ASIC subunits B. Trimeric structure of the ASIC channel (Osmakov, Andreev, & Kozlov, 2014).

ASICs have been shown to be present in most regions of the mammalian brain and in all sensory ganglia. However, it has been reported that ASIC channels are found in neurons but not in glial cells. Studies have shown that ASIC1a, ASIC2a, and ASIC2b genes are more expressed in the central nervous system, while ASIC1b and ASIC3 genes are more expressed in the peripheral nervous system (Chu et al., 2011). ASIC subtypes exhibit different electrophysiological and pharmacological properties (Hesselager, Timmermann, & Ahring, 2004). While all ASICs only conduct Na⁺ ions, ASIC1a conducts calcium ions along with sodium (Waldmann, Champigny, Bassilana, Heurteaux, & Lazdunski, 1997). Zinc (Zn²⁺) enhances homomeric and heteromeric ASIC2a currents, while attenuating other ASIC currents (110). Psalmotoxin1 is a specific ASIC1a inhibitor (Baron et al., 2001). Lead (Pb²⁺) inhibits ASIC1a currents but not other ASICs (Wang et al., 2006). Salicylic acid blocks only the ASIC3 subunit (Voileyet et al., 2001). The kinetic properties of the ion currents generated by ASIC subunits are different from each other (Figure 2) (Osmakov, Andreev, & Kozlov, 2014).

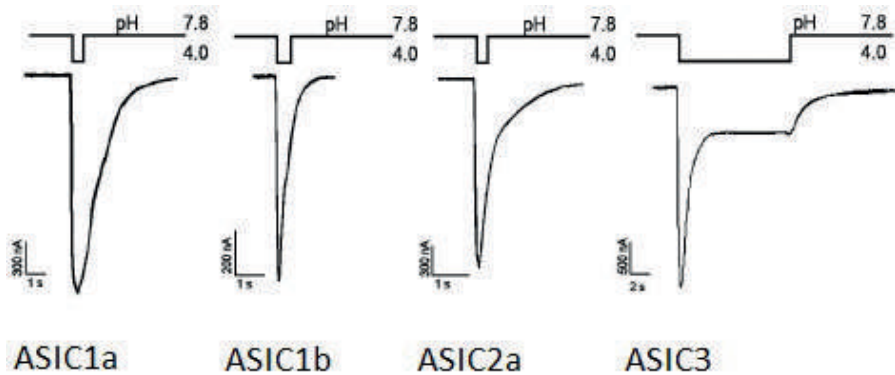


Figure 2. Currents generated by ASIC subunits expressed in frog (*Xenopus laevis*) oocyte cells (Osmakov, Andreev, & Kozlov, 2014).

Studies have shown that acid-sensitive ion channels play a role in many physiological events such as nociception (Chen et al., 2002), touch (Price et al., 2001), taste and smell transmission (Lin et al., 2002), long-term potentiation (LTP), synaptic transmission, memory/learning (Wemmie et al., 2002), sensory transmission and retinal integrity (Ettaiche et al., 2004) and in many pathological events such as pain (Duan et al., 2007), ischemia-related brain damage (Xiong et al., 2004), stroke and epilepsy (Biagini et al., 2001). In the auditory system, the presence of ASICs has been shown in hair cells in the cochlea (Ugawa et al. 2006), spiral ganglion neurons (Peng et al., 2004), vestibular organs (Mercado et al., 2006) and inferior colliculus neurons (Zhang et al., 2008). It has been determined that the ASIC2a subunit is related to noise sensitivity in mice (Peng et al., 2004), and that the absence of ASIC2 does not impair hearing (Roza et al., 2004). In a study conducted on mice that did not express the ASIC3 gene, hearing loss was found in mice lacking the ASIC3 subunit (h, ldebrand et al., 2004).

Acid-sensitive ion channels help in the transduction of stimuli in various physiological and pathophysiological conditions (Askwith et al., 2001). ASIC inhibitors are isolated from animal venoms, which are natural peptide toxins, as non-specific molecules (Diochot et al., 2007). Pepsin toxins, neuropeptides, organic compounds and some di/trivalent cations obtained from some animal venoms are involved in the modulation of ASICs.

MATERIAL AND METHODS

Material

Albino BALB/c mice aged 14-17 days were used in patch clamping studies. Cochlear nucleus brain sections were prepared to record patch clamping. Patch clamp recordings were taken using patch clamp micropipettes. These procedures were performed in constantly oxygenated normal CSF fluid.

Methods

Patch clamp technique

The patch clamp technique is a widely used electrophysiological technique to study ion currents of channels in the cell membrane. The patch clamp technique is based on measuring voltage changes in cells via the electrode in the pipette. This technique is important for determining the biophysical, physiological and pharmacological properties of ion channels. When the patch clamp technique was first discovered, it was used to control the voltage of a small piece of cell membrane. Now, it is used for both voltage clamping and current clamping on the membrane with the help of a micropipette.

Statistical Analysis

Statistical evaluation in this study; It was done using the Statistical Package for Social Science (SPSS) Version 23.0 (SPSS inc, Chicago, USA) package program. In patch clamping studies, it was examined whether there was a difference between the recordings taken before the application of the acidic solution and the recordings taken during the application of the acidic solution. Student's t test was used to determine the difference during acidic solution. Descriptive statistics for numerical variables were expressed as group mean \pm standard error (S.H.). In statistical evaluation, $p < 0.05$ value was accepted as significant.

RESULTS

To investigate acid-induced currents in neurons of the cochlear nucleus, the responses of these neurons to a decrease in pH in the extracellular solution were recorded using the whole-cell patch-clamp technique. In the present experiments, the average resting membrane potential of stellate cells was found to be -63.7 ± 0.71 mV ($n = 42$). Therefore, recordings were made by keeping the cells at -62 mV holding potential. The acidic solution stimulated most of the neurons by generating inward currents. The amplitudes of acid-induced currents showed high sensitivity to pH (Figure 3). Acidic solutions with pH ranging from 7.4 to 4 were applied. The

relationship between the peak values of the currents occurring in different acidic solutions and the pH values obtained was shown graphically.

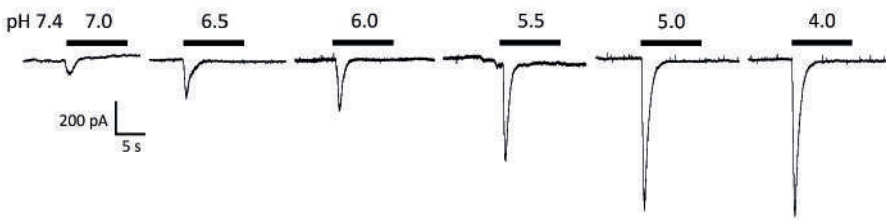


Figure 3. Typical traces showing the inward currents that are activated by extracellular solutions with different pH.

