



Institute of Pure and Applied Sciences

International Journal of Advances
in Engineering and Pure Sciences

MARMARA JEPS

UYIK 2024 SPECIAL ISSUE

(UYIK 2024 ÖZEL SAYISI)



"PREFACE - Special Issue"

This special issue contains 10 papers submitted by the participants of the *V. International Applied Statistics Congress*, held in Istanbul from May 21 to May 23, 2024.

The congress aimed to bring together researchers using statistical methods across diverse disciplines. Organized in honor of the author, economist, and academician Alev ALATLI, the event was hosted by Marmara University in collaboration with Tokat Gaziosmanpaşa University, Marmara University, and Selçuk University.

The papers included in this issue can be categorized into the following topics: applied statistics, computational statistics, statistical analysis, statistical data science, and spatial statistics.

As the guest editor, I would like to cordially thank all authors for their valuable contributions and the reviewers for their insightful comments, which significantly enhanced the quality of this issue.

I would also like to express my sincere appreciation to the Editor-in-Chief of the International Journal of Advances in Engineering and Pure Sciences, Asst. Prof. Dr. Merve ER, and all the co-editors for their kind support throughout this process.

We hope that the readers will find the papers of this special issue both informative and inspiring.

Asst. Prof. Dr. Gizel BAKICIERLER SEZER

International Applied Statistics Congress – Special Issue**CONFERENCE CHAIRMANS**

Prof. Dr. Rüstem CANGI
Tokat Gaziosmanpaşa University, Türkiye

Prof. Dr. Deniz İNAN
Marmara University, Türkiye

Prof. Dr. Semahat KÜÇÜKKOLBAŞI
Selçuk University, Türkiye

Prof. Dr. Soner ÇANKAYA
Ondokuz Mayıs University, Türkiye

Assoc. Prof. Dr. Yalçın TAHTALI
Tokat Gaziosmanpaşa University, Türkiye

SPECIAL ISSUE FIELD EDITOR

Asst. Prof. Dr. Gizel BAKICIERLER SEZER

JOURNAL EDITOR-IN-CHIEF

Asst. Prof. Dr. Merve ER

JOURNAL EDITORS

Prof. Dr. Bulent EKİCİ
Prof. Dr. Hayriye KORKMAZ

PUBLISHING EDITOR

Oğuz Eren SÜSLÜ

İÇİNDEKİLER/ CONTENTS

1. Robust Function-on-Function Regression: A Penalized Tau-based Estimation Approach

Ufuk BEYAZTAŞ

Sayfa/Page: 1-6

2. A Robust Approach using M-Estimation for Dynamic Panel Autoregressive Model

Beste Hamiye BEYAZTAŞ

Sayfa/Page: 7-20

3. Türkiye, Hindistan ve Brezilya'daki Petrol Kiralarının Otoregresif Kesirli Bütünleşik Hareketli Ortalama (ARFIMA) Modeli ile Belirlenmesi

Semanur SARIÇAM, Barış AŞIKGİL

Sayfa/Page: 21-29

4. Performance of Deep Learning Models on Imputed Time Series Data: A Simulation Study and Application to Leading Airline Companies' Stock Price

İsmail YENİLMEZ, Kürşat ATMACA

Sayfa/Page: 30-39

5. Mekânsal Nokta Verilerinin Mekânsal İstatistiklerle Analizi: Yenilenebilir Enerji Kaynakları Uygulaması

Mehmet Burak BARTAN, Günseli AYTAÇ CANKURTARAN, Özlem TÜRKŞEN

Sayfa/Page: 40-49

6. Türkiye'deki Ekonomik Büyüme Üzerine Gecikmesi Dağıtılmış Otoregresif Sınır Testi Yaklaşımı

Sinem GEÇGEL KARAGÖL, Barış AŞIKGİL

Sayfa/Page: 50-58

7. Türkiye'de Zeytin Fiyatı ile Zeytin Üretim Alanı ve Üretim Miktarı Arasındaki İlişkinin Toda-Yamamoto Testi ile Belirlenmesi

Kaan KAPLAN, Halil KIZILASLAN

Sayfa/Page: 59-64

8. Analysis of Artificial Intelligence Methods in Classifying Heart Attack Risk: Black-Box Models vs. Glass-Box Models

Ebru GEÇİCİ, Eyüp Ensar IŞIK, Mısra ŞİMŞİR, Mehmet GÜNEŞ

Sayfa/Page: 65-76

9. Examining the Relationship Between Climate Change, Energy Consumption, and the Financial System: The Case of E7 Countries

Naime İrem DURAN, Özge DEMİRKALE

Sayfa/Page: 77-87

10. Elektrokimyasal Ayırma Yöntemi ile Sulu Çözeltilerden Bor Giderim Prosesinin Optimizasyonu

Sevgi POLAT

Sayfa/Page: 88-95

Robust Function-on-Function Regression: A Penalized Tau-based Estimation Approach

Ufuk BEYAZTAŞ¹ 

¹Marmara Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 34722, İstanbul, Türkiye

Abstract

This study introduces a novel penalized estimation method tailored for function-on-function regression models, combining the robustness of the Tau estimator with penalization techniques to enhance resistance to outliers. Function-on-function regression is essential for modeling intricate relationships between functional predictors and response variables across diverse fields. However, traditional methods often struggle with outliers, leading to biased estimates and diminished predictive performance. Our proposed approach addresses this challenge by integrating robust Tau estimation with penalization, promoting both robustness and parsimony in parameter estimation. Theoretical foundations of the penalized Tau estimator within function-on-function regression are discussed, along with empirical validations through simulation studies and an empirical data analysis. By incorporating penalization, our method not only ensures robust estimation of regression parameters but also promotes model simplicity, offering enhanced interpretability and generalization capabilities in functional data analysis.

Keywords: Functional data, Penalization, Regression, Tau estimator

Öz

Bu çalışma, fonksiyon-fonksiyon regresyon modellerine yönelik yeni bir cezalandırılmış tahmin yöntemini tanıtmaktadır ve Tau tahmin edicisinin sağlamlığını cezalandırma teknikleriyle birleştirerek aykırı değerlere karşı direnci artırmaktadır. Fonksiyon-fonksiyon regresyon, fonksiyonel bağımsız değişkenler ile yanıt değişkenleri arasındaki karmaşık ilişkileri modellemek için çeşitli alanlarda gereklidir. Ancak geleneksel yöntemler genellikle aykırı değerlerle başa çıkmakta zorlanır ve bu durum yanlış tahminlere ve zayıf tahmin performansına yol açar. Önerilen yaklaşımımız, sağlam Tau tahminini cezalandırma ile birleştirerek bu zorluğun üstesinden gelmekte ve parametre tahmininde hem sağlamlığı hem de tutarlılığı sağlamaktadır. Fonksiyon-fonksiyon regresyon içinde cezalandırılmış Tau tahmin edicisinin teorik temelleri tartışılmakta, simülasyon çalışmaları ve ampirik veri analizleri yoluyla ampirik doğrulamalar sunulmaktadır. Cezalandırmayı dahil ederek, yöntemimiz yalnızca regresyon parametrelerinin sağlam tahminini sağlamakla kalmaz, aynı zamanda modelin basitliğini teşvik ederek fonksiyonel veri analizinde daha iyi yorumlanabilirlik ve genelleme yetenekleri sunar.

Anahtar Kelimeler: Fonksiyonel veri, Cezalandırma, Regresyon, Tau tahmincisi

I. INTRODUCTION

Recent advances in data collection technology have markedly enhanced access to high-dimensional, intricately structured datasets, known as functional data. Consequently, there is a growing demand for analytical tools designed for functional data analysis. For a comprehensive review of the latest theoretical and practical advancements in this field, refer to [1], [2], and [3]. One notable method within this domain is function-on-function regression (FoFR), which has gained popularity for examining the relationships between a response and predictors, where both the response and predictors are represented as random curves.

Let (Y, X) denote a pair where Y is the response and X is the predictor. Here, Y and X are presumed to be stochastic processes whose elements belong to the L_2 Hilbert space; specifically, $Y \in L_2(I)$ and $X \in L_2(S)$, where I and S are bounded and closed intervals on the real line. Consider $\{Y_i(t), X_i(s); i=1, \dots, n\}$ to be a random sample (Y, X) , with $t \in I$ and $s \in S$. The FoFR is then defined as follows:

$$Y_i(t) = \alpha(t) + \int_{-\infty}^S X_i(s) \beta(s, t) ds + \epsilon_i(t), \quad (1)$$

where $\alpha(t)$ is the constant function, $\beta(s, t)$ is the slope function, and $\epsilon_i(t)$ is the functional noise. We presume that this functional noise is independent of the predictor variable.

The main objective in model (1) is to estimate the slope function $\beta(s,t)$. Various methodologies have been developed for this purpose, including those by [4], [1], [5], [6], [7], and [8]. However, these methods typically rely on least squares estimation procedures, which are significantly affected by outliers—observations that deviate markedly from the bulk of the data. In the presence of outliers, least squares type estimators may cause biased estimates and unreliable inference.

To address this issue, several robust estimation methods for the slope function in model (1) have been proposed. In their work, [9] developed a regression model that maintains Fisher consistency and incorporates a decomposition method using functional principal components for the observed functions. In a different study, [10] introduced a method for robustly estimating model parameters using functional partial least squares, designed to handle outliers effectively. However, the efficacy of both approaches hinges on the selected basis dimension for the predictor. This choice dictates the degree of smoothness in the estimated functional parameter. As referenced in [7], this decision can lead to significant under-smoothing, particularly when the functional parameter inherently possesses a smoother nature in contrast to the higher-order components obtained from the partial least squares and principal component methods.

To achieve robust and smooth estimates for the regression coefficient function $\beta(s,t)$, [11] recently proposed a robust penalized M-estimation strategy. Their numerical analyses demonstrated that the robust penalized M-estimator provides improved parameter estimates and model inferences in the presence of outliers compared to available methods. Conversely, the robust penalized M-estimator presented by [11] lacks the integration of both a high breakdown point and high efficiency. The breakdown point measures the estimator's resistance to the influence of outliers, whereas high efficiency refers to the estimator maintaining a variance comparable to that of the least squares estimator under normal distribution conditions, as elaborated by [12].

In this study, we propose a robust penalized tau estimator designed to yield a smooth and robust estimate for $\beta(s,t)$, integrating both a high breakdown point and high-efficiency characteristics. The method represents the slope functions using a tensor product of B-spline expansion. In addition, the quadratic penalties are applied to the expansion coefficients to ensure smooth estimates. The regression functions are obtained using the tau estimator from [13], known for its high breakdown point and asymptotic efficiency under normal conditions. Our approach surpasses unpenalized estimators by ensuring a level of smoothness that mitigates overfitting. Unlike the M-estimator described by [11], our method produces

estimates that maintain both high breakdown point and asymptotic efficiency under normality. The optimal smoothness degree is governed by the penalization term, and the optimum values of the smoothing parameters are determined through a grid-search approach with the Bayesian Information Criterion (BIC).

The rest of this paper is structured as follows. Section 2 introduces the proposed robust penalized tau estimator. In Section 3, the empirical performance of the proposed method is evaluated via Monte-Carlo experiments. Section 4 presents the results of empirical data analysis results. Finally, Section 5 concludes the paper.

II. METHODOLOGY

Let us consider the FoFR model in (1). To derive penalized estimates for the model's parameters, specifically $\alpha(t)$ and $\beta(s,t)$, we tackle the following minimization problem:

$$\underset{\beta_0, \beta}{\operatorname{argmin}} \sum_{i=1}^n \rho[Y_i(t) - \alpha(t) - \int_S X_i(s)\beta(s,t)ds] + \frac{\lambda_1}{2} J_1(\alpha) + \frac{\lambda_2}{2} J_2(\beta). \quad (2)$$

Here, ρ stands for a loss function, J_1 and J_2 represent penalty functions applied to α and β , respectively, and λ_1 and λ_2 serve as smoothing parameters regulating the degree of shrinkage in α and β , respectively.

To derive estimates for α and β , we initially adopt the basis representation approach for the functional random variables, akin to the methodologies outlined in [7], [11], and [14]. Initially, we express $Y_i(t) = Y_i(t_{ij})$ and $X_i(s) = X_i(s_{ir})$, where $j = 1, \dots, M_i$ and $r = 1, \dots, G_i$, representing the number of observations on the response and predictor, respectively. For the basis representation of the functional objects in (2), we presume that $\beta_0(t)$ follows a B-spline basis expansion with K_0 basis expansion functions; $\alpha(t) = \sum_{k=1}^{K_0} \alpha_k \phi_k(t)$, where $\phi_k(t)$ (for $k = 1, \dots, K_0$) denotes the B-spline basis expansion function, and α_k signifies the expansion coefficient. Additionally, we posit that $\beta(s,t)$ adopts a basis expansion representation with the truncation constants K_y and K_x as follows:

$$\beta(s,t) = \sum_{l=1}^{K_y} \sum_{p=1}^{K_x} b_{lp} \psi_l(t) \vartheta_p(s), \quad (3)$$

where $\psi_l(t)$ and $\vartheta_p(s)$ denote the expansion functions and b_{lp} denotes the corresponding expansion coefficient. Let $\Delta_r = s_{r+1} - s_r$ represent the length of r -th interval in S . Subsequently, the integral component in the minimization problem (3) can be approximated using numerical integration, which can be expressed as follows:

$$\begin{aligned} \int_s X_i(s)\beta(t,s)ds &\approx \sum_{r=1}^{G-1} \Delta_r \beta(t,s_r) X_i(s_r) \\ &= \sum_{r=1}^{G-1} \Delta_r \sum_{l=1}^{K_y} \alpha_k \sum_{p=1}^{K_x} b_{lp} \psi_l(t) \vartheta_p(s_r) X_i(s_r), \\ &= \sum_{l=1}^{K_y} \sum_{p=1}^{K_x} b_{lp} \psi_l(t) \tilde{\vartheta}_{p,i} \end{aligned} \quad (4)$$

where $\tilde{\vartheta}_{p,i} = \sum_{r=1}^{G-1} \Delta_r \vartheta_p(s_r) X_i(s_r)$. Replacing the basis expansion approximation in (1) gives:

$$Y_i(t) = \sum_{k=1}^{K_0} \alpha_k \phi_k(t) + \sum_{l=1}^{K_y} \sum_{p=1}^{K_x} b_{lp} \psi_l(t) \tilde{\vartheta}_{p,i}. \quad (5)$$

For the penalty functionals, $J_1(\alpha)$ and $J_2(\beta)$, we employ quadratic penalties based on the second derivatives of the basis functions. Let $\alpha = [\alpha_1, \dots, \alpha_{K_0}]^T$ represent the vector consisting of expansion coefficient for the constant function. Then, the penalty functional $J_1(\alpha)$ is approximated as $\tilde{J}_1(\alpha) = \int_I [\alpha^{(2)}(t)]^2 dt = \alpha^T \mathbf{P}_\alpha \alpha$, where $\alpha^{(2)}(t)$ denotes the second derivative of $\alpha(t)$, and \mathbf{P}_α denotes penalty matrix with dimension $K_0 \times K_0$ with entries $\int_I \phi_k^{(2)}(t) \phi_l^{(2)}(t) dt$. Let $\mathbf{b} = (b_{lp})_{lp}$ denote the $K_y \times K_x$ dimensional basis expansion coefficients. Then, the penalty functional $J_2(\beta)$ is approximated as follows:

$$\tilde{J}_2(\beta) = \int_I \int_s \left[\frac{\partial^2}{\partial t^2} \beta(t,s) \right]^2 ds dt + \int_I \int_s \left[\frac{\partial^2}{\partial s^2} \beta(t,s) \right]^2 ds dt = \mathbf{b}^T (\boldsymbol{\vartheta} \otimes \mathbf{P}_y + \mathbf{P}_x \otimes \boldsymbol{\psi}) \mathbf{b}, \quad (6)$$

where $\boldsymbol{\psi} = \int_I \boldsymbol{\psi}(t) \boldsymbol{\psi}(t)^T dt$ with $\boldsymbol{\psi}(t) = [\boldsymbol{\psi}_1(t), \dots, \boldsymbol{\psi}_{K_y}(t)]^T$, $\boldsymbol{\vartheta} = \int_s \boldsymbol{\vartheta}(s) \boldsymbol{\vartheta}(s)^T ds$ with $\boldsymbol{\vartheta}(s) = [\boldsymbol{\vartheta}_1(s), \dots, \boldsymbol{\vartheta}_{K_x}(s)]^T$. Here, \mathbf{P}_y and \mathbf{P}_x are the matrices of penalty terms whose elements are computed from the derivatives of the expansion coefficients.

Subsequently, leveraging the approximate penalty functionals and the expansion representation of the functionals, we can rewrite the minimization problem in (2) as follows:

$$\underset{\alpha, \mathbf{b}}{\operatorname{argmin}} \sum_{i=1}^n \sum_{j=1}^M \rho \left[Y_i(t_j) - \phi^T(t_j) \alpha - (\tilde{\boldsymbol{\vartheta}}_i^T \otimes \boldsymbol{\psi}^T(t_j)) \mathbf{b} \right] + \frac{\lambda_1}{2} \tilde{J}_1(\alpha) + \frac{\lambda_2}{2} \tilde{J}_2(\beta). \quad (7)$$

To obtain robust estimates for α and \mathbf{b} , we consider the τ -estimator proposed by [13]. Let $\boldsymbol{\theta} = [\alpha^T, \mathbf{b}^T]^T$, $\boldsymbol{\Pi} = [\boldsymbol{\Pi}_1, \dots, \boldsymbol{\Pi}_n]^T$ with $\boldsymbol{\Pi}_i = [\boldsymbol{\phi}^T(t) \tilde{\boldsymbol{\vartheta}}_i^T \otimes \boldsymbol{\psi}^T(t)]^T$, and $\mathbf{P}(\lambda_1, \lambda_2)$ is a block diagonal matrix whose elements are $\lambda_1 \mathbf{P}_\alpha$ and $\lambda_2 (\boldsymbol{\vartheta} \otimes \mathbf{P}_y + \mathbf{P}_x \otimes \boldsymbol{\psi})$. Subsequently, we consider the following optimization problem in a matrix form:

$$\operatorname{argmin}_{\boldsymbol{\theta}} \sum_{i=1}^n \rho[Y_i(t) - \boldsymbol{\Pi}_i \boldsymbol{\theta}] + \mathbf{P}(\lambda_1, \lambda_2) \boldsymbol{\theta} \quad (8)$$

The τ -estimator for $\boldsymbol{\theta}$ is defined as follows:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \tau(\boldsymbol{\theta}), \quad (9)$$

where the τ -scale estimator $\tau(\boldsymbol{\theta})$ is given by

$$\tau^2(\boldsymbol{\theta}) = s^2(\boldsymbol{\theta}) \frac{1}{n u_2} \sum_{i=1}^n \rho_2 \left[\frac{Y_i(t) - \boldsymbol{\theta}^T \boldsymbol{\Pi}_i}{s(\boldsymbol{\theta})} \right], \quad (10)$$

with $s(\boldsymbol{\theta})$ is an M-estimator that solves $\frac{1}{n} \sum_{i=1}^n \rho_1 \left[\frac{Y_i(t) - \boldsymbol{\theta}^T \boldsymbol{\Pi}_i}{s(\boldsymbol{\theta})} \right] = u_1$. Here, the loss functions ρ_1 and ρ_2 are symmetric, continuously differentiable, and bounded functions. The parameters u_1 and u_2 , conversely, act as tuning parameters utilized to achieve consistency under normally distributed error terms. The selection of loss functions ρ_1 and ρ_2 holds significant practical and theoretical relevance. In this investigation, we adopt the optimal loss function proposed by [15]:

$$\rho(v) = \begin{cases} 1.38 \left(\frac{v}{c}\right)^2, & \left|\frac{v}{c}\right| \leq \frac{2}{3} \\ 0.55 - 2.69 \left(\frac{v}{c}\right)^2 + 10.76 \left(\frac{v}{c}\right)^4 - 11.66 \left(\frac{v}{c}\right)^6 + 4.04 \left(\frac{v}{c}\right)^8, & \frac{2}{3} < \left|\frac{v}{c}\right| \leq 1 \\ 1, & \left|\frac{v}{c}\right| > 1 \end{cases} \quad (11)$$

Following the recommendation of [15], we opt for $c_1 = 1.214$ and $b_1 = 0.5$ for ρ_1 and $c_2 = 3.270$ and $b_2 = 0.128$ for ρ_2 . With these parameter selections, the τ -estimator achieves a 50% breakdown and 95% efficiency under normally distributed error terms, as demonstrated by [15].

The τ -estimator $\hat{\boldsymbol{\theta}}$ is computed via an iterative algorithm. We employ a random resampling-based fast estimation algorithm for this purpose. Initially, random subsamples are drawn from the entire dataset. For each subsample, the iteratively reweighted least squares algorithm is iterated multiple times to obtain potential estimates. This process continues until convergence, and the final estimator is selected from the potential estimates, providing the minimum scale estimate. Let $\hat{\boldsymbol{\theta}} = [\hat{\boldsymbol{\alpha}}, \hat{\mathbf{b}}]^T$ represent the τ -estimate of $\boldsymbol{\theta}$. Subsequently, the robust estimates of the intercept function and bivariate coefficient function are obtained as follows:

$$\hat{\alpha}(t) = \hat{\boldsymbol{\phi}}^T(\hat{\boldsymbol{\alpha}}), \quad \hat{\beta}(t,s) = [\boldsymbol{\vartheta}^T(s) \otimes \boldsymbol{\psi}^T(t)] \hat{\mathbf{b}}. \quad (12)$$

III. MONTE CARLO SIMULATIONS

We implement a series of simulations to assess the estimation and predictive performance of the proposed method, referred to as "tau." This method's empirical performance is benchmarked against functional principal component regression (fpcr), functional partial least squares regression (fpls), and the penalized function-on-function linear regression model introduced by [7] (pffr). For the simulations, we adopt the data generation process outlined in [11].

In the data generation process, the predictor is created at fifty equally-spaced grid points within the unit interval. Similarly, the elements of response variable are simulated at sixty equally distance points within the same interval. The generation of the functional predictor follows:

$$X_i(s) = \sum_{r=1}^{50} \frac{1}{r^2} \{ \zeta_{i1,r} \sqrt{2} \sin(r \pi s) \zeta_{i2,r} \sqrt{2} \cos(r \pi s) \},$$

where $\zeta_{i1,r}$ and $\zeta_{i2,r}$ i.i.d. random variables from the normal distribution with zero mean and unit variance. Following this, the functional response is generated using the specified methodology:

$$Y_i(t) = \alpha(t) + \int_0^1 X_i(s) \beta(s, t) ds + \epsilon_i(t),$$

where $\alpha(t) = 2 \exp(-(t - 1)^2)$, $\beta(s, t) = 4 \cos(2 \pi t) \sin(\pi s)$, and ϵ_i is the random noise where each $\epsilon_i(t_j) \sim N(0, 0.01)^2$. Replacing 5% and 10% of the data points with outliers, we utilize $\alpha^*(t) = 4 \exp(-(t)^2)$, $\beta^*(s, t) = 6 \sin(4 \pi t) \sin(2 \pi s)$ to generate atypical observations. Figure 1 presents a visual depiction of the generated data alongside the parameter functions utilized in the process.

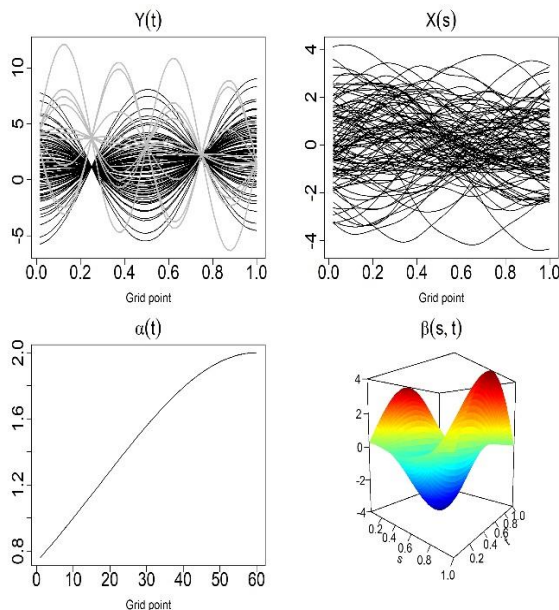


Figure 1. The graphical display showcases the response (left-top panel), predictor (right-top panel), constant function (left-bottom panel), and slope function (right-bottom panel).

In the experiments, a training sample of fixed size $n_{train} = 250$ is generated. Based on the training sample, we build the models and compute the root relative integrated squared percentage estimation errors (RISPEE) for the constant and slope functions for assessing the estimation performance of the methods:

$$RRISPEE(\hat{\alpha}) = 100 \times \sqrt{\frac{\|\hat{\alpha}(t) - \alpha(t)\|_2^2}{\|\alpha(t)\|_2^2}},$$

$$RRISPEE(\hat{\beta}) = 100 \times \sqrt{\frac{\|\hat{\beta}(s, t) - \beta(s, t)\|_2^2}{\|\beta(s, t)\|_2^2}},$$

where $\|\cdot\|_2$ denotes the L_2 norm. It's essential to mention that the methods fpcr and fpls presume that the response and predictor variables are centered, so that they have mean-zero. That is, $RRISPEE(\hat{\alpha})$ is computed only for pffr and tau methods. To assess the predictive performance of the methods, a test sample of size $n_{test} = 100$ is generated. The models fitted using the training sets are then applied to the test samples, and we compute the root mean squared percentage error (RMSPE) as follows:

$$RMSPE = 100 \times \sqrt{\frac{\|\hat{Y}(t) - Y(t)\|_2^2}{\|Y(t)\|_2^2}}$$

In the simulations, 100 Monte Carlo replications are conducted. For constructing the models using pffr, fpcr, fpls, and the proposed tau method, a fixed 15 basis expansion functions are employed.

The computed mean $RRISPEE(\hat{\alpha})$, $RRISPEE(\hat{\beta})$, and $RMSPE$, along with their standard errors, are presented in Table 1. When no outliers are present in the data, the proposed tau estimator demonstrates superior parameter estimation for the intercept function and achieves better prediction accuracy, as indicated by lower RMSPE values, compared to all other methods. This improvement may be due to the random data generation process, where a 0% contamination level might still produce small-magnitude outliers. The proposed method effectively mitigates the impact of these outliers, yielding enhanced results, whereas other non-robust methods are influenced by these outliers, resulting in biased outcomes. However, in this scenario, the tau method performs worst for $RRISPEE(\hat{\beta})$, with fpcr and fpls showing the best results. When outliers are introduced into the data, regardless of the contamination level, the proposed tau method consistently outperforms all competitors across all performance metrics. The non-robust methods exhibit significantly poorer estimation and predictive performance in the presence of outliers compared to their results with 0% contamination. In contrast, the proposed method effectively down-weights the influence of outliers, maintaining performance comparable to that achieved with no contamination.

Table 1. The mean $RRISPEE(\hat{\alpha})$, $RRISPEE(\hat{\beta})$, and $RMSPE$, along with their standard errors (given in brackets), are computed over 100 Monte Carlo replications.

%	Metric	pffr	fpcr	fpls	tau
0	$RRISPEE(\hat{\alpha})$	0.0958 (0.0940)	-	-	0.0253 (0.0052)
	$RRISPEE(\hat{\beta})$	2.3722 (0.0073)	1.7457 (0.6129)	0.8209 (0.1517)	3.5597 (1.0817)
	$RMSPE$	0.7463 (0.0279)	3.9089 (2.5550)	3.9166 (2.5646)	0.0729 (0.0035)
5	$RRISPEE(\hat{\alpha})$	6.8535 (1.1207)	-	-	0.0347 (0.0053)
	$RRISPEE(\hat{\beta})$	32.5246 (13.7395)	8.6258 (6.4341)	9.6036 (7.334)	4.1152 (1.4877)
	$RMSPE$	7.4112 (1.0931)	7.9222 (1.5537)	8.1117 (1.6796)	0.0850 (0.0059)
10	$RRISPEE(\hat{\alpha})$	12.3480 (0.6299)	-	-	0.0403 (0.0089)
	$RRISPEE(\hat{\beta})$	32.6744 (11.6223)	15.9920 (6.7921)	12.3848 (2.1659)	6.0510 (2.7341)
	$RMSPE$	12.2392 (0.8134)	12.4304 (1.0484)	12.2301 (1.0691)	0.0991 (0.0058)

Moreover, we compare the performance of the proposed method with existing non-robust methods, namely, pffr, fpcr, and fpls, in terms of their computing times. A single Monte Carlo simulation is performed with a sample size of 250, and the elapsed computing time for all the methods is recorded. The computations are executed on a desktop PC with an Intel® Core™ i5-9500 CPU at 3.00 GHz and 8 GB RAM. The computing times (in seconds) are recorded as 6.83, 0.25, 0.95, and 72.63 for pffr, fpcr, fpls, and the proposed tau method, respectively. From the results, it is evident that the classical non-robust methods require considerably less computing time than the proposed method. This result is due to the proposed method utilizing an iterative approach for estimating both model parameters and smoothing parameters (grid-search algorithm).

IV. EMPIRICAL DATA ANALYSIS

We employ the Oman weather dataset from the National Center for Statistics & Information (<https://data.gov.om>). This dataset comprises monthly maximum humidity (in percentage) and evaporation (in millimeters) measurements from 49 weather stations across Oman, spanning from January 2018 to December 2018. Each observation is treated as a function of the months, resulting in a total of $n = 49$ functional observations $\{Y(t), X(s): 1 \leq t, s \leq 24\}$. Figure 2 presents the graphical display of the functional observations.

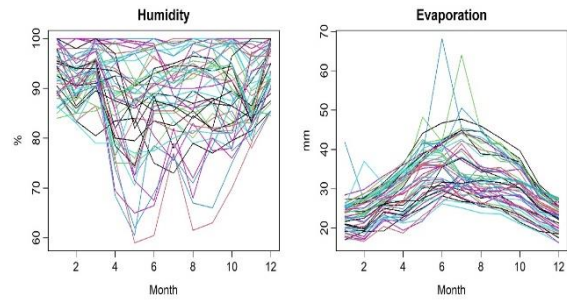


Figure 2. Graphical display of the maximum humidity (left panel) and evaporation (right panel) variables for the Oman weather data.

To assess the predictive performance of the methods, we repeat the following procedure 100 times: 1) Randomly split the dataset into training and test samples, with sizes 33 and 16, respectively. 2) Construct a model using the training sample curves, employing 8 basis functions determined by the generalized cross-validation procedure. 3) Use this model to predict 13 curves in the test sample. 4) Calculate the RMSPE for each replication to compare the predictive accuracy of the methods.

Figure 2 reveals that the Oman weather data contains clear atypical observations in the response (humidity). Hence, it's expected that the proposed robust method would deliver superior prediction results compared to its non-robust counterparts, namely pffr, fpcr, and fpls. The mean RMSPE values computed from the methods and their standard errors given in brackets are as follows: 5.9523 (0.9518) for pffr, 5.8333 (0.9255) for fpcr, 5.2234 (0.8309) for fpls, and 3.1239 (0.4478) for the proposed tau method. These findings suggest that our method achieves improved predictive performance, as indicated by lower RMSPE values compared with its competitors.

V. CONCLUSION

The FoFR model has emerged as a pivotal tool for investigating the functional relationship between a functional response and a set of functional predictor variables. Numerous methods have been put forth to accurately estimate the parameters of this model. However, many of these methods suffer from a lack of robustness and can be substantially influenced by the existence of outliers. Consequently, traditional approaches may produce biased estimates for the regression parameters, leading to subpar predictive performance.

This study introduces a novel penalized robust estimation method, named "tau," tailored to acquire outlier-resistant estimates for the slope function of the FoFR model. The proposed method employs B-spline expansion to project functional object into a finite-dimensional space and the penalty functionals obtained from their second derivatives are applied to the expansion coefficients to control the smoothness of the estimates. To assess the estimation and predictive performance, a series of Monte Carlo experiments and empirical data analyses are conducted, comparing the results favorably with existing methods. The findings indicate that the proposed method yields comparable estimation and predictive performance to existing non-robust methods in outlier-free data scenarios. However, notably, our method demonstrates improved estimation and predictive accuracy when data are contaminated by outliers, surpassing the performance of existing methods in such scenarios.

The estimation approach proposed in this study can be extended in several research directions. For example, the current functional regression model includes only one functional predictor. The proposed method can be easily extended to models that include multiple functional predictors. Additionally, the considered model includes only a functional predictor. However, fields such as health and medicine often require both functional and scalar predictors. Therefore, the proposed method can be extended to robustly estimate model parameters when both functional and scalar predictors are included. Moreover, our Monte Carlo experiments indicate that the proposed method requires significantly more computing time than existing methods. To address this, several algorithms, such as parallel computing, can be applied to reduce the computational burden of the proposed method.

REFERENCES

- [1] Ramsay, J.O., & Silverman, B.W. (2006). *Functional data analysis*, 2nd edn. Springer, New York.
- [2] Horvath, L., & Kokoszka, P. (2012). *Inference for functional data with applications*, Springer, New York.
- [3] Kokoszka, P., & Reimherr, M. (2015). *Introduction to functional data analysis*, CRC Press, London.
- [4] Yamanishi, Y., & Tanaka, Y. (2003). Geographically weighted functional multiple regression analysis: A numerical investigation, *Journal of the Japanese Society of Computational Statistics*, 15, 307–317.
- [5] Yao, F., Müller, H.-G., & Wang, J.-L. (2005). Functional linear regression analysis for longitudinal data, *The Annals of Statistics*, 33, 2873–2903.
- [6] Matsui, H., Kawano, S., & Konishi, S. (2009). Regularized functional regression modeling for functional response and predictors, *Journal of Math-for-Industry*, 1, 17–25.
- [7] Ivanescu, A. E., Staicu, A.-M., Scheipl, F., & Greven, S. (2015). Penalized function-on-function regression, *Computational Statistics*, 30, 539–568.
- [8] Beyaztas, U., & Shang, H.L. (2020). On function-on-function regression: Partial least squares approach, *Environmental and Ecological Statistics*, 27, 95–114.
- [9] Hullait, H., Leslie, D.S., Pavlidis, N. G., & King, S. (2021). Robust function-on-function regression, *Technometrics*, 63, 396–409.
- [10] Beyaztas, U., & Shang, H.L. (2022). A robust partial least squares approach for function-on-function regression, *Brazilian Journal of Probability and Statistics*, 36, 199–219.
- [11] Cai, X., Xue, L., & Cao, J. (2021). Robust penalized M-estimation for function-on-function linear regression, *Stat* 10, e390.
- [12] Beyaztas, U., Shang, H.L., & Mandal, A. (2023). Robust function-on-function interaction regression, *Statistical Modelling*, in-press.
- [13] Yohai, V.J., & Zamar, R.H. (1988). High breakdown-point estimates of regression by means of the minimization of an efficient scale, *Journal of the American Statistical Association: Theory and Methods*, 83, 406–413.
- [14] Beyaztas, U., Shang, H.L., & Saricam, S. (2024). Penalized function-on-function linear quantile regression, *Computational Statistics*, in-press.
- [15] Salibian-Barrera, M., Willems, G., & Zamar, R. (2008). The fast- τ estimator for regression, *Journal of Computational and Graphical Statistics*, 17, 659–682.

A Robust Approach using M-Estimation for Dynamic Panel Autoregressive Model

Beste Hamiye BEYAZTAŞ ¹ 

¹*İstanbul Medeniyet Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, İstatistik Bölümü, İstanbul, Türkiye*

Abstract

This paper presents a robust M-estimation approach for first-order panel autoregressive models, addressing the challenges posed by high persistence levels of the autoregressive parameter and individual heterogeneity. Generalized method of moments estimators widely used in dynamic panel models exhibit substantial finite sample biases and are sensitive to weak instruments, particularly as the autoregressive parameter gets close to unity. Our proposed weighted M-estimator, which uses a power function for the scale parameter in Huber's loss function, offers a robust alternative. By minimizing the variance of model parameters through an optimal tuning parameter, our method enhances the efficiency and robustness of parameter estimates. We demonstrate the superiority of the proposed approach through several Monte-Carlo simulations and an application to hydro-electric power output data, providing comprehensive comparisons with existing generalized method of moments estimators.

Keywords: Panel autoregressive models, Robust estimation, M-estimation, Generalized method of moments

I. INTRODUCTION

Dynamic panel data (DPD) models play a pivotal role not only in econometrics but also in engineering and the natural sciences, serving as fundamental analytical tools, especially when dealing with data that evolves over time and across individual units. These models allow us to capture the temporal dynamics of data, thus facilitating the understanding of various phenomena in different fields of research, such as physics, biology, environmental science, engineering, econometrics and so on. For instance, DPD models can be used to analyze the long-term effects of greenhouse gas emissions, oceanic circulation patterns, and temperature fluctuations over time in the field of climatology. As another example, in biology, the growth and development of organisms, the spread of diseases, and the interactions between species in ecosystems can be investigated via these models, considering both short-term and long-term dynamics. The behavior of complex systems such as electrical circuits, mechanical systems, or chemical processes in engineering can also be addressed with these models to discover hidden patterns, relationships, and trends in data that may not be evident with simpler models.