DISCUSSION AND CONCLUSION

Acid-sensitive ion channels (ASICs) are voltage-insensitive sodium channels that are activated by acidification of the extracellular environment. ASICs are widely expressed in the central and peripheral nervous systems. There are reports that ASICs activated by acid exposure play a role in various physiological mechanisms and pathophysiological events. Although local pH decreases in various cellular structures are important enough to activate ASICs, it is thought that pH in tissues is tightly regulated by homeostatic mechanisms (Chesler & Kaila, 1992). For example, synaptic vesicles are generally acidic and have a pH of 5.7 (Yuste et al., 2000). Studies in hippocampal neurons have shown that the extracellular pH in the synaptic cleft temporarily drops below 6 after vesicle release in synapses showing intense synaptic activity (Miesenbock et al., 1998). Again, studies in retinal cone receptors have shown that ASICs can affect synaptic transmission (DeVries SH., 2001). It is thought that many cellular stresses in the cochlea (e.g. ischemia and inflammation) can induce local acidosis. Although the cause is not well understood, it is known that inflammation or ischemia results in sudden hearing loss (Roza et al., 2004). In this context, it is possible that ASICs may underlie some hearing loss caused by non-mechanical causes. It is also thought that ASICs can act as a sensor against harmful stimuli and may have an important role in some pathological cases. It has been shown that intercellular acidosis affects and activates ASICs, and ASICs trigger excessive excitatory activities associated with epileptic seizures and ischemia. Therefore, it is thought that ASICs play a role in the pathogenesis of these diseases (Varming T. 1999).

References

- Akopian, A. N., Chen, C. C., Ding, Y., Cesare, P., & Wood, J. N. (2000). A new member of the acid-sensing ion channel family. *Neuroreport*, 11(10), 2217-2222. doi: 10.1097/00001756-200007140-00031
- Askwith CC, Benson CJ, Welsh MJ, Snyder PM. DEG/ENaC ion channels involved in sensory transduction are modulated by cold temperature. *Proceedings of the National Academy of Sciences of the United States of America*. 2001;98(11):6459-6463
- Baron A, Schaefer L, Lingueglia E, Champigny G, Lazdunski M. Zn²⁺ and H⁺ are coactivators of acid-sensing ion channels. *The Journal of biological chemistry*. 2001;276(38):35361-35367
- Biagini G, Babinski K, Avoli M, Marcinkiewicz M, Seguela P. Regional and subunit-specific downregulation of acid-sensing ion channels in the pilocarpine model of epilepsy. *Neurobiology of disease*. 2001;8(1):45-58
- Calorini, L., Peppicelli, S., & Bianchini, F. (2012). Extracellular acidity as favouring factor of tumor progression and metastatic dissemination. *Exp Oncol*, 34(2), 79-84
- Chen, C. C., England, S., Akopian, A. N., & Wood, J. N. (1998). A sensory neuron-specific, proton-gated ion channel. *Proc Natl Acad Sci U S A*, 95(17), 10240-10245. doi: 10.1073/pnas.95.17.10240
- Chen CC, Zimmer A, Sun WH, Hall J, Brownstein MJ, Zimmer A. A role for ASIC3 in the modulation of high-intensity pain stimuli. *Proceedings of the National Academy of Sciences of the United States of America*. 2002;99(13):8992-8997
- Chesler M, Kaila K. Modulation of pH by neuronal activity. *Trends in neurosciences*. 1992;15(10):396-402
- Chu XP, Papasian CJ, Wang JQ, Xiong ZG. Modulation of acid-sensing ion channels: molecular mechanisms and therapeutic potential. *International journal of physiology, pathophysiology and pharmacology*. 2011;3(4):288-309
- DeVries SH. Exocytosed protons feedback to suppress the Ca²⁺ current in mammalian cone photoreceptors. *Neuron*. 2001;32(6):1107-1117
- Diochot S, Salinas M, Baron A, Escoubas P, Lazdunski M. Peptides inhibitors of acid-sensing ion channels. *Toxicon : official journal of the International Society on Toxinology*. 2007;49(2):271-284
- Duan B, Wu LJ, Yu YQ, Ding Y, Jing L, Xu L, et al. Upregulation of acid-sensing ion channel ASIC1a in spinal dorsal horn neurons contributes to inflammatory pain hypersensitivity. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2007;27(41):11139-11148

- Ettaiche M, Guy N, Hofman P, Lazdunski M, Waldmann R. Acid-sensing ion channel 2 is important for retinal function and protects against light-induced retinal degeneration. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2004;24(5):1005-1012
- Grunder, S., & Chen, X. (2010). Structure, function, and pharmacology of acid-sensing ion channels (ASICs): focus on ASIC1a. *Int J Physiol Pathophysiol Pharmacol*, 2(2), 73-94
- Hesslager, M., Timmermann, D. B., & Ahring, P. K. (2004). pH Dependency and desensitization kinetics of heterologously expressed combinations of acid-sensing ion channel subunits. *J Biol Chem*, 279(12), 11006-11015. doi: 10.1074/jbc.M313507200
- Hildebrand MS, de Silva MG, Klockars T, Rose E, Price M, Smith RJ, et al. Characterisation of DRASIC in the mouse inner ear. *Hearing research*. 2004;190(1-2):149-160
- Jasti, J., Furukawa, H., Gonzales, E. B., & Gouaux, E. (2007). Structure of acid-sensing ion channel 1 at 1.9 Å resolution and low pH. *Nature*, 449(7160), 316-323. doi: 10.1038/nature06163
- Korkushko, A. O., & Kryshchal, O. A. (1984). [Blocking of proton-activated sodium permeability of the membranes of trigeminal ganglion neurons in the rat by organic cations]. *Neirofiziologija*, 16(4), 557-561
- Krishtal, O. A., & Pidoplichko, V. I. (1980). A receptor for protons in the nerve cell membrane. *Neuroscience*, 5(12), 2325-2327. doi: 10.1016/0306-4522(80)90149-9
- Krishtal, O. A., & Pidoplichko, V. I. (1981a). A receptor for protons in the membrane of sensory neurons may participate in nociception. *Neuroscience*, 6(12), 2599-2601. doi: 10.1016/0306-4522(81)90105-6
- Krishtal, O. A., & Pidoplichko, V. I. (1981b). A "receptor" for protons in small neurons of trigeminal ganglia: possible role in nociception. *Neurosci Lett*, 24(3), 243-246. doi: 10.1016/0304-3940(81)90164-6
- Lin W, Ogura T, Kinnamon SC. Acid-activated cation currents in rat vallate taste receptor cells. *Journal of neurophysiology*. 2002;88(1):133-141
- Lingueglia, E., de Weille, J. R., Bassilana, F., Heurteaux, C., Sakai, H., Waldmann, R., & Lazdunski, M. (1997). A modulatory subunit of acid sensing ion channels in brain and dorsal root ganglion cells. *J Biol Chem*, 272(47), 29778-29783. doi: 10.1074/jbc.272.47.29778
- Mercado F, Lopez IA, Acuna D, Vega R, Soto E. Acid-sensing ionic channels in the rat vestibular endorgans and ganglia. *Journal of neurophysiology*. 2006;96(3):1615-1624
- Miesenbock G, De Angelis DA, Rothman JE. Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature*. 1998;394(6689):192-5