The flexibility of DPD models highlights their importance as a powerful analytical tool across various disciplines. By incorporating the unobserved individual-specific effects, these models account for unobserved heterogeneity across individuals, leading to enhanced insights (cf. [1]). Additionally, the inclusion of lagged dependent variables as explanatory variables in DPD models is a crucial feature that distinguishes them from static panel data models. This feature enables them to capture both the short- and long-term dynamics of the data and allows for modeling of persistence within the data.

Estimating DPD models involves addressing several issues stemming from endogeneity, potential correlation between the individual-specific effects and the explanatory variables, and unobserved heterogeneity. Using well-known least squares (LS) techniques for dynamic models may result in obtaining inconsistent estimates of the parameters when dealing with panel data with a small time dimension. This inconsistency arises due to the presence of endogenous explanatory variables, which introduce correlation between the regressors and error terms. Even with large samples, LS techniques, such as fixed effects (LSDV) or random effects (GLS), may still exhibit bias, as noted in [2]. Furthermore, [3, 4] address the inconsistency of the maximum likelihood estimator (MLE) when dealing with a large number of individuals (N) and a fixed number of time periods (T), which arises from the increase in parameters with the increasing number of individuals, resulting in an incidental parameter problem. This has prompted likelihood-based approaches aimed at addressing this issue, such as the conditional likelihood estimator outlined in [5], and estimators based on the first differences, as proposed by [6, 7, 8]. Also, for a detailed discussion on the finite sample properties of the MLE within the scope of dynamic

panel data models, see [9]. More recently, likelihood-based estimators for autoregressive panel data models, which are robust in the presence of heteroskedasticity, have been proposed by [10].

The primary focus in literature for the estimation of dynamic panel data models has been on a class of generalized method of moments (GMM) estimators. GMM estimation offers a flexible approach by exploiting moment conditions derived from the sample moment counterparts of population moment conditions, often referred to as orthogonality conditions, of the data-generating model (cf. [11]). As pointed out by [11], the main sources leading to the widespread use of GMM estimators include: (i) Providing a simple approach for demonstrating the asymptotic properties of GMM estimators, and (ii) The capability to construct them without specifying the complete data generation process. For comprehensive discussions of GMM estimation with a wide range of applications, see [12, 13, 14, 15]. The GMM estimators based on first-difference transformation proposed by [16, 17, 18] have led to the beginning of an extensive literature. Although the first-difference based GMM estimators yield consistent estimates for large cross-sectional size, they exhibit substantial finite sample bias, particularly when dealing with strongly persistent data and weak instruments (cf. [1, 19, 20, 21]). In order to enhance the finite sample properties of standard GMM estimators, several alternative estimators, such as level GMM (LEV) estimator of [22] and system GMM (SYS) estimator of [19], have been developed. These estimators can be considered as extensions of the standard GMM estimators by incorporating additional moment conditions derived from the level equations for LEV estimator and from the model in first differences and levels for SYS estimator. Though exploiting many instruments, leads to an improvement in the efficiency of GMM estimators and addresses weak instrument and incidental parameter issues, as noted in [23], these estimators still exhibit bias. Moreover, the SYS estimator may result in increasingly biased estimates and weak instrument issues in the presence of a large variance ratio of the individual-specific effects to the idiosyncratic errors and/or an autoregressive coefficient that is close to unity (cf. [2, 24]). The finite sample biases of the SYS estimators have been investigated by [25]. Furthermore, [26] have proposed a consistent GMM estimator with less bias in the presence an autoregressive coefficient that is close to unity.

The aim of this paper is to develop a robust M-estimator when the value of the autoregressive parameter is near unity and/or the variance of individual effects differs from the variance of the error terms, where the class of GMM estimators is highly sensitive to the increasing level of persistence and

individual heterogeneity. In this paper, we propose an extension of the weighted M-estimation approach introduced by [27] to estimate the parameters of the first-order autoregressive panel data models. The proposed robust estimator, based on Huber's loss function, is obtained by weighting the M-estimator with a power function for the scale parameter. Also, the optimal value of the tuning parameter related to the loss function has been chosen with the aim of minimizing the variance of the model parameters and based on the data distribution, as in [28, 29].

The rest of the paper is organized as follows. In Section 2, we begin by presenting comprehensive information on first-order autoregressive panel data models and existing GMM estimators. Subsequently, we describe our approach to obtain the proposed M-estimator, which is weighted by a power function applied to the scale parameter in Huber's loss function. The finite sample properties of the proposed estimator are demonstrated through an extensive simulation study, and the results are compared with those of traditional GMM estimators in Section 3. In Section 4, we apply our proposed method to hydro-electric power output data to further validate its applicability. Finally, a few concluding remarks are provided in Section 5.

II. METHODOLOGY

2.1. First-Order Autoregressive Panel Model and GMM Estimators

We consider the first-order autoregressive panel model described as follows

$$y_{it} = \varphi y_{i,t-1} + \alpha_i + \varepsilon_{it}; \quad i = 1, \dots, N; \quad t = 2, \dots, T \quad (1)$$

where α_i , ε_{it} and y_{it} respectively represent unobserved individual-specific effects, idiosyncratic error terms and the response variable for an individual i observed at time t and φ is the autoregressive parameter under the stationarity assumption that $|\varphi| < 1$. For this simple DPD model, α_i 's are assumed to be independent and identically distributed (iid) across individuals, with $E(\alpha_i) = 0$, $Var(\alpha_i) = \sigma_\alpha^2$, and ε_{it} 's are iid across time and individuals, with $E(\varepsilon_{it}) = 0$, $Var(\varepsilon_{it}) = \sigma_\varepsilon^2$ (cf. [19]). Also, for mean stationarity on the process, it is assumed that $E(y_{i1}\varepsilon_{it}) = 0$ and $E(\alpha_i\varepsilon_{it}) = 0$ (cf. [19]). An additional assumption developed by [30] has been imposed on initial observations as follows

$$y_{i1} = \frac{\alpha_i}{1-\varphi} + \mu_{i1} \text{ for } i = 1, \dots, N \quad (2)$$

where $\mu_{i1} = \sum_{j=0}^{\infty} \varphi^j \varepsilon_{i,1-j}$ is independent of α_i . By defining $y_i = (y_{i3}, \dots, y_{iT})'$, $y_{i,-1} = (y_{i2}, \dots, y_{i,T-1})'$

and $u_i = (u_{i3}, \dots, u_{iT})'$, Equation (2) can be expressed as

$$y_i = \varphi y_{i-1} + u_i \quad (3)$$

where $u_{it} = \alpha_i + \varepsilon_{it}$.

Under the assumptions given above, we examine three commonly used GMM estimators: the first difference (DIF) GMM estimator, the LEV GMM estimator, and the SYS GMM estimator. The GMM estimators are constructed using some moment conditions, with the asymptotic covariance of these moment conditions as the weight matrix (cf. [2]). Employing two-step procedures improves the asymptotic efficiency of the standard GMM estimators. The one-step GMM estimate is derived using an initial positive semidefinite weight matrix, which is independent of estimated parameters (cf. [31]). Then, the weight matrix, which includes residuals from the one-step estimation, is used to obtain the two-step GMM estimate. Also, estimated standard errors using two-step procedure tend to show a downward bias in small samples, leading to a preference for one-step estimates with robust standard errors as noted in [1, 18, 31]. Next, we briefly discuss the one-step and two-step DIF GMM, LEV GMM and SYS GMM estimation for first-order autoregressive panel data models.

2.1.1. First difference GMM estimator

The DIF GMM estimator transforms the model, given in Equation (3), into a system of equations in first differences to address the correlation between the lagged endogenous variable ($y_{i,-1}$) and the error term (u_i) stemming from the individual effect (α_i). Therefore, to eliminate the individual effects, [18] employ the first differences of Equation (3) as follows:

$$\Delta y_i = \varphi \Delta y_{i,-1} + \Delta u_i$$

where $\Delta y_i = (y_{i3} - y_{i2}, \dots, y_{iT} - y_{i,T-1})'$, $\Delta y_{i,-1} = (y_{i2} - y_{i1}, \dots, y_{i,T-1} - y_{i,T-2})'$, and $\Delta u_i = (u_{i3} - u_{i2}, \dots, u_{iT} - u_{i,T-1})'$. By exploiting $m_D = (1/2)(T-1)(T-2)$ orthogonality conditions, $E(Z_i^D \Delta u_i) = 0$, where Z_i^D denotes a $(T-2) \times m_D$ instrumental variable matrix given below,

$$Z_i^D = \begin{pmatrix} y_{i1} & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & y_{i1} & y_{i2} & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & y_{i1} & \dots & y_{i,T-2} \end{pmatrix}$$

one-step DIF GMM estimator (DIF₁) of [18] for φ is calculated as

$$\hat{\varphi}_{DIF_1} = (\Delta y'_{-1} Z^D W^D Z^D \Delta y_{-1})^{-1} \Delta y'_{-1} Z^D W^D Z^D \Delta y$$

where $\Delta y_{-1} = (\Delta y'_{1,-1}, \dots, \Delta y'_{N,-1})'$, $Z^D = (Z_1^D, \dots, Z_N^D)'$, $\Delta y = (\Delta y'_1, \dots, \Delta y'_N)'$, and

$$W^D = \left(\frac{1}{N} \sum_{i=1}^N Z_i^D D Z_i^D \right)^{-1}$$

where D is a $(T-2) \times (T-2)$ square Toeplitz matrix as follows

$$D = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & \dots & -1 & 2 \end{pmatrix}$$

Using the residuals from DIF₁, $\Delta \hat{u}_i$, two-step DIF-GMM (DIF₂) estimator is obtained as

$$\hat{\varphi}_{DIF_2} = (\Delta y'_{-1} Z^D W_2^D Z^D \Delta y_{-1})^{-1} \Delta y'_{-1} Z^D W_2^D Z^D \Delta y$$

where W_2^D denotes the weighting matrix as $W_2^D = \left(\frac{1}{N} \sum_{i=1}^N Z_i^D \Delta \hat{u}_i \Delta \hat{u}_i' Z_i^D \right)^{-1}$.

[19] demonstrate that using lagged levels as instruments to obtain the first difference GMM estimator results in weak instrumental variables, and the instruments become invalid when φ is close to unity and/or when $\sigma_\alpha^2 / \sigma_\varepsilon^2$ increases.

2.1.2. Level GMM estimator

The primary motivation behind level GMM estimator developed by [22] has been to remove individual effects from instrumental variables to address the correlation problem between error terms, u_i and lagged endogenous variables, $y_{i,-1}$ caused by the presence of individual effects, α_i . For consistency with a large N and a fixed time period T , LEV GMM estimator requires the mean stationarity of the process, which implies that the assumption on the initial observations is satisfied.

The level model described in Equation (3) is used to obtain LEV GMM estimator, with the assumption that the $m_L = (1/2)(T-1)(T-2)$ moment conditions, $E(Z_i^L u_i) = 0$, hold. Here, Z_i^L represents a $(T-2) \times m_L$ instrumental variable matrix as follows

$$Z_i^L = \begin{pmatrix} \Delta y_{i2} & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & \Delta y_{i2} \Delta y_{i3} & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & \Delta y_{i2} & \dots & \Delta y_{i,T-1} \end{pmatrix}$$

Using these orthogonal conditions, the matrix of instruments, $Z^L = (Z_1^L, \dots, Z_N^L)'$, and weighting matrix, W^L , described as,

$$W^L = \left(\frac{1}{N} \sum_{i=1}^N Z_i^L Z_i^{L'} \right)^{-1}$$

one-step LEV GMM estimator (LEV₁) of φ is obtained as follows

$$\hat{\varphi}_{LEV_1} = (y'_{-1} Z^L W^L Z^{L'} y_{-1})^{-1} y'_{-1} Z^L W^L Z^{L'} y$$

where $y_{-1} = (y'_{1,-1}, \dots, y'_{N,-1})'$ and $y = (y'_1, \dots, y'_N)'$.

In the second step, the weighting matrix W_2^L is constructed by using the instrument matrix, which is weighted by the fitted residuals from the LEV₁ estimator, \hat{u}_i 's, as follows

$$W_2^L = \left(\frac{1}{N} \sum_{i=1}^N Z_i^L \hat{u}_i \hat{u}_i' Z_i^{L'} \right)^{-1}$$

and the two-step LEV GMM estimator (LEV₂) for φ is computed as

$$\hat{\varphi}_{LEV_2} = (y'_{-1} Z^L W_2^L Z^{L'} y_{-1})^{-1} y'_{-1} Z^L W_2^L Z^{L'} y$$

2.1.3. System GMM estimator

A System GMM estimator (SYS) introduced by [19] combines the moment conditions of the DIF and LEV approaches to handle weak instrument problem and enhance the efficiency of the estimator. The model incorporating all equations both in first differences and in levels can be reformulated as a system of equations as follows.

$$\begin{pmatrix} \Delta y_i \\ y_i \end{pmatrix} = \varphi \begin{pmatrix} \Delta y_{i-1} \\ y_{i-1} \end{pmatrix} + \begin{pmatrix} \Delta u_i \\ u_i \end{pmatrix}$$

To obtain the SYS estimator, a full set of $m_S = (1/2)(T+1)(T-2)$ moment conditions and a $2(T-2) \times m_S$ block diagonal matrix, Z_i^S , are described by the following equations, respectively.

$$E(Z_i^S u_i^S) = 0 \text{ where } u_i^S = (\Delta u_i', u_i')'$$

$$Z_i^S = \begin{pmatrix} Z_i^D & 0 \\ 0 & Z_i^L \end{pmatrix}$$

Using the matrix of instruments, $Z^S = (Z_1^S, \dots, Z_N^S)'$, the one-step SYS estimator (SYS₁) of φ is obtained as

$$\hat{\varphi}_{SYS_1} = (y^{S'} Z^S W^S Z^{S'} y_{-1}^S)^{-1} y^{S'} Z^S W^S Z^{S'} y^S$$

where $y^S = [(\Delta y'_1, y'_1), \dots, (\Delta y'_N, y'_N)]'$, $y_{-1}^S = [(\Delta y'_{1,-1}, y'_{1,-1}), \dots, (\Delta y'_{N,-1}, y'_{N,-1})]'$, and $W^S = \left(\frac{1}{N} \sum_{i=1}^N Z_i^S G Z_i^{S'} \right)^{-1}$ with $G = \begin{pmatrix} D & 0 \\ 0 & I_{T-2} \end{pmatrix}$.

In the second step, using the residuals (\hat{u}_i^S) from SYS₁, a weighting matrix, W_2^S , and the two-step SYS estimator (SYS₂) are calculated respectively as follows

$$W_2^S = \left(\frac{1}{N} \sum_{i=1}^N Z_i^S \hat{u}_i^S \hat{u}_i^{S'} Z_i^{S'} \right)^{-1}$$

$$\hat{\varphi}_{SYS_2} = (y^{S'} Z^S W_2^S Z^{S'} y_{-1}^S)^{-1} y^{S'} Z^S W_2^S Z^{S'} y^S$$

Although the SYS estimator yields more efficient estimates compared to the LEV estimator, the bias of SYS estimator significantly increases in the presence of high persistence level of the autoregressive parameter and/or when the ratio of the variance of individual effects to that of the error term deviates from unity (cf. [2]).

In this study, we consider an alternative class of robust estimators, i.e., M-estimators, since the class of GMM estimators results in distorted parameter estimates in the presence of a high persistence level of the autoregressive parameter and/or when the ratio of the variance of the individual effects to the error term deviates from one. Next, we describe M-estimation approach weighted by power function to estimate first-order autoregressive panel model.

2.2. A Weighted-M Estimation Using Power Function for Dispersion Parameter

The main objective of this paper is to improve efficiency using the robust M-estimation approach for processes where the weak instrument problem arises. M-estimation approaches involve the minimization of a loss function that changes slowly in the presence of abnormal residuals. The novelty of our approach relies on obtaining weighted M-estimation using a power function for the scale parameter and choosing the value of the tuning parameter in the loss function by minimizing the variance of the estimators, as in [27, 28, 29].

The power function proposed by [32] for the scale parameter, σ , can be expressed in the first-order autoregressive panel data model as follows:

$$\sigma = \tau |\varphi y_{-1}|^\gamma$$

where τ is an unknown dispersion parameter with an unknown parameter vector γ , and $y_{-1} = (y'_{1,-1}, \dots, y'_{N,-1})'$. For the first-order autoregressive panel data models, the loss function, $\rho(\cdot)$ is defined as

$$\sum_{i=1}^N \sum_{t=2}^T \rho \left(\frac{y_{it} - \varphi y_{i,t-1} - \alpha_i}{\hat{\sigma}} \right)$$

where $\hat{\sigma}$ denotes an estimated value of scale parameter. To obtain the robust M-estimates of the autoregressive parameter, φ , and the individual-specific effects, α_i 's, the loss function, which varies slowly in the presence of outliers, is minimized. One of the most commonly employed loss function is the Huber's loss function described as follows

$$\rho(u) = \begin{cases} \frac{u^2}{2} & \text{if } |u| \leq c \\ c|u| - \frac{c^2}{2} & \text{if } |u| > c \end{cases} \quad (4)$$

where c is a tuning constant that regulates the level of robustness and is chosen within a range of values of 0 to 3. The default value of c in the R package *rlm* is set to 1.345 to achieve 95% asymptotic relative efficiency under the assumption of a normally distributed data.

Differentiating the loss function defined in Equation (4) results in the following estimating equations:

$$U(\varphi; \alpha_i) = \sum_{i=1}^N \sum_{t=2}^T \left(\frac{y_{i,t-1}}{\hat{\sigma}} \right) \psi \left(\frac{y_{it} - \varphi y_{i,t-1} - \alpha_i}{\hat{\sigma}} \right) = 0$$

$$U_\alpha(\varphi; \alpha_i) = \sum_{i=1}^N \sum_{t=2}^T \left(\frac{1}{\hat{\sigma}} \right) \psi \left(\frac{y_{it} - \varphi y_{i,t-1} - \alpha_i}{\hat{\sigma}} \right) = 0$$

where $\psi(u) = \min(c, \max(u, -c))$ denotes the sub-gradient function of the Huber's loss function.

Let us consider the idiosyncratic error terms in the model, $\varepsilon_{it} = y_{it} - \varphi y_{i,t-1} - \alpha_i$. Then, the solutions of this estimating functions are obtained by rewriting $U(\varphi; \alpha_i)$ and $U_\alpha(\varphi; \alpha_i)$ as in the form of the weighted score functions as follows:

$$U(\varphi; \alpha_i) = \sum_{i=1}^N \sum_{t=2}^T y_{i,t-1} W_{it} \bar{\varepsilon}_{it} = 0$$

$$U_\alpha(\varphi; \alpha_i) = \sum_{i=1}^N \sum_{t=2}^T W_{it} \bar{\varepsilon}_{it} = 0$$

where $\bar{\varepsilon}_{it} = \frac{(y_{it} - \varphi y_{i,t-1} - \alpha_i)}{\hat{\sigma}}$ and $W_{it} = \frac{\psi(\bar{\varepsilon}_{it})}{\bar{\varepsilon}_{it}}$

respectively denote the Pearson residuals and weights. Then, by solving the weighted score equations, the robust estimators of φ , $\hat{\varphi}_{POWER}$, and α_i , $\hat{\alpha}_i^{POWER}$, can be calculated as follows, respectively.

$$\hat{\varphi}_{POWER} = \left(\sum_{i=1}^N \sum_{t=2}^T y'_{i,t-1} W_{it} y_{i,t-1} \right)^{-1} \left(\sum_{i=1}^N \sum_{t=2}^T y'_{i,t-1} W_{it} y_{it} \right) \quad (5)$$

$$\hat{\alpha}_i^{POWER} = \frac{\sum_{t=2}^T (y_{it} - \hat{\varphi}_{POWER} y_{i,t-1})}{T-1} \quad (6)$$

Here, an iterative method is required since W_{it} depends on φ , α_i , and σ . This method is based on the pseudolikelihood approach and entails fixing the

model parameter φ and variance parameters (τ and γ) alternatively (cf. [27]). To estimate γ , we use a robust estimator with a breakdown point of %50 defined as

$$\sum_{i=1}^N \sum_{t=2}^T \chi \left(\frac{y_{it} - \hat{\varphi} y_{i,t-1} - \hat{\alpha}_i}{\hat{\tau} |\hat{\varphi} y_{i,t-1}|^\gamma} \right) \frac{(|\hat{\varphi} y_{i,t-1}|^\gamma)'}{|\hat{\varphi} y_{i,t-1}|^\gamma} = 0 \quad (7)$$

where $\chi(u) = \min\left(\frac{u^2}{1.0412^2}, 1\right) - 0.5$ is a bounded function developed by [33, 34]. Also, to estimate the dispersion parameter τ , we employ the MAD estimator expressed as

$$\hat{\tau} = \text{Median} \left\{ \frac{|y_{it} - \hat{\varphi} y_{i,t-1} - \hat{\alpha}_i|}{|\hat{\varphi} y_{i,t-1}|^\gamma} \right\} / 0.6745 \quad (8)$$

In robust approaches, it is crucial to appropriately select the tuning parameter, c , related to the loss function, as it controls the level of robustness in the estimation. In traditional robust methods, it is important to pre-specify the tuning constant c in any chosen loss function based on the desired level of robustness. When the errors follow a normal distribution, the optimal value for this parameter is $+\infty$. However, in heavy-tailed distributions, then c should be selected as a small positive value. The selection of the tuning parameter requires careful consideration as robustness often entails a sacrifice in efficiency. The value of the tuning parameter c should be selected based on the potential percentage of outliers in the data or the data distribution to maximize the asymptotic efficiency of the estimators. This is because the primary goal is to enhance the efficiency of the estimators while maintaining robustness.

In this study, following the approach outlined by [27, 29], the tuning parameter that minimizes the variance of the model parameters is referred to as the optimal one. Thus, it is suggested to iteratively estimate the tuning parameter c from the data for values ranging between 0 and 3 in Huber's loss function and select the value that minimizes the total estimated variances of the estimators in the model.

The algorithm below provides a summary of the entire estimating process.

1. Obtain initial M-estimates $\hat{\varphi}_0$ and $\hat{\alpha}_i^0 = \frac{\sum_{t=2}^T (y_{it} - \hat{\varphi}_0 y_{i,t-1})}{T-1}$ for $i = 1, \dots, N$ using the default value of c under the assumption of a constant variance $|\varphi y_{i,t-1}|^\gamma = 1$.
2. Calculate the residuals of this model by fixing $\hat{\varphi} = \hat{\varphi}_0$ and $\hat{\alpha}_i = \hat{\alpha}_i^0$ for $i = 1, \dots, N$ and obtain the robust estimates of the variance parameters, $(\hat{\gamma}, \hat{\tau})$, using Equations (7)-(8).
3. Update $\hat{\varphi}$ and $\hat{\alpha}_i$'s with the Equations (5)-(6), i.e., using weighted M-estimators, by setting the robust estimates of variance parameters, $(\hat{\gamma}, \hat{\tau})$, obtained in the previous step.

4. Repeat steps 2-3 till the optimal value of the tuning parameter, c , is obtained from a finite set of values ranging from 0 to 3. The best tuning parameter yields the smallest sum of the estimated variance of the model parameters.
5. Finally calculate the robust estimates of φ and α_i 's using the best tuning constant.

III. NUMERICAL RESULTS

In this section, we carry out a comprehensive simulation study to investigate the performance of our proposed estimator (referred to as "POWER" in the graphs), and we compare our results with the one- and two-step GMM estimators: DIF, LEV and SYS. Note that the subscripts used in the abbreviations of the GMM estimators indicate the number of steps applied in the GMM estimation. For example, DIF₁ is used to show the performance of the one-step DIF estimator, while DIF₂ is associated with the two-step DIF estimator. To assess the finite-sample properties of each estimator, we consider three different scenarios: (i) different sample sizes, (ii) varying levels of individual heterogeneity, and (iii) varying levels of the persistence. For the data generation process (DGP), dynamic panel autoregressive model of order one given in Equation (1) is considered.

To assess the performance of the proposed and conventional GMM estimators using $S = 1000$ simulations, we calculate the bias and the root mean squared error (RMSE) of the estimator of autoregressive parameter, $\hat{\varphi}$, as follows:

$$\text{Bias} = \frac{1}{S} \sum_{s=1}^S (\hat{\varphi}^s - \varphi)$$

$$\text{RMSE} = \sqrt{\frac{1}{S} \sum_{s=1}^S \text{Var}(\hat{\varphi}^s) + \text{Bias}^2(\hat{\varphi}^s)}$$

where $\hat{\varphi}^s, s = 1, \dots, S$ are the estimates of the autoregressive parameter, φ , obtained from the simulated samples. Throughout the experiments, the initial observations are generated considering the initial conditions developed by [30] as $y_{i1} = \frac{\alpha_i}{1-\varphi} + \mu_{i1}$ where $\mu_{i1} \sim N(0,1)$, independent of α_i and ε_{it} .

3.1. Sample Sizes

To investigate the performance of the estimators under different sample sizes, the individual effects and the errors are generated from a normal distribution with mean zero and equal variances of $\sigma_\alpha^2 = 0.25$ and $\sigma_\varepsilon^2 = 0.25$, respectively. This implies that the ratio of the variance of the individual effect to that of the error term, $r = \frac{\sigma_\alpha^2}{\sigma_\varepsilon^2}$, kept fixed at $r = 1$. Thus, we exclude the examination of individual heterogeneity and focus

solely on investigating the impact of sample sizes on the performance of the estimators in this subsection. To this end, we consider the increasing values of the cross-sectional dimension $N = 25, 50, 100, 200, 500$, with a fixed time period $T = 8$, and the increasing values of the time dimension $T = 4, 8, 12, 18, 24$ while keeping the number of cross-sectional units fixed at $N = 100$. The RMSE and bias of the estimators are compared when a moderate level of persistence is present with an autoregressive parameter $\varphi = 0.5$. Figure 1 illustrates the calculated RMSE and bias of both the conventional GMM estimators and the proposed weighted M-estimator with a power function as the cross-sectional dimension increases. Figure 2 displays the simulation results for the increasing time periods. In general, the RMSEs and bias of all the methods tend to decrease with increasing N and T , as expected. These figures demonstrate that the RMSEs calculated for the one- and two-step LEV estimators are largest relative to the other estimators, while the one- and two-step DIF estimators have the largest negative bias values in all cases. Also, although our proposed weighted M-estimator with a power function produces slightly increasing RMSEs and bias with increasing T , our proposed method (POWER) yields the lowest RMSE and bias values, particularly when considering small values of T . It is conspicuous from these figures that our proposed weighted M-estimator outperforms all the conventional GMM estimators in all situations.

3.2. Levels of Individual Heterogeneity

In this subsection, we focus on the impacts of different values of the variance ratio, $r = \frac{\sigma_\alpha^2}{\sigma_\varepsilon^2}$, which indicates the level of individual heterogeneity, on the performance of the estimators because conventional GMM estimators result in obtaining distorted estimates of the parameters as heterogeneity level increases. Thus, we investigate the robustness against increasing heterogeneity for our proposed estimator, following the experimental design used by [1]. To consider the scenarios where the variance ratio is $r < 1, r = 1$, and $r > 1$, we select four different pairs of σ_α^2 and σ_ε^2 as $(\sigma_\alpha^2, \sigma_\varepsilon^2) = (0.25, 0.50), (0.25, 0.25), (2.5, 0.5), (2.5, 0.25)$ when $(N, T) = (100, 8)$. We report the results for the selected values of autoregressive parameter $\varphi = 0.1, 0.3, 0.5, 0.7, 0.9$, considering a range from weak persistence to strong persistence.

The plots of the calculated RMSEs of the estimators versus the variance ratio, $r = 0.5, 1, 5, 10$, are given in Figure 3, whereas Figure 4 presents the bias values of the estimators versus r . Figure 3 illustrates that under weak and moderate levels of persistence, all the estimators, including our proposed estimator, tend to exhibit increasing RMSEs as the variance ratio r rises.

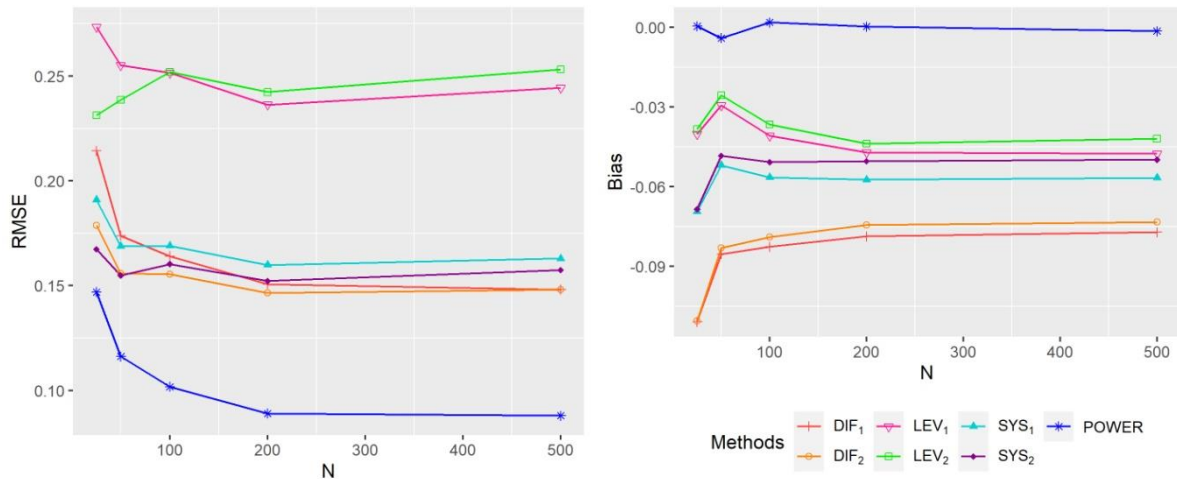


Figure 1: RMSE and Bias of conventional GMM estimators and the proposed weighted M-estimator with a power function when $N = 25, 50, 100, 200, 500$ with $T = 8, r = 1,$ and $\varphi = 0.5$.

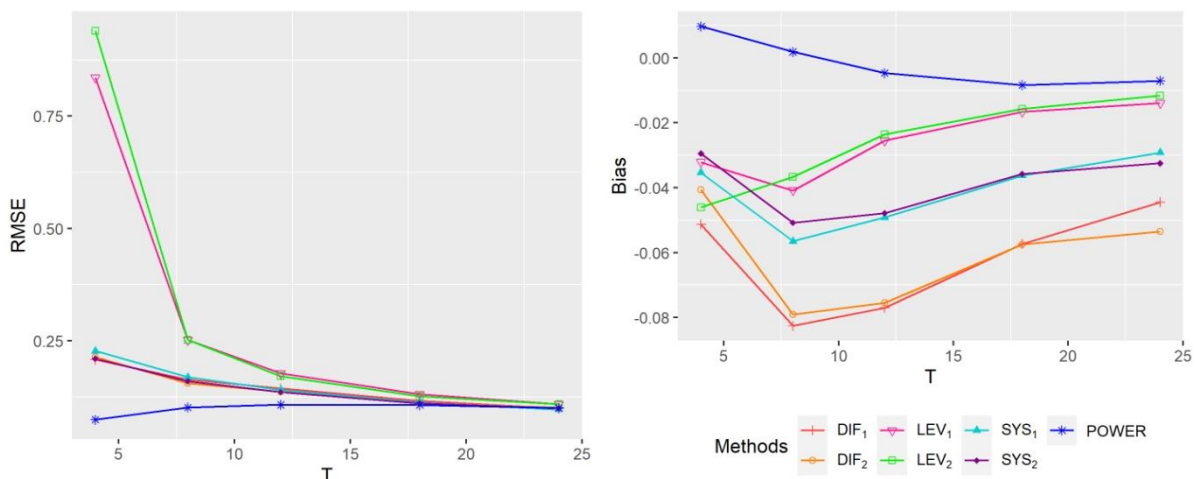


Figure 2: RMSE and Bias of conventional GMM estimators and the proposed weighted M-estimator with a power function when $T = 4, 8, 12, 18, 24$ with $N = 100, r = 1$ and $\varphi = 0.5$.

For weak and moderate levels of persistence, the LEV-GMM estimators have the highest RMSEs among the other estimators. On the other hand, when $\varphi = 0.9$, indicating strong persistence, all the methods except the DIF-GMM methods produce decreasing values of RMSEs as r increases, with the DIF-GMM estimators providing the largest RMSEs in that scenario. Our proposed weighted-M estimator with a power function, yields the smallest RMSE values as r increases for all values of the autoregressive parameter considered in this study. Figure 4 demonstrates that the DIF estimators have the largest bias values among all estimators and exhibit increasing bias as the level of heterogeneity increases. This indicates that the DIF estimators are more efficient than the LEV estimators at the weak and moderate levels of persistence when taking into account both performance metrics, i.e., RMSE and bias. The bias values obtained by the proposed POWER estimator are the smaller than those of the class of conventional GMM estimators for the weak and moderate level of persistence, even as the

level of heterogeneity increases. At $\varphi = 0.9$, both LEV-GMM and our proposed estimator yield competitive results when bias values are evaluated. In general, our results demonstrate that that proposed POWER estimator outperforms conventional GMM estimators as the heterogeneity level increases across all persistence levels.

3.3. Level of Persistence

Following the same experimental design used in the previous subsection, we examine the finite sample performance of the estimators against increasing persistence level. To this end, we plot the RMSEs and bias values of the estimators versus the various levels of persistence in Figures 5-6. In Figure 5, both the proposed and conventional GMM estimators result in decreasing RMSE values as the persistence level increases across all heterogeneity levels considered. Under weak heterogeneity ($r = 0.5$ and $r = 1$), the one- and two-step LEV-GMM estimators have higher RMSEs than the other GMM estimators.

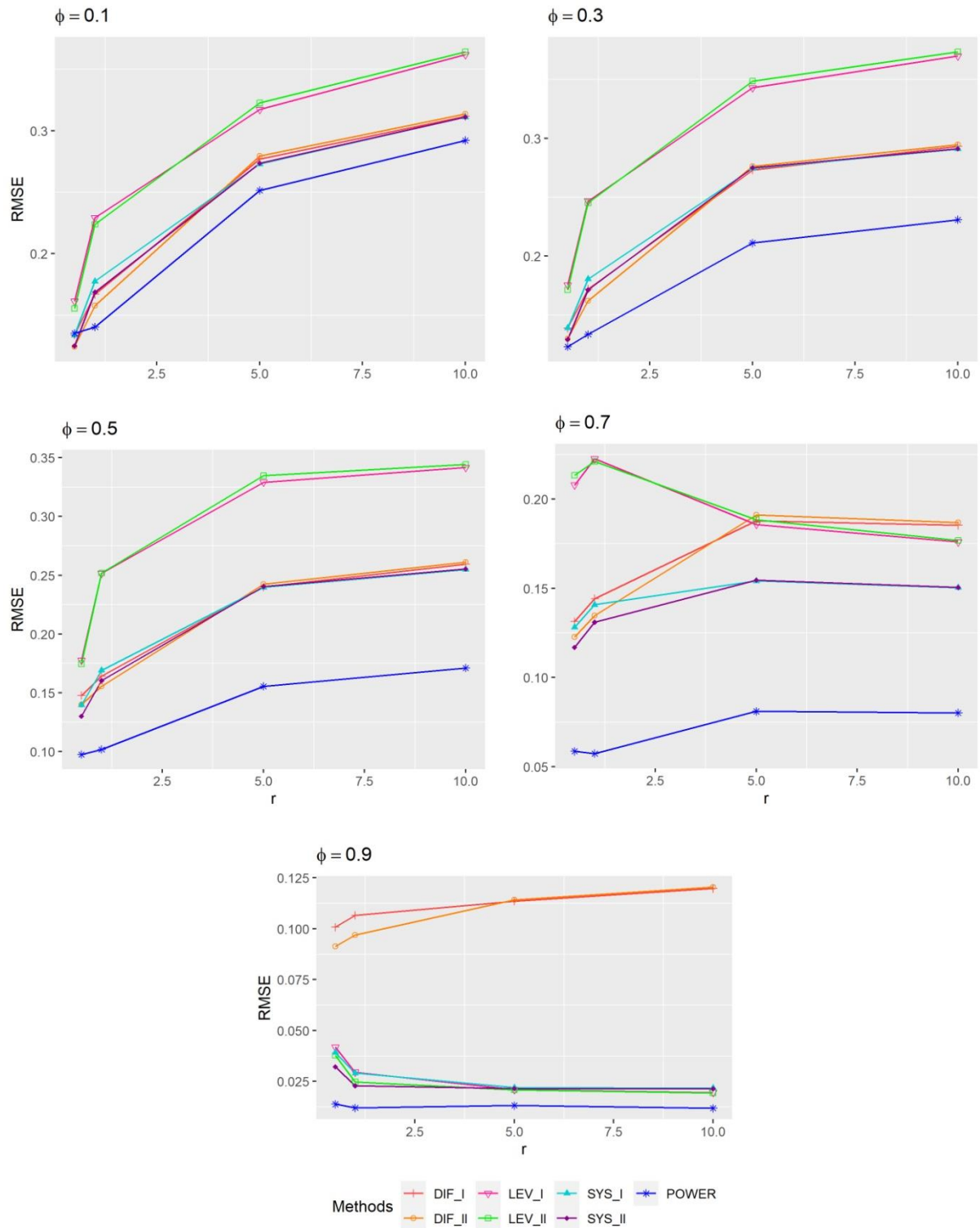


Figure 3: RMSE of conventional GMM estimators and the proposed weighted M-estimator with a power function for different values of $\phi = 0.1, 0.3, 0.5, 0.7, 0.9$ when $(N, T) = (100, 8)$ and $r = 0.5, 1, 5, 10$.