- Osmakov, D. I., Andreev, Y. A., & Kozlov, S. A. (2014). Acid-sensing ion channels and their modulators. *Biochemistry (Mosc)*, 79(13), 1528-1545. doi: 10.1134/S0006297914130069
- Peng BG, Ahmad S, Chen S, Chen P, Price MP, Lin X. Acid-sensing ion channel 2 contributes a major component to acid-evoked excitatory responses in spiral ganglion neurons and plays a role in noise susceptibility of mice. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2004;24(45):10167-10175
- Price, M. P., Snyder, P. M., & Welsh, M. J. (1996). Cloning and expression of a novel human brain Na⁺ channel. *J Biol Chem*, 271(14), 7879-7882. doi: 10.1074/jbc.271.14.7879
- Price MP, McIlwrath SL, Xie J, Cheng C, Qiao J, Tarr DE, et al. The DRASIC cation channel contributes to the detection of cutaneous touch and acid stimuli in mice. *Neuron*. 2001;32(6):1071-1083
- Roza C, Puel JL, Kress M, Baron A, Diochot S, Lazdunski M, et al. Knockout of the ASIC2 channel in mice does not impair cutaneous mechanosensation, visceral mechanonociception and hearing. *The Journal of physiology*. 2004;558(Pt 2):659-669
- Ugawa S, Inagaki A, Yamamura H, Ueda T, Ishida Y, Kajita K, et al. Acid-sensing ion channel-1b in the stereocilia of mammalian cochlear hair cells. *Neuroreport*. 2006;17(12):1235-1239
- Varming T. Proton-gated ion channels in cultured mouse cortical neurons. *Neuropharmacology*. 1999;38(12):1875-1881
- Voilley N, de Weille J, Mamet J, Lazdunski M. Nonsteroid anti-inflammatory drugs inhibit both the activity and the inflammation-induced expression of acid-sensing ion channels in nociceptors. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2001;21(20):8026-8033
- Waldmann, R., Champigny, G., Bassilana, F., Heurteaux, C., & Lazdunski, M. (1997). A proton-gated cation channel involved in acid-sensing. *Nature*, 386(6621), 173-177. doi: 10.1038/386173a0
- Waldmann, R., & Lazdunski, M. (1998). H(+)-gated cation channels: neuronal acid sensors in the NaC/DEG family of ion channels. *Curr Opin Neurobiol*, 8(3), 418-424. doi: 10.1016/s0959-4388(98)80070-6
- Wang W, Duan B, Xu H, Xu L, Xu TL. Calcium-permeable acid-sensing ion channel is a molecular target of the neurotoxic metal ion lead. *The Journal of biological chemistry*. 2006;281(5):2497-2505
- Wemmie JA, Chen J, Askwith CC, Hruska-Hageman AM, Price MP, Nolan BC, et al. The acid-activated ion channel ASIC contributes to synaptic plasticity, learning, and memory. *Neuron*. 2002;34(3):463-477

- Xiong ZG, Zhu XM, Chu XP, Minami M, Hey J, Wei WL, et al. Neuroprotection in ischemia: blocking calcium-permeable acid-sensing ion channels. *Cell*. 2004;118(6):687-698
- Yuste R, Miller RB, Holthoff K, Zhang S, Miesenbock G. Synapto-pHluorins: chimeras between pH-sensitive mutants of green fluorescent protein and synaptic vesicle membrane proteins as reporters of neurotransmitter release. *Methods in enzymology*. 2000;327:522-546
- Zhang, M., Gong, N., Lu, Y. G., Jia, N. L., Xu, T. L., & Chen, L. (2008). Functional characterization of acid-sensing ion channels in cultured neurons of rat inferior colliculus. *Neuroscience*, 154(2), 461-472. doi: 10.1016/j.neuroscience.2008.03.040

Acknowledgment

We extend our gratitude to all the experimenters who contributed to this research. We also acknowledge the Gaziantep University, School of Medicine, Research Center, Gaziantep, Turkey, for providing the necessary facilities to carry out this study.

Conflict of Interest

The authors have disclosed that they have no competing interests.

Author Contributions

ZC: Conceptualization, Project administration, Resources, Visualization, Data curation, Formal Analysis, Software, Resources, Writing – original draft, Writing.

Assessment of Ankara University ERASMUS+ Programme Outgoing Students by Using Statistical Methodologies for the 2023-2024 Period

Cafer Yıldırım¹

Özlem Türkşen²

Necdet Ünüvar³

İlker Astarıcı⁴

Abstract

Data analysis has a crucial importance to realize student profiles within universities. The data should be analyzed with proper statistical methods and obtained results could be taken into consideration to make decision for supporting student activities and success. This study aims to realize the profile of Ankara University ERASMUS+ Programme Outgoing Students for the 2023-2024 period. Data set is obtained from the ERASMUS Office database and data preprocessing is achieved to organize the data. The well-organized data is analysed according to the categorical and numerical data type by using statistical explanatory data analysis, e.g. descriptive statistics, data visualization. Several parametric and non-parametric statistical tests are applied to the data set. In this context, chi square analysis and correspondence analysis are performed to understand whether there is a relationship between categorical data. One-way analysis of variance (ANOVA) is performed to understand whether there is a difference between groups containing numerical data. Additionally, classification models were created to predict students'

1 Ankara University, European Union Education Programs Coordinatorship and Project Office, Türkiye, Ankara University, Faculty of Dentistry, Department of Basic Medical Sciences, Türkiye

2 Ankara University, Faculty of Science, Department of Statistics, Türkiye

3 Rectorate of Ankara University, Türkiye

4 Turkish National Agency, Türkiye

success and failure through machine learning classification algorithms, called Logistic Regression, k-Nearest Neighbors, Support Vector Machine and Random Forest. It can be said from the results that statistical methodologies help to identify patterns of the data to get knowledge for administrative processes at higher education in European level.

INTRODUCTION

In today's data-driven world, effective data analysis is crucial for understanding trends, optimizing processes, and predicting future outcomes in the field of education as in many areas. Just as universities are important in higher education, ERASMUS offices that provide international connections at universities are also important. The ERASMUS Office holds a central position within the university structure as a key unit responsible for coordinating international academic mobility programs, particularly under the ERASMUS+ framework. Data-driven decision making allows to optimize operations, manage budgets and plan for future growth for administrative duties. Moreover, the statistical data analysis plays a critical role by enabling evidence-based decision making across academic, administrative and research activities. It also helps to assess student performance, improve teaching methods, allocate resources efficiently and enhance effectiveness (Ünivar et al., 2023).

The statistical data analysis process typically includes several stages: data collection, data cleaning, data exploration, data modeling or transformation, and interpretation of results to make clear decision. Each step plays a vital role in ensuring the accuracy and reliability of the insights generated. Statistical methodologies such as descriptive statistics, data visualization, classification analysis, regression analysis and clustering are frequently applied to explore and present data effectively (Türkşen, 2024). Researchers rely on analytical tools to test hypotheses, validate findings and draw meaningful conclusions from large and complex datasets. While data analysis focuses on interpreting existing data to generate insights, data science involves building predictive models and algorithms using advanced mathematics and programming (Geron, 2017). Data scientists often use machine learning methods and data analysts mainly work with descriptive analytics whereas statisticians use both machine learning methods and descriptive analytics. Common tools used in data analysis include Excel, SQL, Python, R, Tableau, and Power BI.

In this study, it is aimed to assess the profile of Ankara University ERASMUS+ Programme Outgoing Students by using statistical methodologies for the 2023-2024 period since the ERASMUS Office plays an important role in supporting international academic mobility

and fostering global connections within universities. The main aim of this study is to reveal the current profile of the Ankara University ERASMUS Office and to strength administrative performance with data-based decision making in several parts e.g. collaboration with partner universities across Europe, promoting academic cooperation, joint research projects, and best practices in education. The rest of the paper is organized as follows. Applied statistical methodologies are given in Materials and Methods section. The analysis results are presented in Results section. Finally, conclusion is given in the last section.

MATERIAL AND METHODS

Material

Data set, in which undergraduate students are taken into consideration, is obtained from the ERASMUS Office database. Data preprocessing stage is achieved with checking missing data, data cleaning and feature selection. Features, namely variables, are defined as Faculty, Research Field (Science, Health and Social), Department, Visited University, Country, Zone (West Europe, East Europe and South Europe), Accomation Time, Visiting Time Period and Success Status. The Success Status, which has categorical data, considered as dependent variable while the others are considered as independent variables which are composed with categorical and continuous data. In addition, a new continuous variable, called Success (%), is calculated according to the achieved ECTS score from the raw data. Data exploration is done by calculating summary statistics (descriptive statistics - central tendency and spread) and data visualization (histogram, barplot, pieplot, boxplot etc.). Some non-parametric statistical tests, (i) Chi-Square Analysis, (ii) Correspondence Analysis, and (iii) Kruskal-Wallis test, are applied to the data set. In this study, the problem is considered as classification problem. Then, Machine Learning Classification Algorithms, called Logistic Regression (LR), k-Nearest Neighbors (KNN), Support Vector Machine (SVM) and Random Forest (RF), are applied to predict students' success status. All the calculations are done by using Python libraries, called numpy, pandas, matplotlib, seaborn, scikit-learn.

Methods

Exploratory Data Analysis

Exploratory Data Analysis (EDA) is an approach to analyze datasets for summarizing their main characteristics, often with visual methods. The main goal of the EDA is to understand the data better before applying any formal statistical models or machine learning algorithms. By using plots like

histograms, scatter plots, and boxplots, as well as summary statistics such as mean, median, and standard deviation, analysts can gain insights into the underlying structure of the data, identify important variables, and detect outliers or missing values that may affect further analysis.

Non-parametric Statistical Analysis

i. Chi-Square Analysis

Chi-Square Analysis, also known as Chi-Squared Test, is a statistical method used to determine if there is a significant association between two categorical variables. It compares the observed frequencies in each category of a contingency table with the expected frequencies if the variables were independent. The main purpose of the Chi-Square Test is to assess whether there is a statistically significant difference between the observed and expected frequencies in one or more categories. The Chi-Square Statistic is calculated using the formula

$$\chi^2 = \sum_{i=1}^r \frac{(O_i - E_i)^2}{E_i}$$

where r is the number of category, O_i is observed frequency and E_i is expected frequency. A higher Chi-Square value indicates a greater difference between observed and expected values. The result is compared to a critical value from the chi-square distribution table or evaluated using a p-value to determine significance (Karagöz, 2019). In this study, Chi-Square Test is used to determine if there is a significant association between two categorical variables, called Chi-Square Test of Independence.

ii. Correspondence Analysis

Correspondence Analysis is a multivariate statistical method used to visualize and interpret relationships between categorical variables in a contingency table. It reduces the dimensions of the data, allowing researchers to explore associations between rows and columns by projecting them into a lower-dimensional space, typically a 2D plot. The main goal of the Correspondence Analysis is to summarize and visualize patterns in large categorical datasets (Alpar, 2025).

iii. Kruskal-Wallis Test

The Kruskal-Wallis test is a non-parametric statistical method used to determine if there are statistically significant differences between the medians of three or more independent groups. It is considered the non-parametric

alternative to the one-way ANOVA, and it does not assume that the data follows a normal distribution (Karagöz, 2019).

Machine Learning Classification Algorithms

Classification algorithms in machine learning are supervised learning methods used to predict categorical dependent variables based on independent variables. These algorithms learn patterns from labeled training data and apply that knowledge to classify new, unseen instances into one of the predefined classes (Müller and Guido, 2017).

LR is a classification method. It models the probability of a binary outcome using a logistic function. It works well when the relationship between features and the target is approximately linear.

KNN is a simple instance-based classification algorithm that classifies a new data point based on the majority class among its k nearest neighbors in the feature space.

SVM finds the optimal boundary (hyperplane) that best separates different classes. It works well in high-dimensional spaces and with small datasets.

RF is an ensemble method that builds multiple decision trees and combines their outputs to improve accuracy and reduce overfitting. It's robust and effective for many real-world problems.

The tuning parameters of these algorithms are defined by using expert knowledge. The performance comparison of the classification algorithms is achieved by using several metrics, e.g. Accuracy, Precision, Recall, F1-Score and AUC (Ulu Metin, 2024). If the data set is imbalanced the F1-Score metric can be considered as the most preferable one.

RESULTS

The raw data of the Ankara University ERASMUS Office is well-organized according to the defined data preprocessing process. Exploratory data analysis is applied and obtained results are presented in Figure 1.(a)-(c) for the ERASMUS+ Programme outgoing students.

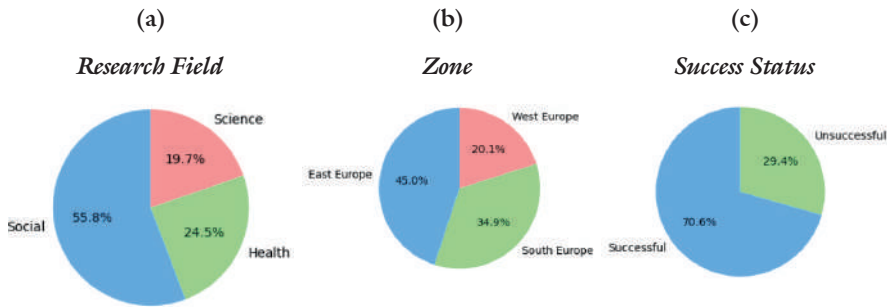


Figure 1. Pie plots of Research Field, Zone and Success Status variables

The distribution of Research Fields can be seen in Figure 1.(a). It can be said from Figure 1.a that the majority of the students are from the social research field. From Figure 1.(b), it is possible to say that the students prefer East Europe. And also, the majority of the students are successful, as presented in Figure 1.(c).

The Success (%) and summary statistics of success quantity can be seen in Figure 2. It is seen from the Figure 2 that the distribution of the Success (%) is left-skewed which means that high-achieving students are the majority. It can be easily said from the summary statistics that number of students, minimum, maximum, mean, median, standart deviation, range and inter quartile range values of the Success (%) are 269, 0, 100, 69.86, 75.76, 27.13, 100 and 36.36, respectively.

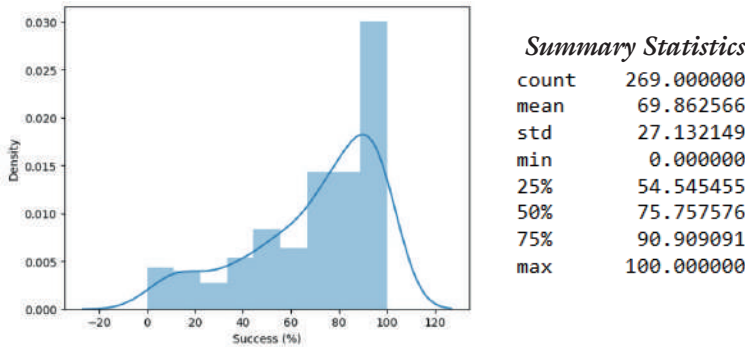


Figure 2. The distribution of the Success (%) and the summary statistics

The distribution of data about students Visiting Time Period has two-peaked distribution, presented in Figure 3. It is seen from Figure 3 that the most preferred visiting time periods are 5 and 10 months for the students.