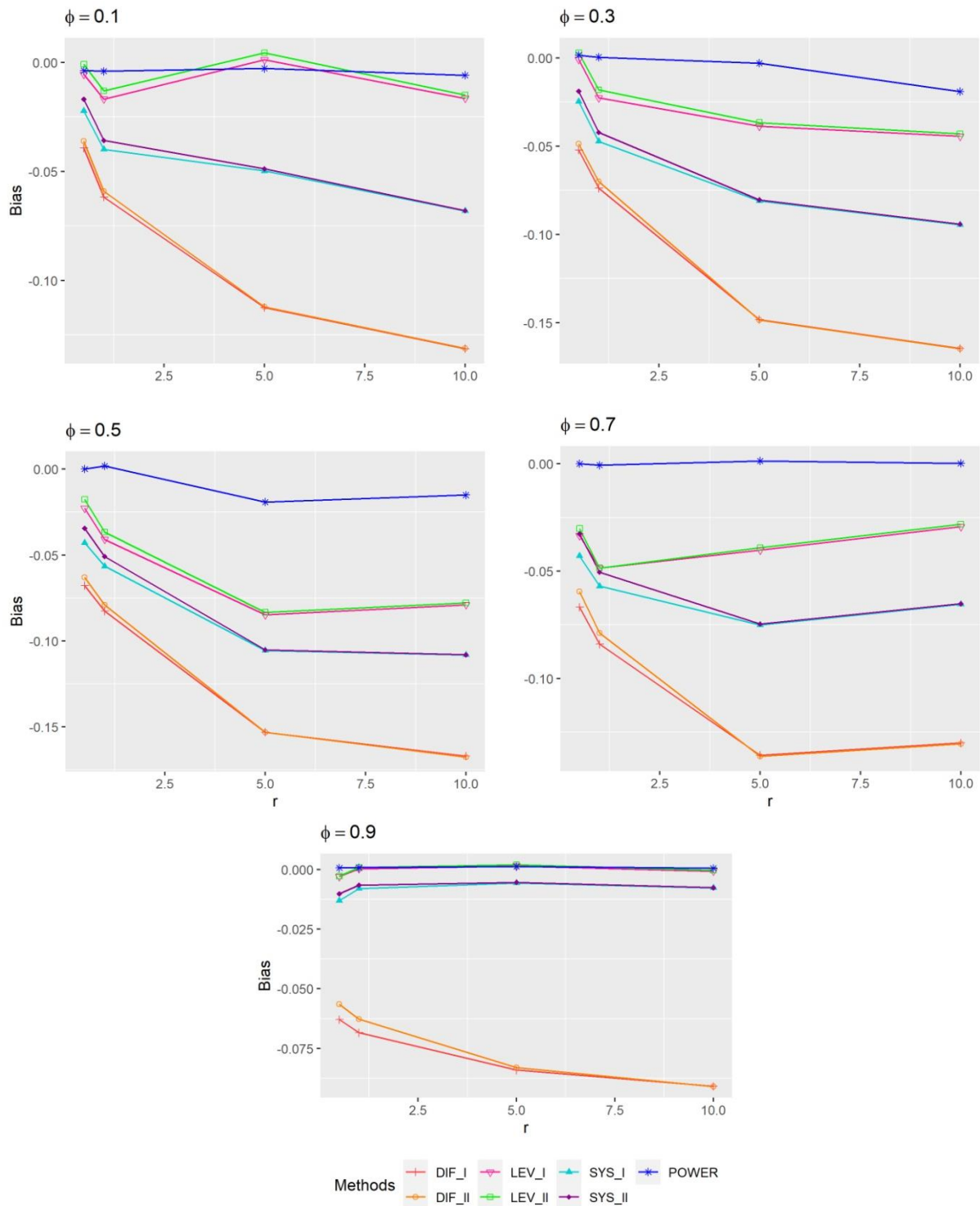


Figure 4: Bias of conventional GMM estimators and the proposed weighted M-estimator with a power function for different values of $\phi = 0.1, 0.3, 0.5, 0.7, 0.9$ when $(N, T) = (100, 8)$ and $r = 0.5, 1, 5, 10$.

However, their performance tends to converge closely with that of the DIF-GMM and SYS-GMM estimators as the persistence level increases ($\phi = 0.9$). Furthermore, as the heterogeneity level increases ($r = 5$ and $r = 10$), LEV-GMM estimators yield competitive results to SYS-GMM estimators for moderate and strong levels of persistence ($\phi = 0.7$ and $\phi = 0.9$). The POWER estimator, proposed in

this study, demonstrates superior performance overall, except for the case $r = 0.5$ and $\phi = 0.1$. This is despite the fact that the two-step SYS estimator achieves the lowest RMSEs among all GMM estimators, particularly with higher values of the autoregressive parameter. Based on Figure 6, it can be observed that the largest negative bias values are obtained by the one-step DIF-GMM estimator.

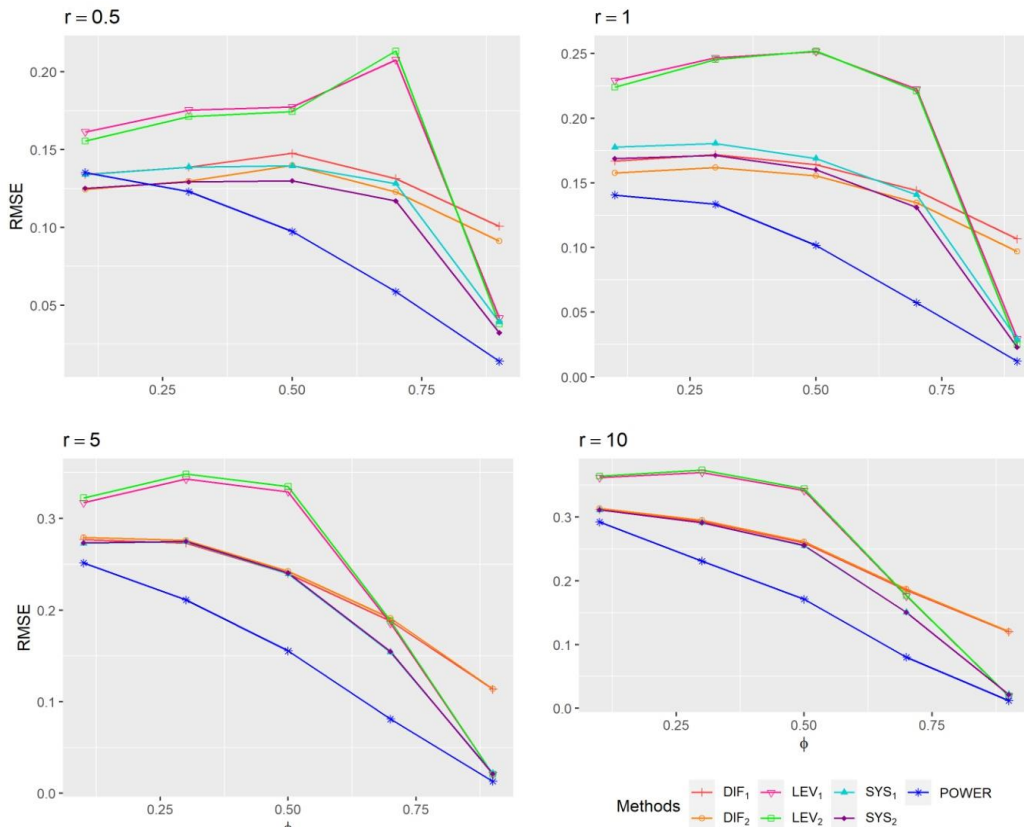


Figure 5: RMSE of conventional GMM estimators and the proposed weighted M-estimator with a power function for different levels of variance ratio $r = 0.5, 1, 5, 10$ when $(N, T) = (100, 8)$.

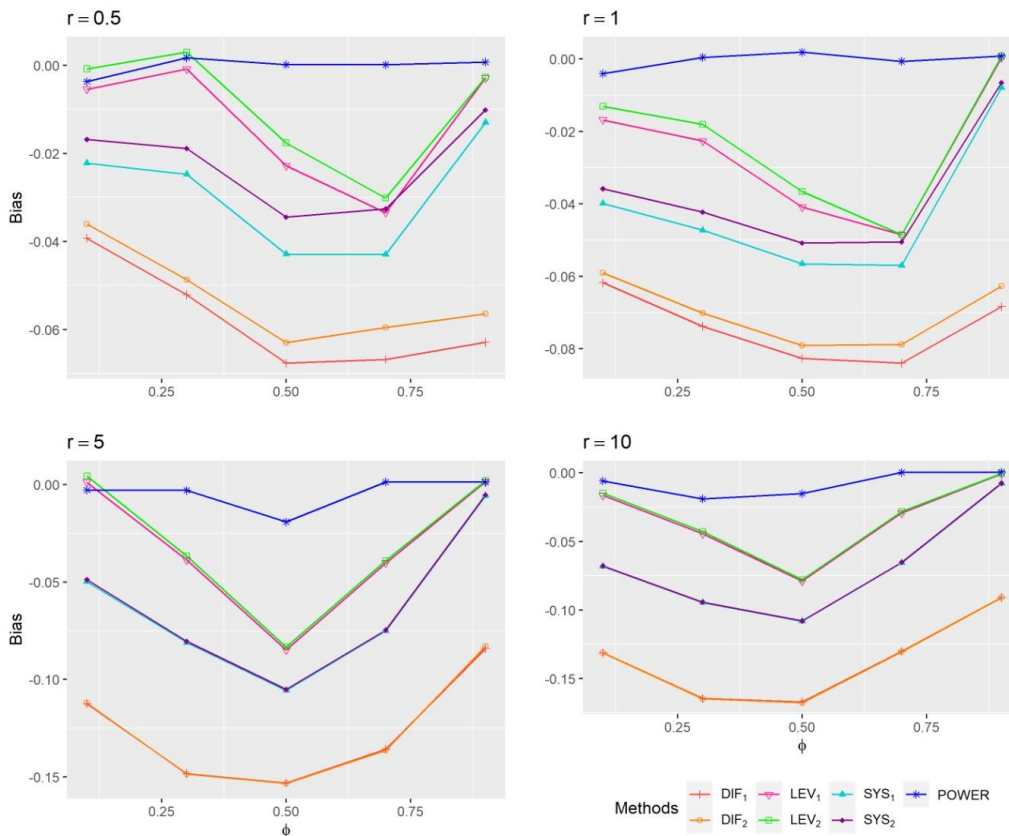


Figure 6: Bias of conventional GMM estimators and the proposed weighted M-estimator with a power function for different levels of variance ratio $r = 0.5, 1, 5, 10$ when $(N, T) = (100, 8)$.

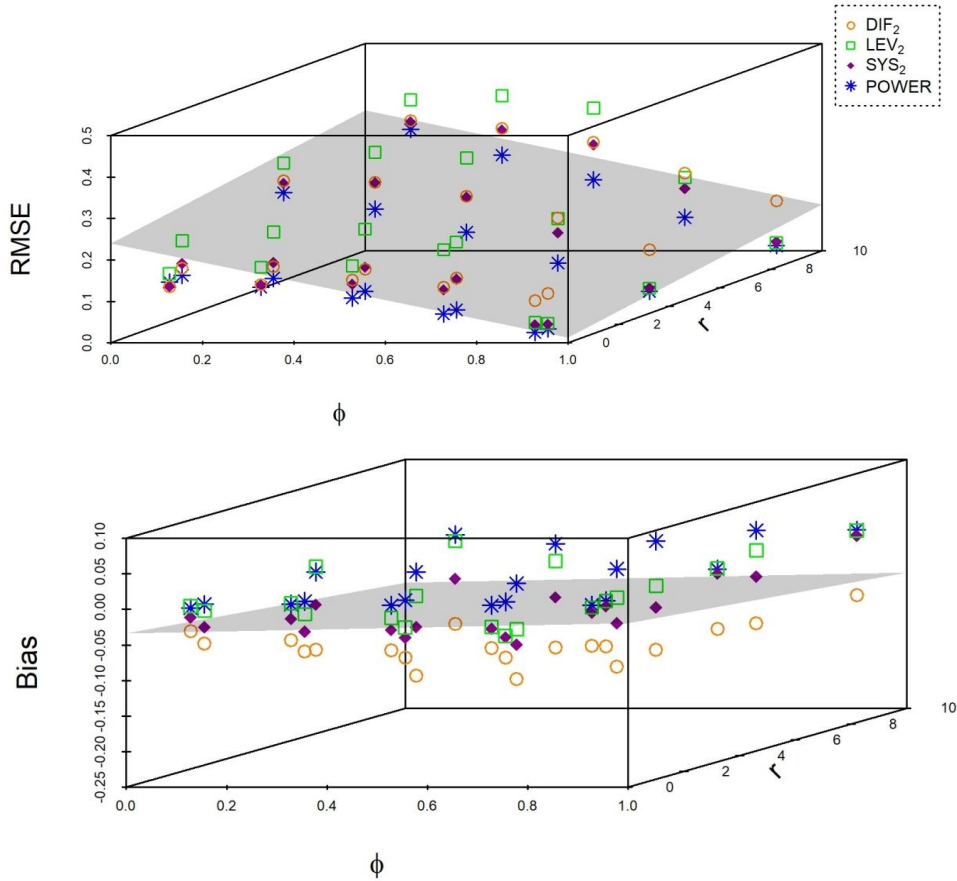


Figure 7: 3D-scatterplots of RMSE and Bias for conventional GMM estimators and the proposed weighted M-estimator with a power function across $\varphi = 0.1, 0.3, 0.5, 0.7, 0.9$ and $r = 0.5, 1, 5, 10$ when $(N, T) = (100, 8)$.

For all GMM estimators, the absolute bias values increase until a moderate level of persistence. However, they tend to produce decreasing bias in absolute value when $\varphi = 0.7$ and $\varphi = 0.9$. Indeed, the autoregressive parameter value of 0.5 acts as a critical threshold for all methods. In all scenarios, our proposed POWER estimator demonstrates improved performance over conventional GMM estimators based on both performance metrics, except for cases of weak persistence and weak heterogeneity level ($\varphi = 0.1$ and $r = 0.5$).

The simulation results for different levels of heterogeneity and persistence are depicted by a three dimensional scatter plot (3D-scatter plot) in Figure 7. Figure 7 presents that the RMSE and bias of the estimators plotted as functions of the autoregressive parameter φ and variance ratio r . An increase in the variance ratio r have a more significant impact on RMSE compared to an increase in the autoregressive parameter φ . The larger RMSE values are obtained for higher values of r when the level of persistence is weak and moderate. The bias values of the GMM estimators are often negative, whereas the proposed estimator yields bias values that are closest to zero. In summary, our proposed POWER estimator is

considerably less affected by increasing level of heterogeneity and/or persistence.

IV. CASE STUDY

In this section, we employ the proposed robust procedure to examine the monthly hydro-electric power output data (100 million kwh), available from the National Bureau of Statistics of China at <https://data.stats.gov.cn/english/easyquery.htm>. This dataset includes 220 observations ($N = 22, T = 10$), representing a cross section of 22 regions across China from March 2023 to December 2023. Table 1 presents the regions of China included in this study. Figure 8 displays the marginal distributions of the hydro-electric power output across 22 regions of China throughout different months. This figure suggests that there is heterogeneity in the hydro-electric power output among different regions of China, with some regions including outliers in their hydro-electric power output.

A first-order panel autoregressive model, given in Equation (1), is fitted to the data. For this data, y_{it} represents the hydro-electric power output, and the indices $i = 1, \dots, 22$ and $t = 1, \dots, 10$ denote the regions of China and months, respectively.

Table 1: List of regions of China

Beijing	Jilin	Fujian	Hainan	Qinghai
Hebei	Heilongjiang	Jiangxi	Chongqing	Ningxia
Shanxi	Jiangsu	Shandong	Tibet	
Inner Mongolia	Zhejiang	Henan	Shaanxi	
Liaoning	Anhui	Guangxi	Gansu	

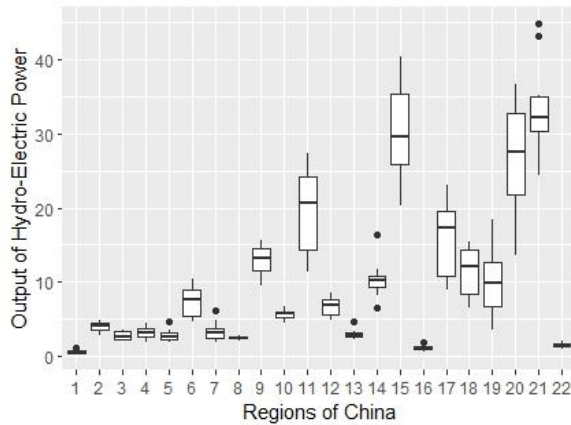


Figure 8: Boxplots of the hydro-electric power output across 22 regions of China

To assess the predictive performance of all estimators, we calculate the trimmed mean prediction errors due to the lack of prior knowledge about outliers in empirical data analysis. To this end, a specified percentage (e.g., $p\%$) of observations is trimmed by considering those with the highest squared prediction errors, followed by computing the mean prediction error for the remaining portion of the dataset. The mean prediction error (MPE), excluding outliers, is calculated as defined below

$$MPE = \frac{1}{22 \times 10 - m} \sum_{i=1}^{22} \sum_{t=1}^{10} (1 - I_{it})(y_{it} - \hat{y}_{it})^2$$

where y_{it} and \hat{y}_{it} respectively denote the observed and predicted values of the hydro-electric power output, and $m = \frac{p}{100} \times 22 \times 10$ represents the number of outliers in the dataset. Here, I_{it} represents an indicator variable that takes the value 1 if y_{it} is an outlying observation, and 0 otherwise. The estimates of the autoregressive parameter and trimmed MPEs are reported in Table 2.