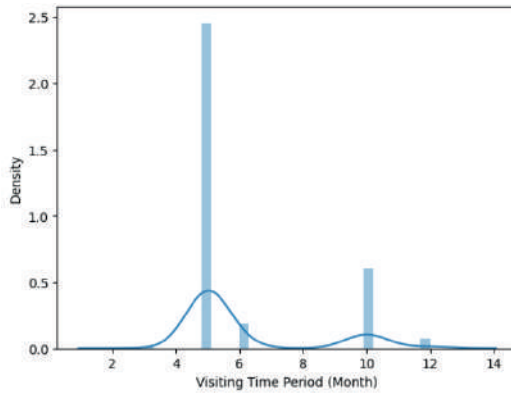


Figure 3. The distribution of the students visiting time period

Chi-Square Analysis is applied to understand whether there is a dependency between success status and Research Field, Success Status and Zone. Contingency tables of Success Status-Research Field and Success Status-Zone can be seen in Table 1 and Table 2, respectively. In Tables 1-2, the frequencies of the categories and expected values, given in parentheses, are presented. Chi-Square statistics is obtained as 6.04 (p -value=0.048) for Success Status and Research Field and the Chi-Square statistics value is equal to 22.98 (p -value=0.0000102) for Success Status and Zone. According to this statistics, Success Status has dependency with Research Field and Zone with %95 confidence.

Table 1. Contingency table of Success Status and Research Field

	Social	Science	Health	Total
Successful	98 (105.95)	38 (37.43)	54 (46.62)	190
Unsuccessful	52 (44.05)	15 (15.57)	12 (19.38)	79
Total	150	53	66	269

Table 2. Contingency table of Success Status and Zone

	East Europe	West Europe	South Europe	Total
Successful	103 (85.46)	34 (38.14)	53 (66.39)	190
Unsuccessful	18 (35.54)	20 (15.86)	41 (27.61)	79
Total	121	54	94	269

The results of Correspondence Analysis are presented in Figure 4. It can be said from Figure 4 that the ERASMUS+ Programme outgoing students from Science, Health and Social research fields are preferred to go to the West Europe, the East Europe and the South Europe, respectively.

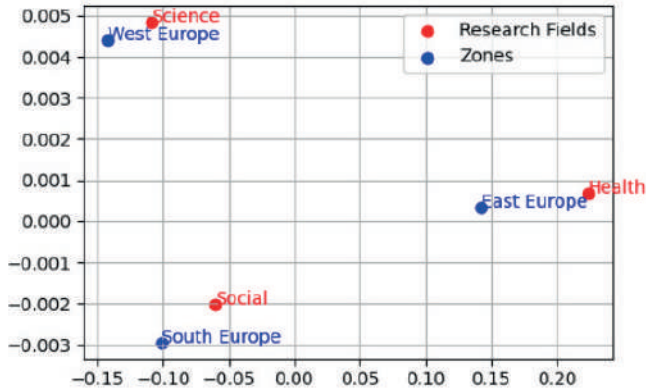


Figure 4. The Correspondence Analysis plot for Research Fields and Zones

It is seen from Figure 5 that the normality assumption is not provided according to the Success (%) in the research fields. Therefore, Kruskal-Wallis test is applied to determine whether there is a difference in Success (%) between the the Research Fields. The Kruskal-Wallis statistics is calculated as 1.6034 and p -value is equal to 0.4486. Thus, it is possible to say that there is no difference between the research fields according to the Success (%) with %95 confidence. In addition, this result can be seen from the boxplots clearly as given in Figure 6.

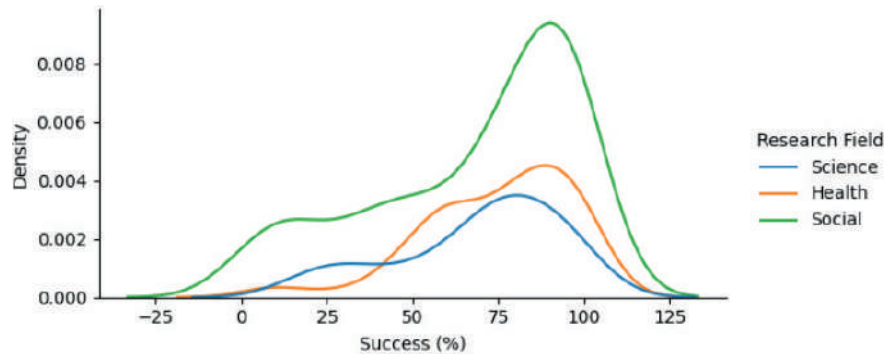


Figure 5. Distribution of the Success (%) for the Research Fields

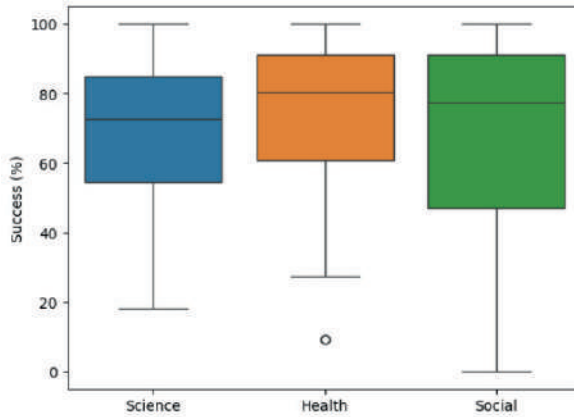


Figure 6. Box-plots of the Success (%) for the research fields

In this study, Success Status is considered as dependent variable with two categories, Successful and Unsuccessful. The values of the dependent variable has imbalanced distribution as can be seen from the pie plot in Figure1.(c). Machine Learning Classification Algorithms are applied to the data set. The obtained performance metrics of the algorithms are given in Table 3. It can be said from the Table 3 that the RF and the LR algorithms have better performance than the KNN and the SVM algorithms for the classification of ERASMUS+ Programme outgoing students according to the their Success Status.

Table 3. Performance metric values of Machine Learning Classification Algorithms

	Accuracy	Precision	Recall	F1-Score	AUC
LR	0.703704	0.720000	0.947368	0.818182	0.657895
KNN	0.666667	0.717391	0.868421	0.785714	0.648849
SVM	0.648148	0.720930	0.815789	0.765432	0.532895
RF	0.703704	0.720000	0.947368	0.818182	0.536184

DISCUSSION AND CONCLUSION

In higher education, the ERASMUS Office, as a key administrative unit within universities, plays a crucial role in facilitating international mobility and cooperation under the ERASMUS+ program, which is one of the most significant educational initiatives of the European Union. Accordingly, it would be appropriate to say that the statistical analysis of the data belonging

to the ERASMUS Office is important to track student progress, identify at-risk learners, personalize educational experiences, and data-driven decision making from an administrative perspective.

This study presents the statistical data analysis results of Ankara University ERASMUS Office for Outgoing Students during the 2023-2024 period for the first time. It is seen from the results that majority of the ERASMUS+ Programme outgoing students prefer to go East Europe, majority of them are from social research field and also successful. It is seen from the results that the most preferred visiting time periods are 5 and 10 months for the students. Chi-Squared statistics showed that the success status has dependency with Research Fields and Zones with %95 confidence. The Correspondence Analysis plot helps to realize closeness relationship of the Research Fields and Zones. According to the performance metrics of Machine Learning Classification Algorithms, the LR and the RF can be used for forecasting of Success Status of ERASMUS+ Programme Outgoing Students. It can be also said from the results that the LR is slightly better than the RF for all performance metrics considering the imbalanced data.

References

- Alpar R, 2025. Uygulamalı İstatistik ve Geçerlik-Güvenirlik. Detay Yayıncılık, Ankara.
- Geron A, 2017. Hands-On Machine Learning with Scikit-Learn and TensorFlow. O'Reilly, USA.
- Karagöz Y, 2019. SPSS, AMOSA, META Uygulamalı İstatistiksel Analizler. Nobel Yayınevi, Ankara.
- Müller AC and Guido S, 2017. Introduction to Machine Learning with Python A Guide for Data Scientists. O'Reilly, USA.
- Türkşen Ö, 2024. Veri Analizi ve Madenciliği. Ankara Üniversitesi Yayınevi, Ankara.
- Ulu Metin G, 2024. Ar-Ge ve Tasarım Merkezlerinin İstatistiksel Olarak Değerlendirilmesi: Veri Madenciliği Yöntemleri ile Hibrit Karar Verme. Doktora Tezi, Ankara Üniversitesi, Ankara.
- Ünüvar N, Apaydın A, Kutlu Ö, Vural MR, Türkşen Ö, Arıca Akkök E, 2023. Ankara Üniversitesi'nde Eğitim: Düşünce Atölyesi. Ankara Üniversitesi Yayınevi, Ankara.