Table 2: The estimates of autoregressive parameter φ obtained by the conventional GMM estimators and the proposed weighted M-estimator with a power function, along with MPEs for various trimming percentages, p .

Estimator	Full Data		MPE $p = 5\%$	MPE $p = 10\%$	MPE $p = 15\%$	MPE $p = 20\%$
	$\hat{\varphi}$	MPE $p = 0\%$				
POWER	0.9946	10.7786	5.2866	2.9803	1.8708	1.2068
DIF₁	0.4749	61.9876	38.7937	26.0449	17.6564	12.5514
DIF₂	0.4738	62.2326	38.9658	26.1661	17.7376	12.6114
LEV₁	0.9535	10.8289	5.5147	3.0137	1.9170	1.2982
LEV₂	0.9535	10.8296	5.5145	3.0136	1.9169	1.2982
SYS₁	0.9357	11.0653	5.7130	3.1685	1.9643	1.3014
SYS₂	0.9356	11.0666	5.7141	3.1694	1.9645	1.3015

We consider trimming percentages as $p = 5\%, 10\%, 15\%$, and 20% , with the value zero representing the full dataset. From Table 2, it is evident that the estimated value of the autoregressive parameter to be greater than 0.9, indicating strong persistence within the dataset. The strong persistence is evident since the DIF-GMM estimators result in the largest MPEs with estimated values of the

parameter is considerably high, approaching unity, for the proposed POWER estimator. Additionally, the GMM estimators, with the exception of the DIF-GMM estimators, have estimated the autoregressive parameter near the moderate level of 0.5. As the trimming percentages increase, all estimators exhibit improved predictive performance with decreasing MPEs. The proposed POWER

estimator produces the lowest MPE among all estimators considered across all trimming percentages. Although the smallest MPEs are provided by the LEV-GMM estimators among the GMM estimators, our proposed POWER estimator yields better predictions, ranging from 0.47% to 7.57% across various trimming percentages, compared to the LEV-GMM estimators. Moreover, the proposed estimator produces more accurate predictions, ranging from 2.65% to 8.08, compared to SYS-GMM estimators. All empirical findings clearly indicate that the proposed POWER estimator is robust in the presence of strong persistence and heterogeneity, demonstrating substantially improved predictive ability compared to conventional GMM methods.

V. CONCLUSION

In this study, we propose a robust weighted M-estimation approach for first-order dynamic panel autoregressive model, which considerably enhances the precision and robustness of parameter estimates under challenging conditions of high persistence and individual heterogeneity. The proposed estimator is constructed by weighting the M-estimator with a power function for the scale parameter used in the Huber's loss function, with optimization of the tuning parameter by minimizing the variance of the model parameters. Through extensive simulations and empirical data analysis, we have demonstrated that our method outperforms traditional GMM estimators, particularly in the presence of weak instruments and near-unity autoregressive parameters. This robust estimator provides a valuable tool for researchers across various fields, including econometrics, biology, and engineering, facilitating more accurate analysis of dynamic panel data.

REFERENCES

- [1] Youssef, A. H. & Abonazel, M. R. (2017). Alternative GMM estimators for first-order autoregressive panel model: An improving efficiency approach. *Communications in Statistics-Simulation and Computation*, 46(4), 3112-3128.
- [2] Bun, M. J. G. & Kiviet, J. F. (2006). The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models. *Journal of Econometrics*, 132(2), 409-444.
- [3] Neyman, J. & Scott, E. L. (1948). Consistent estimates based on partially consistent observations. *Econometrica*, 16(1), 1-32.
- [4] Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6), 1417-1426.
- [5] Lancaster, T. (2002). Orthogonal parameters and panel data. *Review of Economic Studies*, 69(3), 647-666.
- [6] Hsiao, C., Pesaran, M. H. & Tahmiscioglu, A. K. (2002). Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. *Journal of Econometrics*, 109(1), 107-150.
- [7] Binder, M., Hsiao, C. & Pesaran, M. H. (2005). Estimation and inference in short panel vector autoregressions with unit roots and cointegration. *Econometric Theory*, 21(4), 795-837.
- [8] Hayakawa, K. & Pesaran, M. H. (2015). Robust standard errors in transformed likelihood estimation of dynamic panel data models with cross-sectional heteroskedasticity. *Journal of Econometrics*, 188(1), 111-134.
- [9] Bun M. J. G., Carree, M. A. & Juodis, A. (2017). On Maximum likelihood estimation of dynamic panel data models. *Oxford Bulletin of Economics and Statistics*, 79(4), 463-494.
- [10] Alvarez, J. & Arellano, M. (2022). Robust likelihood estimation of dynamic panel data models. *Journal of Econometrics*, 226(1), 21-61.
- [11] Hansen, L. P. (2010). Generalized method of moments estimation. In: *Macroeconometrics and Time Series Analysis*, Durlauf, S. N., Blume, L. E. (eds), Palgrave Macmillan, London, p. 105-118.
- [12] Cochrane, J. H. (2001). *Asset pricing*. Princeton University Press, Princeton, New Jersey.
- [13] Arellano, M. (2003). *Panel data econometrics*. Oxford University Press, New York.
- [14] Hall, A. R. (2005). *Generalized method of moments*. Oxford University Press, New York.
- [15] Singleton, K. J. (2006). *Empirical dynamic asset pricing: Model specification and econometric assessment*. Princeton University Press, Princeton, New Jersey.
- [16] Anderson, T. W. & Hsiao, C. (1982). Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, 18(1), 47-82.
- [17] Holtz-Eakin, D., Newey, W. & Rosen, H. S. (1988). Estimating vector autoregressions with panel data. *Econometrica*, 56(6), 1371-1395.
- [18] Arellano, M. & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58(2), 277-297.
- [19] Blundell, R. & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115-143.
- [20] Alonso-Borrego, C. & Arellano, M. (1999). Symmetrically normalized instrumental variable estimation using panel data. *Journal of Business & Economic Statistics*, 17(1), 36-49.
- [21] Han, C., Phillips, P. C. B. & Sul, D. (2014). X-differencing and dynamic panel model estimation. *Econometric Theory*, 30(1), 201-251.

- [22] Arellano, M. & Bover, O. (1995). Another look at the instrumental variable estimation of error components models. *Journal of Econometrics*, 68(1), 29-51.
- [23] Ahn, S. C. & Schmidt, P. (1995). Efficient estimation of models for dynamic panel data. *Journal of Econometrics*, 68(1), 5-27.
- [24] Bun, M. J. G. & Windmeijer, F. (2010). The weak instrument problem of the system GMM estimator in dynamic panel data models. *Econometrics Journal*, 13(1), 95-126.
- [25] Hayakawa, K. (2007). Small sample bias properties of the system GMM estimator in dynamic panel data models. *Economics Letters*, 95(1), 32-38.
- [26] Han, C. & Philips, P. (2010). GMM estimation for dynamic panels with fixed effects and strong instruments at unity. *Econometric Theory*, 26(1), 119-151.
- [27] Callens, A., Wang, Y. G., Fu, L. & Liqueet, B. (2021). Robust estimation procedure for autoregressive models with heterogeneity. *Environmental Modeling & Assessment*, 26, 313-323.
- [28] Wang, Y. G., Lin, X., Zhu, M. & Bai, Z. (2007). Robust estimation using the Huber function with a data-dependent tuning constant. *Journal of Computational and Graphical Statistics*, 16(2), 468-481.
- [29] Wang, N., Wang, Y. G., Hu, S., Hu, Z. H., Xu, J., Tang, H. & Jin, G. (2018). Robust regression with data-dependent regularization parameters and autoregressive temporal correlations. *Environmental Modeling & Assessment*, 23, 779-786.
- [30] Alvarez, J. & Arellano, M. (2003). The time series and cross-section asymptotics of dynamic panel data estimators. *Econometrica*, 71(4), 1121-1159.
- [31] Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, 126(1), 25-51.
- [32] Box, G. E. P. & Hill, W. J. (1974). Correcting inhomogeneity of variance with power transformation weighting. *Technometrics*, 16(3), 385-389.
- [33] Croux, C. (1994). Efficient high-breakdown M-estimators of scale. *Statistics and Probability Letters*, 19(5), 371-379.
- [34] Bianco, A., Boente, G. & Di Rienzo, J. (2000). Some results for robust GM-based estimators in heteroscedastic regression models. *Journal of Statistical Planning and Inference*, 89(1-2), 215-242.

Türkiye, Hindistan ve Brezilya'daki Petrol Kiralarının Otoregresif Kesirli Bütünleşik Hareketli Ortalama (ARFIMA) Modeli ile Belirlenmesi

Determination of Oil Rents in Turkey, India and Brazil with Autoregressive Fractionally Integrated Moving Average (ARFIMA) Model

Semanur SARIÇAM¹, Barış AŞIKGİL²

¹Marmara Üniversitesi, Fen Fakültesi, İstatistik Bölümü, İstanbul, Türkiye

²Mimar Sinan Güzel Sanatlar Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü, İstanbul, Türkiye

Öz

Petrol kiralari, üretim, ithalat ve tüketim vergileri üzerinden elde edilen gelir olduğu için enerji politikalarının bir parçası olarak değerlendirilir. Türkiye, Hindistan ve Brezilya gibi ülkelerin enerji politikaları, petrol ürünlerine yüksek vergiler uygulanmasından dolayı bu ülkelerin gelir elde etme yöntemlerinin bir parçasıdır. Petrol kiralariinin ekonomide önemli rolü göz önüne alındığında, bu çalışma, bu ülkelerin 1970-2016 yıllarını kapsayan petrol kira verilerini kullanarak uzun hafıza modellerinden biri olan otoregresif kesirli bütünleşik hareketli ortalama (ARFIMA) modeliyle incelemeyi amaçlamaktadır. İlk aşamada, Türkiye, Hindistan ve Brezilya'nın petrol kira serilerinin durağan olup olmadığını belirlemek için klasik birim kök testleri uygulanmıştır. Sonrasında, serilerin uzun hafızaya sahip olup olmadığını değerlendirmek amacıyla Hurst'un R/S istatistiği kullanılmıştır. İkinci aşamada ise p ve q değerleri 2'den küçük veya eşit olacak şekilde olası tüm otoregresif hareketli bütünleşik ortalama (ARIMA) ve ARFIMA modelleri test edilmiştir. En iyi model belirlenebilmesi için Akaike Bilgi Kriteri (AIC) ve Bayesci Bilgi Kriteri (BIC) iki temel bilgi kriterlerinden yararlanılmıştır. En uygun model belirlendikten sonra, bu model aracılığıyla geleceğe yönelik tahminler yapılmıştır.

Anahtar kelimeler: Petrol kiralari, Uzun hafıza modelleri, ARFIMA modeli, ARIMA modeli.

Abstract

Oil rents are considered as part of energy policies as they are revenues from taxes on production, imports and consumption. The energy policies of countries such as Turkey, India and Brazil are part of the revenue generation methods of these countries due to high taxes on petroleum products. Considering the important role of oil rents in the economy, this study aims to analyse the oil rent data of these countries covering the years 1970-2016 with the autoregressive fractionally integrated moving average (ARFIMA) model, which is one of the long memory models. In the first stage, classical unit root tests are applied to determine whether the oil rent series of Turkey, India and Brazil are stationary. Then, Hurst's R/S statistic is used to assess whether the series have long memory. In the second stage, all possible autoregressive moving average (ARIMA) and ARFIMA models are tested for p and q values less than or equal to 2. In order to determine the best model, two basic information criteria Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are used. After the most appropriate model is determined, forecasts are made by using this model.

Key words: Oil rents, Long memory models, ARFIMA model, ARIMA model.

1. GİRİŞ

Zamana bağlı olarak kaydedilen finansal ve makroekonomik serilerin modellenmesi ve geleceğe yönelik tahminler yapılabilmesi için ekonometriciler ve istatistikçiler genellikle ARIMA modellerini kullanırlar. ARIMA modelleri, bugünkü gözlemler kullanılarak yarının tahmin edilmesine olanak sağlayan kısa hafızalı modellerdir. Ancak, serinin daha eski gözlemlerinin bugünü etkilemesi sonucunda uzun hafızalı modeller geliştirilmiştir. Bu, serinin otokorelasyon fonksiyonunun hiperbolik olarak azalmasıyla karakterize edilir.

Mevsimsel olmayan Box-Jenkins modelleri genellikle ARIMA(p, d, q) olarak ifade edilir. Burada d parametresi, serinin durağan olup olmadığını belirler; eğer seri durağansa, d = 0 olarak tanımlanır. Çoğu zamana bağlı seriler durağan olmadığından, maksimum iki defaya kadar fark alma işlemi yapılır. Serinin birinci farkı durağansa d = 1, ikinci farkı durağansa d = 2 olur. Ancak, 0 ile 1 arasındaki kesirli değerlerin modele dahil edilmemesi bilgi kaybına yol açabilir. Bu kesirli değerlerin [-0.5, 0.5] aralığında olması ARFIMA(p, d, q) modelini oluşturur. Fark parametresi d, serinin uzun hafızalı mı yoksa kısa hafızalı mı olduğunu belirlemeye yardımcı olur.

Sorumlu Yazar: SEMANUR SARIÇAM, **Tel:** 05465660376, **E-posta:** semanur.saricam@marmara.edu.tr

Gönderilme: 28.06.2024, **Düzenleme:** 30.07.2024, **Kabul:** 03.08.2024

Finansal seriler gibi zaman serisi verileri üzerine uzun hafıza kavramının varlığını araştıran çok sayıda çalışma mevcuttur. Kutlar ve Turgut, ARFIMA modelleri ile tahmin ve öngörülebilirlik üzerine Türkiye'deki başlıca ekonomik serilere odaklanmışlardır [1]. Barışık ve Çevik, işsizlik verisi üzerine Geweke ve Porter-Hudak (GPH) ve modifiye edilmiş log periodogram (MLP) metodlarını kullanarak ARFIMA modellerini incelemişlerdir [2]. Ayrıca, 2001 kriz sonrası 2003-2007 yıllarını kapsayan Türk bankacılık sektörüne ait hisse senetlerini Çevik ve Erdoğan, güçlü hafıza modelleri ile analiz etmişlerdir [3]. Portföy optimizasyonu üzerine Pekkaya, ARFIMA ve kesirli bütünleşik genelleştirilmiş otoregresif koşullu değişen varyans (FIGARCH) modellerini kullanarak inceleme yapmıştır [4]. Çevik, İstanbul Menkul Kıymetler Borsası'nda (İMKB) işlem gören 10 sektör endeksini kullanarak uzun hafıza modellerinin etkin piyasa hipotezinde kullanılmasının önemini vurgulamıştır [5]. Karia, Bujang ve Ahmad, Malezya'daki ham hurma yağı fiyatlarını kullanarak tahmin hesaplamasında ARFIMA modelin daha iyi sonuçlar gösterdiğini kanıtlamışlardır [6]. Pekkaya ve Albayrak, İMKB-30 Endeks verileri üzerinde getiri öngörülerini için ARFIMA modelini ve varyans öngörülerini için de FIGARCH modelini kullanarak tahmin yapmışlardır [7]. Yılmaz, Brezilya, Rusya, Hindistan ve Çin'deki hisse senedi verilerini çalışmasına dahil ederek uzun hafıza davranışının sergilenip sergilenmediğini ve volatilitenin olup olmadığını araştırmak için bir inceleme yapmıştır [8]. Boateng, ve ark., Gana'daki tüketici fiyat endeksi (CPI) için uzun hafızanın sergilendiğini, yarı parametrik ve parametrik yöntemler kullanılarak ARFIMA modelinin en iyisi olduğunu keşfetmişlerdir [9]. Omeraka ve diğ., ARIMA ve ARFIMA modellerini kullanarak Nijerya'daki ticari bankaların likidite oran verisi üzerinde en iyi modelin ARFIMA model olduğunu ileri sürmüşlerdir [10]. Fahreddinoğlu, Azerbaycan'daki Tüketici Fiyat Endeksi (TÜFE) serisini kullanarak ARIMA ve ARFIMA modellerinin performanslarını kıyaslamıştır ve en iyi modelin ARFIMA model olduğuna karar vermiştir [11].

Bu çalışmanın sonraki bölümleri aşağıdaki şekilde düzenlenmiştir. Bölüm 2'de, uzun hafızalı model hakkında temel bilgiler sunulmuştur. Bölüm 3'te, petrol kiralari üç ülke üzerinde mevcut yöntemlerle kıyaslanarak değerlendirilmiştir. Son bölüm makaleyi sonlandırmakta ve sunulan modelinin nasıl daha da genişletilebileceğine dair bazı fikirler sunmaktadır.

II.METODOLOJİ

2.1.Uzun Hafıza Modelleri

Gerçek hayatta, gözetim videoları, gerçek zamanlı sensörler, hisse senedi fiyatları ve astrofizik gözlemleri gibi zaman serisi verileriyle sıkça karşılaşırız. Zaman serisi verilerini analiz ederken, bu verilerdeki çeşitli noktalardaki bağımlılıklar dikkate alınması gereken en önemli unsurlardan biridir. Zaman serisi verileri

arasında herhangi bir bağımlılık var mı? Eğer öyleyse, bu değerler birbirlerini etkilemek için zaman açısından ne kadar uzak olmalıdır? Bu sorulara cevap ararken, uzun hafıza (long memory) kavramı devreye girer.

Uzun hafıza ve kısa hafıza varlığını anlamamızın çeşitli yolları vardır. Kısa hafıza varlığında, noktalar arasındaki zaman farkı arttıkça farklı zamanlardaki değerler arasında bağımlılık hızla azalır. Bu durumda, otokorelasyon fonksiyonu üstel bir şekilde azalır veya belirli bir zaman gecikmesinden sonra noktalar bağımsız hale gelerek 0'a düşer. Diğer bir ifadeyle, eğer bugünün gözlemi ile yarını tahmin etmek istenirse, serinin ilk gecikmesinde otokorelasyon fonksiyonunda kısa hafızalı bir süreç bulunmalıdır. Uzun hafıza süreçlerinde ise otokorelasyon fonksiyonunun azalması daha yavaş olduğundan, yani bağımlılık daha güçlü olduğundan, serinin birkaç zaman önceki gözlemleri bugünü etkiler. Bu durum, sürecin uzun hafızalı süreç olduğunu gösterir [12].