Acknowledgment

This study is the result of research conducted in Ankara University, European Union Education Programs Coordinatorship and Project Office Data founded by Turkish National Agency.

Conflict of Interest

The authors have declared that there is no conflict of interest.

Author Contributions

Cafer Yıldırım: Coordinator, Obtaining data from database, Interpretation, Investigation, Writing – review & editing

Özlem Türkşen: Writing – Original draft, review & editing, Methodology, Statistical Analysis, Data visualization, Interpretation, Investigation

Necdet Ünüvar: Resources, Legal Representative of European Union Projects, Policy maker

İlker Astarıcı: Resources, Funding, Conceptualization

Parameter Estimation in Organic-Based Schottky Diodes Used in Solar Cell Applications with Artificial Intelligence Optimization Algorithms

Murat Açıkgöz¹

Defne Akay²

Özlem Türkşen³

Abstract

The electrical characterization of organic semiconductor materials plays a critical role in the design of advanced electronic devices and the understanding of their performance. Particularly, electrical parameters such as the ideality factor, barrier height, and series resistance, derived from the current-voltage (I–V) characteristics of rubrene-based metal-polymer semiconductor (MPS) structures, provide fundamental insights into the material's intrinsic and interfacial properties. The nonlinear nature of these parameters and the uncertainties associated with experimental data render traditional estimation methods inadequate. Additionally, investigating the effects of external factors, especially Cobalt-60 gamma irradiation, on these electrical parameters is of great importance for device reliability. The aim of this study is to accurately and reliably estimate the electrical parameters of Al/rubrene/p-Si Schottky-Junction solar cell using Artificial Intelligence (AI) optimization algorithms, based on experimental I–V data obtained after Cobalt-60 gamma irradiation. In this context, a total of twelve different AI optimization algorithms were employed, including Genetic Algorithm, Differential Evolution, Flower Pollination Algorithm, Artificial Bee Colony, Ant Colony Optimization, Bat Algorithm, Cuckoo Search Algorithm, Grey Wolf Optimization, Jaya Algorithm, Particle Swarm Optimization, Harmony Search, and Teaching-Learning-Based Optimization. The estimation of electrical parameters was performed using I–V measurements conducted across five different

1 Ankara University, Graduate School of Natural and Applied Sciences, Ankara, Türkiye

2 Ankara University, Faculty of Science, Department of Physics, Ankara, Türkiye

3 Ankara University, Faculty of Science, Department of Statistics, Ankara, Türkiye

forward voltage ranges, with the performance of the algorithms evaluated using *MAPE* and R^2 metrics. This study presents that the AI optimization algorithms can be used as optimization tool for parameter estimation of Schottky diode model and the algorithms may offer more effective results compared to traditional optimization methods in the electrical analysis of organic semiconductor models under Cobalt-60 irradiation.

INTRODUCTION

In the contemporary scientific landscape, the accurate modelling of complex systems has become a prevalent necessity. This has prompted researchers to adopt more flexible modelling approaches that transcend linear assumptions. In this context, the employment of nonlinear models becomes important, particularly in scenarios where the relationships between variables cannot be satisfactorily explained linearly. The nonlinear regression models are powerful statistical tools used when the relationships between dependent and independent variables are parametrically nonlinear. These models enable a more realistic and reliable representation of real-world phenomena across a wide range of disciplines, including engineering, economics, biomedical sciences, environmental sciences, and social sciences. However, the mathematical complexity introduced by the nonlinear structure presents significant challenges, particularly in the parameter estimation process, and limits the effectiveness of classical methods (Gallant, 1975).

Parameter estimation of nonlinear regression models is generally formulated as an optimization problem based on the minimization of the sum of squared errors. Traditional optimization methods used in this process, such as Gauss-Newton, Levenberg-Marquardt, and other derivative-based methods, may yield successful results under certain conditions (Türkşen, 2023). Nevertheless, they are limited by several issues, including the non-differentiability of the objective function, the presence of multiple local minima, high sensitivity to initial parameter values, and inefficiency in high-dimensional search spaces. In this regard, the challenges encountered in parameter estimation are not only technical but also significant in terms of computational efficiency and model reliability.

These limitations are especially important in solar cell modelling, which is becoming more and more important because of its role in condensed matter physics and its importance in the development of efficient photovoltaic structures. Addressing these computational challenges is essential for accurately modeling solar cells, which are key to meeting the global demand for clean and sustainable energy. By converting sunlight directly into electricity without emitting greenhouse gases, these systems represent a

promising and environmentally friendly alternative to conventional fossil fuel-based energy sources. To investigate these systems in greater detail, Schottky diodes with an Al/p-type Si structure were utilized. These diodes were fabricated using boron-doped single-crystal silicon wafers (p-type) as the semiconductor substrate, onto which aluminum was thermally evaporated to form the Schottky contact. The resulting devices were employed to examine the electrical characteristics and interfacial properties of the metal-semiconductor junction, with the aim of enhancing our understanding of charge transport mechanisms that are crucial for improving photovoltaic performance. Accordingly, this study outlines the fabrication process and electrical characterization of an Al/Ru/p-Si Schottky diode constructed on a p-type (111) silicon wafer. All electrical measurements were performed using a microcomputer-assisted setup through an IEEE-488 compatible AC/DC converter card, ensuring precision and repeatability. The analysis of the electrical behavior was carried out based on the thermionic emission theory, which serves as a foundational framework for evaluating metal-semiconductor and metal-organic semiconductor junctions. The thermionic emission theory provides a quantitative description of how charge carriers primarily electrons exceed the potential energy barrier at the metal-semiconductor interface through thermal excitation. This model enables the measurement of critical diode parameters, like the barrier height, ideality factor, and saturation current. This is important for seeing how well Schottky diodes and other optoelectronic and microelectronic devices work. Furthermore, the thermionic emission theory becomes particularly significant in cases where interface-related effects such as localized states, barrier inhomogeneities, and thin interfacial layers play a substantial role in charge transport. Deviations from ideal thermionic emission behavior under these conditions can reveal the presence of interface traps, tunneling mechanisms, or spatial variations in barrier height. Understanding and quantifying these non-idealities are essential for optimizing device architecture, especially in advanced applications including harsh or radiation-exposed environments.

In metal-organic semiconductor Schottky barrier diodes, the barrier height has also been shown to exhibit voltage dependence, further complicating the transport dynamics and necessitating more detailed modeling approaches that incorporate both intrinsic and extrinsic influences on charge flow. The thermionic emission theory (Sze, 1981) predicts that the current-voltage (I-V) characteristics as in the studies (Akay et al. 2019, Akay et al. 2020, Akay et al. 2024)

In recent years, new methods inspired by nature and human behaviour have been developed to overcome these challenges. In particular, computer

programmes that optimise things have been shown to be a good option for nonlinear regression analysis. The AI optimization algorithms, called Genetic Algorithm (GA), Differential Evolution (DE), Flower Pollination Algorithm (FPA), Artificial Bee Colony (ABC), Ant Colony Optimization (ACO), Bat Algorithm (BA), Cuckoo Search Algorithm (CSA), Grey Wolf Optimization (GWO), Jaya Algorithm (JA), Particle Swarm Optimization (PSO), Harmony Search (HS), and Teaching-Learning-Based Optimization (TLO), stand out with their ability to effectively explore large search spaces, significantly overcoming the local minimum traps encountered by traditional algorithms. Moreover, these algorithms operate without requiring derivative information and are well-suited for parallel processing, thereby demonstrating high computational performance even with large datasets. Numerous studies in the literature have shown that AI optimization algorithms are better than traditional methods because they are more accurate and faster (Açıkgöz, 2025).