Zaman serilerinin istatistiksel çıkarımlarda öngörüler yapılabilmesi için serinin durağanlık varsayımı üzerine kurulması gerekmektedir. Diğer bir deyişle, ilk önce zaman serisinin birim kök içerip içermediğine bakılarak durağanlık varsayımı kontrol edilmelidir. Durağan olmayan serilerde varyans, zamanın bir fonksiyonu şeklinde hareket eder. Bu da, geçmişteki beklenmedik etkilerin seri tarafından kalıcı olarak taşınmasına neden olur. Bu yüzden, durağan olmayan serilerde öngörülerin güvenilirliği sorgulanacaktır [2]. Ancak, ne tam olarak durağan yani $I(0)$ ne de tam olarak durağan olmayan yani $I(1)$ serileri gibi olan ekonomik ve finansal zaman serileri, uzun hafıza sürecinin özelliğini sergilerler [1]. Uzun hafızaya sahip olan serilerin bütünleşik derecesi " d " parametresi $[-0.5, 0.5]$ arasındaki bir kesirli sayı olarak modellenmektedir [13].

2.1.1.Otoregresif kesirli bütünleşik hareketli ortalama modeli

Granger ve Joyeux 1980 yılında [14] ve Hosking 1981 yılında [15], uzun hafıza sürecinin belirlenmesinde kullanılan ARFIMA modelini ortaya atmışlardır.

y_t , bir zaman serisi olarak tanımlansın. ARFIMA(p, d, q) modeli,

$$\phi(L)(1-L)^d y_t = \mu + \theta(L)\varepsilon_t, t = 1, 2, \dots, T \quad (1)$$

formunda verilebilir. Burada, $\varepsilon_t \sim iid(0, \sigma^2)$, L gecikme operatörü, $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$, $\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$ ve μ , zamanın herhangi bir deterministik fonksiyonu olabilir. Bu modelde, d değeri kesirli olduğu için kesirli bütünleşik (FI) model olarak adlandırılır. Eğer $d = 0$ ise y_t trend-durağan; $d = 1$ (veya daha büyük tamsayı) ise y_t fark-durağan; $d \in (-\frac{1}{2}, 0) \cup (0, \frac{1}{2})$ ise FI modeller söz konusudur.

ARFIMA model, ARIMA modelin bütünleşik (farklılaştırma) parametresi olan d 'yi, tamsayı ile sınırlamaktan ziyade onun herhangi bir gerçek değeri almasına olanak tanımaktadır. d parametresi, Binom açılımından yararlanarak hareketli ortalama (MA) süreçlerinin sonsuz basamağı olarak aşağıdaki gibi ifade edilebilir:

$$(1 - L)^d = 1 - dL + \frac{d(d-1)}{2!}L^2 - \frac{d(d-1)(d-2)}{3!}L^3 + \dots$$

$$= \sum_{k=0}^{\infty} \frac{\Gamma(-d)L^k}{\Gamma(-d)\Gamma(k+1)}$$

Burada, Gamma fonksiyonu,

$$\Gamma(g) = \int_0^{\infty} x^{g-1} e^{-x} dx$$

olarak tanımlanır. ARFIMA modelinde d parametresinin rolünü görmek için y_t 'nin kovaryans fonksiyonunun bilinmesi gerekmektedir. Varyans γ_0 , otokovaryans fonksiyonu γ_τ ve otokorelasyon fonksiyonu ρ_τ , sırasıyla

$$\gamma_0 = \frac{\sigma^2 \Gamma(1-2d)}{\Gamma(1-d)\Gamma(1-d)}$$

$$\gamma_\tau = \frac{\sigma^2 \Gamma(1-2d)\Gamma(\tau+d)}{\Gamma(d)\Gamma(1-d)\Gamma(\tau+1-d)}, \quad \tau = \pm 1, \pm 2, \dots$$

ve

$$\rho_\tau = \frac{\Gamma(1-d)\Gamma(\tau+d)}{\Gamma(d)\Gamma(\tau+1-d)}$$

biçimindedir.

Eğer $\phi(L)$ ve $\theta(L)$ 'nin kökleri birim çemberin dışında ve $d < |0.5|$ ise y_t süreci hem durağan hem de tersinir seri olarak tanımlanır. Fakat, $\tau \rightarrow \infty$, $\rho_\tau \propto \tau^{2d-1}$ olduğundan otokorelasyon fonksiyonları sonlu bir toplama sahip olmayacaktır. Başka bir ifadeyle, $0 < d < 0.5$ iken ARFIMA modellerinde uzun hafıza sürecinin varlığından bahsedilir. Yani, benzer otoregresif hareketli ortalama (ARMA) süreçlerine kıyasla bu süreçlerin otokorelasyon fonksiyonu daha yavaş bir bozulma sergilemektedir. $0.5 < d < 1.0$ iken γ_0 , varyansı sonlu bir toplama sahip olmadığından y_t süreci durağan değildir. Böylece, Hosking'in formülü olan (6) denklemindeki otokorelasyon fonksiyonu hala 0'a doğru bir azalış sergileyecektir.

Süreçlerin hafızasının sonlu olduğu durumda bir şok verildiği zaman bu süreçlerin ortalama geri döndürme eğiliminde (mean-reverting) olduğu anlamına gelmektedir. $d > 1.0$ için, süreç ortalama geri dönme eğiliminde değil ve sürece bir şok uygulandığında sürecin başlangıç noktasından sapmasına sebep olmaktadır. Böylece, bir ARFIMA modelinin hafıza özelliği önemli bir şekilde d parametresine bağlıdır ve otokorelasyon fonksiyonları, $I(0)$ serilerine nispeten yavaş bir şekilde daha küçük oranda azalış sergiler [16].

2.2. Uzun Hafıza Modellerinin Tahmin Yöntemleri

Literatürde, iki aşamalı ve tek aşamalı yöntem olmak üzere uzun hafıza tahmin yöntemleri iki grup altında

toplanmaktadır. Bu çalışmada iki aşamalı tahmin yöntemi kullanılmıştır [17].

Geweke ve Porter-Hudak (GPH) tarafından önerilen yarı-parametrik iki aşamalı bir yöntem olan GPH metodunun ilk aşamasında d bütünleşik parametresi tahmin edilmektedir [18]. İkinci aşamasında ise diğer otoregresif (AR) ve MA parametreleri tahmin edilmektedir. Uygulamada da oldukça sık tercih edilen bir yöntemdir [16]. GPH yönteminin temeli log-periodogram regresyonuna dayandığından dolayı bir regresyon yöntemi olarak ifade edilir.

Durağan bir serinin spektral yoğunluğu,

$$f(\lambda) = f_0(\lambda)[2 \sin(\lambda/2)]^{-2d}$$

biçiminde tanımlanır. (7) eşitliğinin logaritması alınarak aşağıdaki denklem elde edilir:

$$\log f(\lambda_j) = \log f_0(0) - d \log [2 \sin \frac{\lambda_j}{2}]^2 + \log \left[\frac{f_0(\lambda_j)}{f_0(0)} \right]$$

Burada $\lambda_j = 2\pi j / n$, Fourier frekanslardaki spektral yoğunluktur. Başka bir ifadeyle, $I(\lambda_j)$ periodogramının logaritması aşağıdaki gibi ifade edilmektedir:

$$\log I(\lambda_j) = \log \left[\frac{I(\lambda_j)}{f(\lambda_j)} \right] + \log f(\lambda_j)$$

(8) ve (9) eşitliklerinden yararlanarak,

$$\log I(\lambda_j) = \log f_0(0) - d \log [2 \sin \frac{\lambda_j}{2}]^2 + \log \left\{ \frac{I(\lambda_j)[2 \sin(\lambda_j/2)]^{2d}}{f_0(0)} \right\}$$

denklemini elde edilir. $y_j = \log I(\lambda_j)$, $\alpha = \log f_0(0)$, $\beta = -d$, $x_j = \log [2 \sin(\lambda_j/2)]^2$ ve

$\varepsilon_j = \log \left\{ \frac{I(\lambda_j)[2 \sin(\lambda_j/2)]^{2d}}{f_0(0)} \right\}$ biçiminde olmak üzere, aşağıdaki regresyon denklemi elde edilebilir:

$$y_j = \alpha + \beta x_j + \varepsilon_j, \quad j = 1, \dots, m$$

Bu durumda, $f(\lambda_j) \sim f_0(0)[2 \sin(\lambda_j/2)]^{-2d}$ ve $\varepsilon_j \sim \log \left[\frac{I(\lambda_j)}{f(\lambda_j)} \right]$ olmaktadır. Böylece, uzun hafıza parametresi en küçük kareler tahmini,

$$\hat{d}_m = - \frac{\sum_{j=1}^m (x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^m (x_j - \bar{x})^2}$$

biçiminde verilir. Burada, $\bar{x} = \sum_{j=1}^m x_j / m$ ve $\bar{y} = \sum_{j=1}^m y_j / m$ 'dir [13].

2.3. Model Seçiminde Bilgi Kriterleri

Uygulamalarda parametrik olarak belirlenen modeller için seçim kriterleri yardımıyla en iyi model belirlenmesi yapılmaktadır [19]. ARFIMA modellerinin seçimi için çeşitli uygulamalarda AIC, BIC, Schwarz bilgi kriteri (SIC) ve Hannan-Quinn bilgi kriteri (HQC-HIC) gibi seçim yöntemleri mevcuttur [17, 20, 21, 22]. Ancak, uzun hafızalı modellere bu seçim yöntemlerin uygulanması sonucunda etkinliği ve/veya yanlılığına ilişkin herhangi bir teorik sonuç bulunmamaktadır. Ancak, son zamanlarda pek çok gelişmeler mevcuttur.

Crato ve Ray, geniş kapsamlı bir simülasyon çalışması yaparak, üç farklı tahmin yöntemi ve üç otomatik model seçimi kriterinin tahmin doğruluğuna etkilerini karşılaştırmışlardır [23]. Schmidt ve Tschmig, ARFIMA(0, d, 0) süreçlerinde yanlılığı düzeltilmiş AIC (AICc), BIC ve HQC kriterlerinin performansını değerlendirmek için küçük ölçekli bir simülasyon çalışması gerçekleştirmişlerdir [24]. Bu kriterler temelde bir Gauss ARMA(p, q) sürecinin log olabilirlik fonksiyonunun yaklaşık maksimum değeri parametre sayısı için bir ceza cinsinden ifade edilmektedir. Kriterlerin birbirinden farklı olması, cezanın aşırı uyum eğilimine kaşı koymasından kaynaklanmaktadır. Bilgi kriterleri arasında en küçük değere sahip olan en iyi model olarak belirlenir. Bu bilgi kriterleri eşitlik 13 ve 14'teki gibi tanımlanır:

$$AIC = \ln \hat{\sigma}_N^2 + 2(p + q + 1) \quad (13)$$

$$BIC = N \ln \hat{\sigma}_N^2 + (p + q + 1) \log N \quad (14)$$

Burada, $\hat{\sigma}_N^2$, akgürültü varyansıdır [20].

III.UYGULAMA

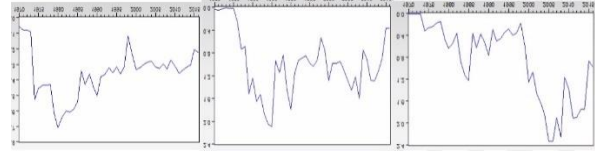
Bu çalışma için petrol kiralaları (OR)'ye ait seriler dikkate alınmıştır [25]. Seriler, Türkiye, Hindistan ve Brezilya ülkeleri için 1970 ile 2016 dönemini kapsayan yıllık değerler içermektedir. Burada OR, dünyadaki ham petrol üretiminin değeri ile toplam üretim değerleri arasındaki fark olarak elde edilmiştir.

Türkiye, Hindistan ve Brezilya için serilerinin normallik varsayımının sağlanıp sağlanmadığı Tablo 1'de sunulmuştur. Türkiye serisinde normallik varsayımı sağlanmadığı için en uygun dönüşüm olan karekök dönüşümü uygulanmıştır. Türkiye, Hindistan ve Brezilya serilerinin sırasıyla tablo değerlerini incelersek karekök dönüşümü uygulanan Türkiye serisinde Jarque-Bera istatistiğine karşılık gelen $p = 0.879894$ olup 0.05 anlam düzeyinden büyüktür. Bu nedenle, bu serinin normal dağıldığı görülmektedir. Aynı şekilde, Hindistan ve Brezilya seriler için de Jarque-Bera istatistiğine karşılık gelen p değerleri, 0.05 anlam düzeyinden büyük olduğundan, bu serilerin de normal dağıldığı açıkça görülmektedir.

Tablo 1. Seriler için tanımlayıcı istatistikler

	TÜRKİYE	HİNDİSTAN	BREZİLYA
Mean	0.359	0.973	0.851
Median	0.335	0.971	0.599
Maximum	0.710	2.090	2.328
Minimum	0.059	0.002	0.000
Std. Dev.	0.151	0.542	0.693
Skewness	0.181	-0.125	0.695
Kurtosis	2.989	2.590	2.251
Jarque Bera	0.256	0.452	4.879
Probability	0.880	0.798	0.087

Türkiye, Hindistan ve Brezilya serilerinin çizgi grafikleri Şekil 1'de sunulmuştur. Her üç serinin durağan olmadığı açıkça görülmektedir. Ayrıca, Türkiye serisi için 1970 ile 1980 yılları arasında, Hindistan serisi için ise 1975 ile 1985 yılları arasında bir artış görülmektedir. Brezilya serisinde ise 2000 ile 2005 yılları arasında bir artış, 2005'ten sonrası için ise bir azalış söz konusudur.



Şekil 1. TÜRKİYE, HİNDİSTAN ve BREZİLYA serilerinin çizgi grafikleri

Tablo 2'de serilerin durağanlıklarını kontrol etmek için Augmented Dickey-Fuller (ADF) ve Phillpps-Perron (PP) birim kök testlerinin sonuçları yer almaktadır. Tablodaki serilerin ADF ve PP birim kök testlerinin test istatistiği ve olasılık değerlerine bakıldığında, Türkiye, Hindistan ve Brezilya serilerinin her iki birim kök testinde p -değerlerinin 0.05 anlam düzeyinden büyük olduğu görülmektedir. Bu nedenle, serilerin birim kök içerdiği ve dolayısıyla durağan olmadığı yorumu yapılabilir.

Tablo 2. Serilerin ADF ve PP birim kök test sonuçları

		ADF Testi		PP Testi	
		Sabit	Trend Sabit	Sabit	Trend Sabit
TÜRKİYE	t-istatistiği	-2.691	-3.152	-2.644	-3.056
	Prob	0.083	0.11	0.092	0.129
HİNDİSTAN	t-istatistiği	-2.613	-2.350	-2.475	2.088
	Prob	0.098	0.08	0.1281	0.539
BREZİLYA	t-istatistiği	-1.968	-2.477	-1.888	-2.456
	Prob	0.300	0.338	0.335	0.347

Tablo 3'te ise birinci farkı alınan her üç serinin birim kök testlerinin sonuçları sunulmuştur. Sonuçlar incelendiğinde, serilerin p -değerleri 0.05 anlam düzeyine göre kıyaslandığında serilerinin durağan oldukları açıkça görülmektedir.

Tablo 3. Birinci fark seriler için ADF ve PP birim kök test sonuçları

		ADF Testi		PP Testi	
		Sabit	Trend Sabit	Sabit	Trend Sabit
TÜRKİYE	t-istatistiği	-7.328	-7.509	-7.450	-7.662
	Prob	0.000	0.000	0.000	0.000
HİNDİSTAN	t-istatistiği	-7.932	-8.132	-8.022	-9.323
	Prob	0.000	0.000	0.000	0.000
BREZİLYA	t-istatistiği	-7.927	-7.907	-7.937	-7.919
	Prob	0.000	0.000	0.000	0.000

Serilerin durağanlık varsayımı incelendikten sonra, uzun hafızanın varlığından bahsedebilmek için Hurst'ün R/S istatistiğinden yararlanılmıştır. Bu test istatistiğinde, güven aralıkları kıyaslanarak serilerde uzun hafıza varlığının olup olmadığı belirlenir. Tablo 4'te bu serilere ait uzun hafıza varlığının sonuçları yer almaktadır. Sonuçlar incelendiğinde, test istatistikleri değerlerinin %90, %95 ve %99 güven düzeylerinin tamamında aralığın dışında kaldığı görülmektedir. Bu durum, her üç serinin uzun hafızaya sahip olduğunu göstermektedir.

Tablo 4. Serilere ait R/S test sonuçları

Güven Düzeyleri	TÜRKİYE	HİNDİSTAN	BREZİLYA
90%	[0.861, 1.747]	[0.861, 1.747]	[0.861, 1.747]
95%	[0.809, 1.862]	[0.809, 1.862]	[0.809, 1.862]
99%	[0.721, 2.098]	[0.721, 2.098]	[0.721, 2.098]
Test İstatistiği	2.51	2.21	2.78

Bu çalışmada, $p, q \leq 2$ için tüm ARFIMA(p, d, q) ve ARIMA(p, d, q) modelleri üzerinde durulmuştur. En iyi modelin belirlenebilmesi için parametrelerin istatistiksel anlamlılığına bakılmıştır. Anlamlı parametreler * olarak işaretlenmiştir. Tablo 5'te Türkiye serisine ait ARFIMA ve ARIMA modellerinin parametre tahmin sonuçları verilmiştir. Her bir model için sonuçlar incelendiğinde, %5 anlam düzeyine göre kıyaslandığında ARFIMA(1, d , 0) modelinde sabit ve AR(1) parametrelerinin anlamlı olduğu, ancak d parametresinin anlamlı olmadığı ve ARFIMA(1, d , 1) modelinde ise sabit ve AR(1)'in anlamlı olduğu, fakat MA(1) ve d 'nin anlamlı olmadığı görülmektedir. Bunun yanı sıra, ARFIMA(2, d , 0) ve ARFIMA(0, d , 2) modellerinde de bazı parametrelerin anlamlı olduğu görülmektedir. ARIMA modellerinden, ARIMA(1, 1, 1) modelinde sadece AR(1) parametresi anlamlı iken, diğer ARIMA modellerinde ise hiçbir parametre tahminleri anlamlı değildir. Yani, serideki ilişkiyi açıklamakta ARIMA modeller yetersiz kalmaktadır. Seri için en iyi modelin ARFIMA(0, 0.423, 1) olduğu açıkça görülmektedir.