The main aim of this study is to comprehensively investigate the applicability of AI optimization algorithms in addressing the parameter estimation problem encountered in nonlinear regression models. In this study, twelve AI optimization algorithms were applied on a parametrically structured nonlinear model called Schottky diode model. Each algorithm's performance was comparatively analyzed in terms of error metrics ($MAPE$ and R^2), computation times, and solution stability. The rest of the paper is organized as follows. Nonlinear regression model of Schottky diode and applied AI optimization algorithms are explained in Materials and Methods section. The analysis results are presented in Results section. Finally, conclusion is given in the last section.

MATERIAL AND METHODS

Material

Nonlinear regression models are powerful statistical tools that provide more realistic and reliable results in situations where the assumption of a linear relationship between dependent and independent variables is invalid. (Akgün, 2018). The general form of a nonlinear regression model can be expressed as,

$$Y_i = f(\mathbf{X}_i, \mathbf{\hat{a}}) + \varepsilon_i \quad , \quad i = 1, 2, \dots, n \quad (1)$$

where, Y_i represents the dependent variable, \mathbf{X}_i denotes the vector of independent variables, $\mathbf{\hat{a}}$ is the parameter vector to be estimated, and ε_i is the error term with zero mean and constant variance. The objective of the model is to determine the parameter vector $\mathbf{\hat{a}}$ in such a way that the

model output best fits the observed data (Khuri and Cornell, 1996). These parameters are estimated numerically by minimizing an objective function based on the sum of squared errors

$$\min \varphi(\boldsymbol{\beta}) = \sum_{i=1}^n [Y_i - f(X_i, \boldsymbol{\beta})]^2. \quad (2)$$

According to the optimization model given in equation (2), the nonlinear f function is defined for the electrical characterization of organic semiconductor materials as follow

$$I = I_0 \exp\left(\frac{qV_0}{nkT}\right) \left[1 - \exp\left(\frac{-qV_0}{nkT}\right)\right]. \quad (3)$$

Here, q is an electron charge, k is the Boltzmann's constant, T is an absolute temperature, n is the ideality factor, I_0 is a reverse saturation current and is equal to $A^*T^2 \exp\left(\frac{-q\Phi_b}{kT}\right)$ in which Φ_b is a barrier height,

A^* is a Richardson Constant. And also, $V_0 = V - IR_s$, where R_s is a

series resistance and V is considered as an input variable while the output variable called I , presented in equation (3). In this study, it is aimed to estimate the key electrical parameters n , R_s and Φ_b of Schottky-barrier solar cells with an Al/rubrene/p-Si structure. Here, parameter vector is defined as $\hat{\mathbf{a}} = [n \ R_s \ \Phi_b]$. Directly estimating the parameter vector $\boldsymbol{\beta}$ from experimental data is highly challenging due to measurement because of the Cobalt-60 gamma radiation uncertainties. At this point, the AI optimization algorithms offer a more reliable alternative to traditional methods, thanks to their derivative-free nature, robustness against initial conditions, and ability to navigate complex landscapes with multiple optima.

In this study, twelve different AI optimization algorithms were applied to the I–V data derived from the described model. Each AI algorithm was assessed for its ability to accurately and consistently estimate the model parameters. These algorithms utilize heuristic strategies in their solution search processes, enabling them to reach both global and local optima. As such, this approach offers not only high accuracy but also computational efficiency and stability, thereby establishing a robust modeling framework.

Methods

The Collection of the Data

The following experimental sets out the fabrication process of an Al/Ru/p-Si Schottky diode on a p-type (111) silicon wafer, with a thickness of $280\text{ }\mu\text{m}$ and a resistivity of $10\text{ }\Omega\cdot\text{cm}$. The wafer was initially subjected to a thorough cleansing process employing the RCA chemical cleaning method. Detailed information about experiment can be seen in the studies of Akay et al. 2019, Akay et al. 2020, Akay et al. 2024. All measurements were conducted with the assistance of a microcomputer via an IEEE-488 AC/DC converter card. The relationship between input (V) and response (I) values can be seen in Figure 1. It can be easily said from Figure 1 that the input and the response values have nonlinear relationship.

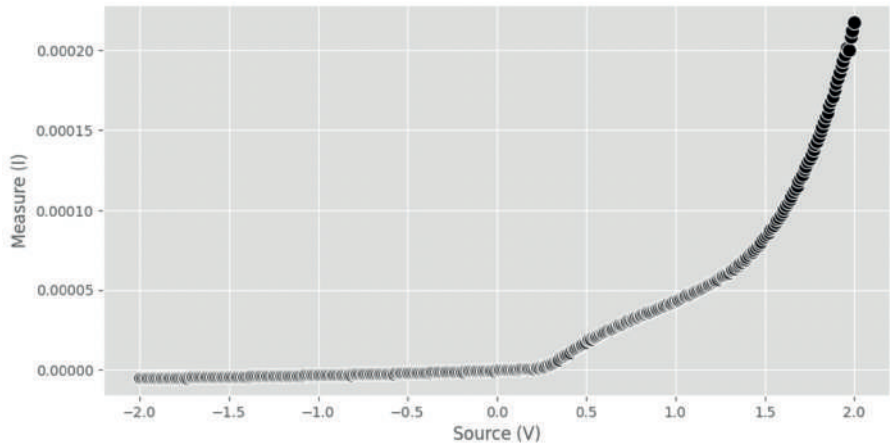


Figure 1. The relationship of input (V) and response (I) values

Statistical Analysis

In this study, the parameter estimation process of a nonlinear electrical model based on experimentally obtained current-voltage (I–V) data was carried out using AI-based algorithms, rather than traditional derivative-based optimization methods. Due to the nonlinear nature of β as well as the inherent uncertainties in the measurement data, classical methods often fail to provide reliable estimations. These types of problems typically involve a complex solution space with multiple local minima and are highly sensitive to initial values. Thus, the AI optimization algorithms, inspired by various disciplines, were used in this study. These algorithms are heuristic approaches derived from nature, social behavior, or learning processes and

aim to reach higher-quality solutions through an evolutionary process, starting with a numerical population. Each algorithm evaluates candidate solutions searching the problem space based on its internal mechanism, improves them, and generates new candidates to reach the optimal solution. Generally, an initial population of randomly generated individuals is created, followed by information exchange and variation operations among selected individuals. This process iteratively continues to enhance the quality of the solutions step by step. A general working schema of the AI optimization algorithms is presented in Figure 2.

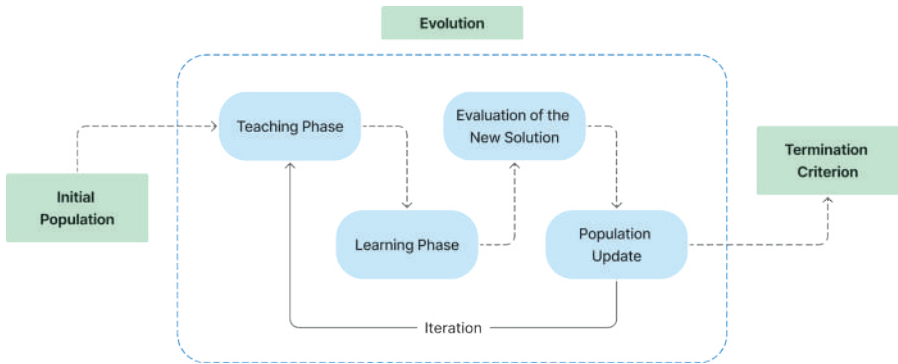


Figure 2. A general schema about working principles of the AI optimization algorithms

Each of these algorithms has been configured to perform parameter estimation for the model built on experimental I–V data, with tuning parameters, carried out for the problem at hand. This process involved optimizing tuning parameters such as population size, number of iterations, crossover rates, learning coefficients, and other control variables to maximize each algorithm's performance. Thus, each algorithm was enabled to function in the most efficient manner appropriate to its structure.

In this study, twelve AI algorithms have been applied in a comparable manner with the aim of minimizing the error function. The tuning parameter values of each AI algorithms are defined by using data-driven expert knowledge and commonly used in the literature. Through the *MAPE* and *R²* metrics, the relative performance of each AI algorithm in terms of accuracy and stability is analyzed. This methodological approach not only enhanced the accuracy of parameter estimation but also provided a robust framework for how AI optimization algorithms can be effectively applied to complex engineering-based modeling problems. In modeling the electrical behavior of electronic structures exposed to external factors such as radiation, these

flexible and data-driven optimization approaches offer significant potential for both academic research and industrial applications.

RESULTS

This study aimed to estimate the electrical parameters of Al/rubrene/p-Si Schottky-barrier solar cells after exposure to Cobalt-60 gamma irradiation within the framework of a nonlinear model with minimum error. The optimization results are presented in Table 1. It can be said from Table 1 that the ABC, the DE and the TLO have better performance according to the R^2 metric. The ACOR has the best performance for computing time. It is seen from the results that model parameters estimation is achieved by using the AI optimization algorithms.

Table 1. Optimization results of the AI algorithms with model parameters estimation

AI	Iteration Number	Total Working Time	Iteration Time	Performance metrics		Model parameters estimation		
				MAPE	R^2	Φ_b	n	R_s
ABC	93	1.3182	0.0126	0.330466	0.702430	0.672212	5.000	100.00
ACOR	67	0.3007	0.0034	0.332134	0.694962	0.672283	5.000	41.31
BA	32	0.5016	0.0046	0.597449	0.395611	0.696066	4.547	45.64
CSA	16	0.7185	0.0069	0.342103	0.666398	0.674036	4.878	54.19
DE	56	0.6041	0.0059	0.330466	0.702430	0.672212	5.000	100.00
FPA	86	0.6742	0.0064	0.333669	0.699231	0.672529	4.988	88.39
GA	99	0.7695	0.0075	0.437047	0.641893	0.683048	4.650	99.69
GWO	100	0.6316	0.0062	0.330507	0.702430	0.672215	5.000	100.00
HS	72	0.6773	0.0065	0.348827	0.691971	0.673716	4.960	65.67
JA	8	0.5124	0.0049	0.338779	0.695660	0.672702	5.000	43.80
PSO	38	0.6175	0.0073	0.330492	0.702430	0.672214	5.000	100.00
TLO	51	1.3489	0.0135	0.330466	0.702430	0.672212	5.000	100.00

These results show that the AI optimization algorithms used in the study not only cover a broader range of parameter values compared to classical methods but also offer more reliable estimations with lower error and higher explanatory power. This highlights the applicability of the AI optimization algorithms in the electrical characterization of organic semiconductor structures, demonstrating significant advantages over classical analytical approaches in terms of accuracy, flexibility, and computational efficiency.

DISCUSSION AND CONCLUSION

In this study, a comprehensive analysis was conducted by integrating nonlinear regression models with the AI optimization algorithms for the electrical characterization of organic solar cells. The results not only revealed the effects of Cobalt-60 gamma irradiation on device performance but also demonstrated that electrical parameters derived from experimental data can be accurately and reliably estimated. The nonlinear modeling approach effectively represented complex system behaviors that are difficult to explain using classical methods, clearly revealing the changes in key parameters, presented as β .

The AI optimization algorithms not only provided balanced performance in modeling but also introduced flexibility into the parameter estimation process by overcoming the limitations associated with traditional techniques. Through AI-based methods, efficient exploration of wide search spaces became feasible, and the derivative-free structure enhanced the generalizability of the modeling process. The findings demonstrate that the AI optimization algorithms offer a robust alternative for analyzing organic electronic devices operating under environmental stress conditions, such as radiation. In this context, AI optimization algorithms are considered applicable to industrial scenarios, particularly for real-time parameter estimation and device performance monitoring.

In conclusion, this study demonstrates the effectiveness of AI-supported parameter estimation processes for both academic research and industrial applications and provides a solid methodological foundation to guide future work in this field.

References

- Açıkgöz M, 2025. Yapay Zeka Optimizasyon Algoritmaları ile Doğrusal Olmayan Regresyon Modellerinde Parametre Tahmini. Yüksek Lisans Tezi, Ankara.
- Akay D, Alkan S, Altin Z, Gökmen U, Ocak SB, 2024. Electrical characterization of Al/SnO₂/PbO/Si double layer MOS under the moderate radiation effect. *Radiat. Phys. Chem.* 216, 111392.
- Akay D, Gokmen U, Ocak SB, 2019. Radiation-induced changes on poly (methyl methacrylate) (PMMA)/lead oxide (PbO) compositenanostructure. *Phys. Scr.* 2019, 94 (11), 115302.
- Akay D, Gokmen U, Ocak SB, 2020. Ionizing radiation influence on rubrene-based metal polymer semiconductors: direct information of intrinsic electrical properties. *JOM*, 72 (6), 2391–2397.
- Akgün F, 2018. Doğrusal Olmayan Regresyon Model Parametrelerinin Nokta ve Aralık Tahmini İçin Bir Yaklaşım. Yüksek Lisans Tezi, Ankara.
- Gallant AR, 1975. Nonlinear Regression. *Journal of American Statistical Association*, 29(2), 73-81.
- Khuri AI, Cornell JA, 1996. Response surfaces: Designs and analyses (2nd ed.). CRC Press.
- Sze SM, 1981. *Physics of Semiconductor Devices*, 2d ed., John Wiley, New York.
- Türkşen Ö, 2023. Optimizasyon Yöntemleri ve MATLAB, Python, R Uygulamaları. Nobel Yayınevi, Ankara.

Acknowledgment

The AI algorithms, included in this study, are within the scope of the master's thesis prepared by the first author under the supervision of the third author.

Conflict of Interest

The authors have declared that there is no conflict of interest.

Author Contributions

Murat Açıkgöz: Writing – original draft, review & editing, Statistical Analysis, Methodology, Investigation, Interpretation

Defne Akay: Writing – original draft, review & editing, Obtaining data, Investigation, Interpretation, Conceptualization

Özlem Türkşen: Writing – original draft, review & editing, Investigation, Interpretation, Methodology, Conceptualization

Current Approaches in Applied Statistics - II

Editors:

Assoc. Prof. Yalçın TAHTALI

Assoc. Prof. İbrahim DEMİR

Assist. Prof. Lütü BAYYURT

Assoc. Prof. Samet Hasan ABACI