Tablo 5. TÜRKİYE serisi için ARFIMA-ARIMA model sonuçları

	Sabit	AR(1)	AR(2)	MA(1)	MA(2)	d
ARFIMA(1, d , 0)	0.323*	0.856*				-0.079
ARFIMA(0, d , 1)	0.320*			0.403*		0.423*
ARFIMA(1, d , 1)	0.327*	0.870*		0.105		-0.172
ARFIMA(2, d , 0)	0.329*	1.008*	-0.118			-0.204
ARFIMA(0, d , 2)	0.321*			0.413*	0.184	0.388
ARIMA(1, 1, 0)	0.004	-0.108				
ARIMA(0, 1, 1)	0.004			-0.144		
ARIMA(1, 1, 1)	0.001	0.864*		-1		
ARIMA(2, 1, 0)	0.004	-0.122	-0.126			
ARIMA(0, 1, 2)	0.004			-0.121	-0.121	

Hindistan serisine ait ARFIMA ve ARIMA modellerinin parametre tahmin sonuçları Tablo 6'da yer almaktadır. Sonuçlara göre, ARIMA modellerde %5

anlam düzeyine göre kıyaslandığında istatistiksel olarak anlamlı parametre tahmini bulunmadığından bu modeller seriyi açıklamada yetersiz kalmaktadır. ARFIMA modellerden ise ARFIMA(0, 0.442, 1) ile ARFIMA(0, 0.342, 2) modelleri dışında diğer ARFIMA modellerinin bazı parametre tahminleri anlamlı olmadığı açıkça görülmektedir.

Tablo 6. HİNDİSTAN serisine için ARFIMA-ARIMA model sonuçları

	Sabit	AR(1)	AR(2)	MA(1)	MA(2)	d
ARFIMA(1, d , 0)	0.857*	0.862*				-0.127
ARFIMA(0, d , 1)	0.182*			0.262*		0.442*
ARFIMA(1, d , 1)	0.854*	0.858*		-0.025		-0.103
ARFIMA(2, d , 0)	0.850*	0.804*	0.043			-0.079
ARFIMA(0, d , 2)	0.836*			0.382*	0.411*	0.342*
ARIMA(1, 1, 0)	0.007	-0.184				
ARIMA(0, 1, 1)	0.008			-0.196		
ARIMA(1, 1, 1)	0.010	0.345		-0.537		
ARIMA(2, 1, 0)	0.007	-0.182	0.013			
ARIMA(0, 1, 2)	0.008			-0.202	-0.031	

Hindistan serisi için iki model arasından en iyiye karar verebilmek amacıyla Tablo 7'de yer alan AIC ve BIC bilgi kriterlerinden yararlanılmıştır. AIC ve BIC bilgi kriterlerinden en küçük olan ARFIMA(0, 0.342, 2) modelinin bu seri için en iyi model olduğu açık bir şekilde görülmektedir.

Tablo 7. HİNDİSTAN serisinin bilgi kriteri sonuçları

	ARFIMA(0, 0.442, 1)	ARFIMA(0, 0.342, 2)
AIC	46.956	43.928
BIC	54.356	53.179

Aynı şekilde, Brezilya serisine ait ARFIMA ve ARIMA modellerinin parametre tahmin sonuçları Tablo 8'de sunulmuştur. Bu tabloya göre, ARIMA modeller ile serideki ilişkiyi açıklamak yetersiz kalırken, ARFIMA modellerde ise sadece ARFIMA(0, 0.462, 1) modeli seriyi açıklamada yeterli olmaktadır.

Tablo 8. BREZİLYA serisi için ARFIMA-ARIMA model sonuçları

	Sabit	AR(1)	AR(2)	MA(1)	MA(2)	d
ARFIMA(1, d , 0)	0.761*	0.927*				-0.136
ARFIMA(0, d , 1)	0.804			0.322*		0.462*
ARFIMA(1, d , 1)	0.751*	0.905		-0.133		-0.013
ARFIMA(2, d , 0)	0.752*	0.785*	0.1097			-0.028
ARFIMA(0, d , 2)	0.788			0.274	0.266	-0.028
ARIMA(1, 1, 0)	0.021	-0.184				
ARIMA(0, 1, 1)	0.022			-0.210		
ARIMA(1, 1, 1)	0.024	0.282		-0.484		
ARIMA(2, 1, 0)	0.022	-0.110	-0.088			
ARIMA(0, 1, 2)	0.023			-0.201	-0.065	

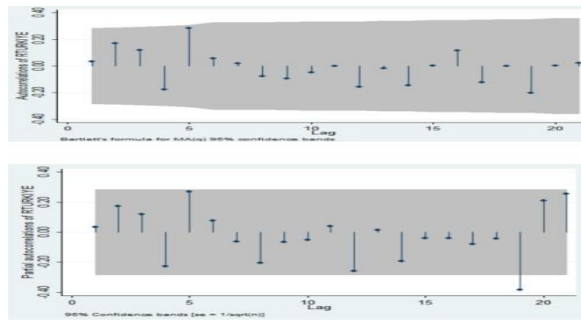
Şekil 2'de birinci farkı alınan Türkiye, Hindistan ve Brezilya serilerinin korelogramları sırasıyla sunulmuştur. Bu korelogramlar incelendiğinde, birinci farkı alınan serilerin sırasıyla otokorelasyon ve kısmi otokorelasyon değerlerinin güven aralığı içinde olduğu görülmektedir. Bir başka deyişle, bu seriler için AR ve MA terimlerinin etkisi devreden çıktığından dolayı ARIMA modeller anlamlı değildir. Böylece, her üç seri

hem uzun hafızaya sahip hem de ilişkiyi açıklamada ARFIMA modelinin en iyi olduğu yorumu yapılabilir.

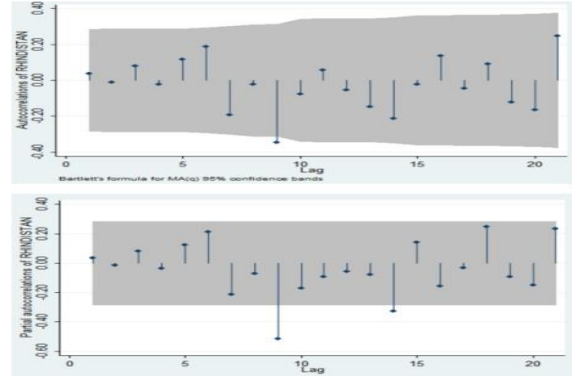
Model	ARFIMA	ARMA	ARIMA	ARFIMA	ARMA	ARIMA	ARFIMA	ARMA	ARIMA
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
12	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
13	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
14	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
16	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
17	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
18	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
19	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
20	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Şekil 2. TÜRKİYE, HİNDİSTAN ve BREZİLYA serilerinin korelogramları

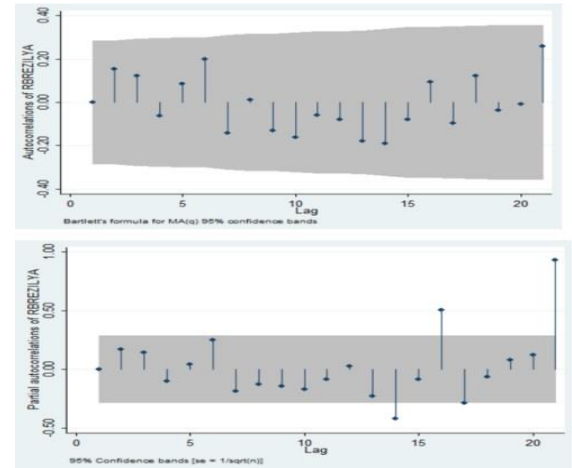
En iyi model belirlendikten sonra, bu modellerin artıkları için varsayımlar kontrol edilmelidir. Başka bir ifadeyle, modellere ait artıkların akgürültü (white-noise) özelliğine sahip olması gerekmektedir. Şekil 3'te Türkiye serisi için ARFIMA(0,0.423,1) modelinin artıklarının otokorelasyon fonksiyonu (ACF) ve kısmi otokorelasyon fonksiyonu (PACF) grafikleri verilmiştir. Bu grafikler incelendiğinde, modelden elde edilen artıkların gecikme değerleri güven bölgesinin içinde yer aldığından ilişkisiz olduğu görülmektedir. Aynı şekilde, Şekil 4'te verilen Hindistan serisi için ARFIMA(0,0.342,2) modelinin artıkları incelendiğinde, artıklar arasında PACF grafiğinde dokuzuncu ve on dördüncü gecikmelerde güven bölgesinin dışına çıktığı görülmüştür. Ancak, bu durum ihmal edilerek artıkların ilişkisiz olduğu yorumu yapılabilir. Brezilya serisi için ARFIMA(0,0.462,1) modelinin artık ilişki grafikleri ise Şekil 5'te sunulmuştur. Burada da PACF grafiğinde birkaç gecikme güven aralığının dışındadır. Fakat, ilk on gecikmenin ötesi ihmal edilerek artıkların ilişkisiz olduğu sonucuna varılabilir.



Şekil 3. TÜRKİYE serisi için ACF (üstteki) ve PACF (alttaki) grafikleri



Şekil 4. HİNDİSTAN serisi için ACF (üstteki) ve PACF (alttaki) grafikleri



Şekil 5. BREZİLYA serisi için ACF (üstteki) ve PACF (alttaki) grafikleri

Artıklar için diğer önemli varsayımlardan biri de otoregresif koşullu değişen varyans (ARCH) testi ile değişen varyanslılığın olup olmadığının belirlenmesidir. Tablo 9'da sırasıyla Türkiye, Hindistan ve Brezilya serileri için ARCH testi p değerlerinin sonuçları yer almaktadır. Sonuçlar incelendiğinde, ilk üç gecikmenin olasılık değerlerinin %5 anlam düzeyinden büyük olduğu görülmektedir. Bu nedenle, her üç serinin artıkları değişen varyansa sahip değildir.

Tablo 9. Seriler için ARCH testi p değerleri sonuçları

	1	2	3
TÜRKİYE	0.286	0.627	0.665
HİNDİSTAN	0.286	0.627	0.665
BREZİLYA	0.743	0.950	0.998

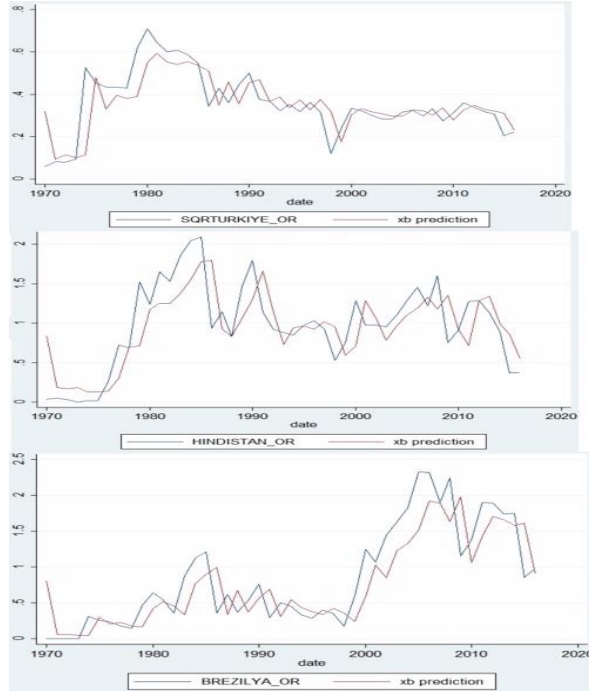
Aslında akgürültü, modelin geçerliliğinde ve tahmin edilmesinde önemli rol oynamaktadır. Artıklar için son varsayımlardan biri de akgürültü özelliğine sahip olması gerekliliğidir. Bunun için Bartlett ve Portmanteau testlerinden yararlanılmıştır. Tablo 10'da yer alan Bartlett ve Portmanteau testlerinin sonuçları incelendiğinde, her iki testin parantez içinde yer alan olasılık değerlerinin 0.05 anlam düzeyinden büyük

olduğu görülmektedir. Bu nedenle, her üç serinin artıkları akgürültü özelliğine sahiptir. Başka bir deyişle, Türkiye, Hindistan ve Brezilya serileri için kurulan modeller geçerlidir ve ileriye yönelik tahmin yapılabilmektedir.

Tablo 10. Seriler için Bartlett ve Portmanteau istatistik sonuçları

	Bartlett'in (B) istatistiği	Portmanteau (Q) istatistiği
TÜRKİYE	0.719(0.679)	8.402(0.135)
HİNDİSTAN	0.698(0.715)	1.170(0.948)
BREZİLYA	0.864(0.445)	2.641(0.755)

Bütün varsayımlar incelendikten sonra, Türkiye, Hindistan ve Brezilya serilerinin gerçek ve kestirim değer grafikleri sırasıyla Şekil 6'da yer almaktadır. Grafiklere göre her üç seri için gerçek değerler ile kestirim değerlerinin birbirine yakın olması sebebiyle ARFIMA modellerinin seriler için uygun olduğu söylenebilmektedir.



Şekil 6. TÜRKİYE, HİNDİSTAN ve BREZİLYA serilerinin gerçek ve kestirim değer grafikleri

Tablo 11'de her üç seri için gelecek 5 yılın tahmin sonuçları ve parantez içinde gerçek değerleri yer almaktadır. ARFIMA modeli ile tahmin sonuçları incelendiğinde, Türkiye serisinde bir artış olduğu ve bu artışın belli bir sürede kalıcı olabileceği görülmektedir. Hindistan serisi için de yine bir artış ve belli bir sürede kalıcılık söz konusudur. Brezilya serisi için ise belli bir dönem artış sonrasında bir azalış olduğu görülmektedir. Hata Kareler Ortalamasının Karekökü (RMSE) değerleri incelendiğinde, ARFIMA modeli Türkiye için çok iyi, Hindistan için makul ve Brezilya için nispeten düşük bir tahmin performansı

göstermektedir. Modelin farklı ülkelerde farklı performans sergilemesi, her ülkenin petrol kiralari dinamiklerinin ve değişkenliklerinin farklı olmasından kaynaklanabilir. Türkiye için ARFIMA modelinin oldukça iyi bir tahmin yapması, modelin uzun hafızalı verilerde başarılı olduğunu ve Türkiye'nin petrol kiralari'nin belirgin bir uzun hafıza varlığı olduğunu düşündürülebilir. Hindistan ve Brezilya için RMSE değerlerinin daha yüksek olması, bu ülkelerin petrol kiralari'nde daha fazla kısa dönemli dalgalanmalar veya modelin yakalayamadığı belirli olaylar olabileceğini gösterebilir.

Tablo 11. Seriler için 2017-2021 tahmini değer sonuçları

	TÜRKİYE	HİNDİSTAN	BREZİLYA
2017	0.070 (0.054)	0.4351 (0.313)	1.120 (1.128)
2018	0.078 (0.108)	0.608 (0.447)	1.1261 (1.901)
2019	0.082 (0.103)	0.719 (0.305)	1.123 (1.676)
2020	0.085 (0.061)	0.7593 (0.144)	1.115 (1.079)
2021	0.087 (0.138)	0.783 (0.326)	1.106 (2.605)
RMSE	0.031	0.310	0.794

IV. TARTIŞMA VE SONUÇLAR

Uzun hafızaya sahip bir zaman serisinin önceki dönemlerdeki davranışlarını modele yansıtmak, FI modellerinde öngörü başarısını artırmaktadır. ARIMA gibi modeller ise kısa hafızalıdır ve genellikle serinin son dönemlerdeki davranışlarına dayanarak modelleme yaparlar. Ayrıca, AR(10) ve AR(20) gibi birçok parametre içeren modeller, basitlik (parsimoni) ilkesine aykırı oldukları ve başarı seviyeleri düşük olduğu için uygun olmayabilir. Bu sebeple, ARFIMA modelleri, uzun hafızalı olmaları ve genellikle az sayıda parametre ile tahmin edilebilmeleri nedeniyle daha avantajlıdır [26].

Bu çalışmada, ekonomik seri olan petrol kiralari üzerinden verimli bir tahmin yapabilmek için ARFIMA modelini sunuyoruz. Bölüm 2'de tartışılan teorik çalışma, petrol kiralari'nin tahmin edilmesinde klasik modele kıyasla parametreleri tahmin etmede ve ileriye yönelik tahminlerde daha iyi sonuçlar vermiştir. Klasik ARIMA modelleri parametre tahmininde düşük bir performans sergilemiştir. Sayısal sonuçlara ve grafiğe baktığımızda da ARIMA modeli ile parametre tahmin etmede ARFIMA modeli alternatif olarak düşünülebilir.

Bu çalışmada sunulan ARFIMA modeli, diğer yarı-parametrik veya parametrik yöntemlerle genişletilebilir. Örneğin, model performansı ve belirlenmesi, Whittle metodu gibi tekniklerle sağlanabilir. Yarı-parametrik veya parametrik yöntemlerle yapılan karşılaştırmalı analizler, modelin avantajlarını ve sınırlamalarını daha iyi anlamamıza olanak tanır. Bu karşılaştırmalar, ARFIMA modelinin

farklı veri setlerinde ve koşullarda nasıl davrandığını incelemek için faydalı olabilir.

ARFIMA modeli çalışmasında, modelin belirlenmesi ve optimizasyonu için bilgi kriterlerinden yararlanılmaktadır. Modelin, en iyi bilgi kriterine göre değerlendirilmesi, farklı yapılandırmalarda nasıl performans gösterdiğini ve en uygun parametrelerin nasıl belirlendiğini ortaya koyabilir. Ayrıca, ARFIMA modelinin çeşitli veri setleri ve uygulama alanlarında test edilmesi, modelin geliştirilebilirliğini ve pratik faydasını artırabilir. Örneğin, finansal zaman serileri, iklim verileri veya ekonomik göstergeler gibi farklı alanlarda modelin performansı incelenebilir. Bu, modelin doğruluğunu ve genel performansını artırma potansiyeline sahiptir.

TEŞEKKÜR

Bu çalışma, Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü İstatistik Bölümünde hazırlanmış “Otoregresif Kesirli Bütünleşik Hareketli Ortalama (Arfima) Modelinin Belirlenmesi: Türkiye, Hindistan ve Brezilya Ülkeleri Üzerine Uygulamalar” adlı yüksek lisans tezinden üretilmiştir.

KAYNAKLAR

- [1] Kutlar, A. ve Turgut, T. (2006). Türkiye’deki Başlıca Ekonomi Serilerinin ARFIMA Modelleri ile Tahmini ve Öngörülebilirliği. *Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 11(1), 120-149.
- [2] Barışık, S. ve Çevik, E. İ. (2008). İşsizlikte Histeri Etkisi: Uzun Hafıza Modelleri. *Kamu-İş*, 9(4), 1-36.
- [3] Çevik, E. İ. ve Erdoğan, S. (2009). Bankacılık Sektörü Hisse Senedi Piyasasının Etkinliği: Yapısal Kırılma ve Güçlü Hafıza. *Doğuş Üniversitesi Dergisi*, 10 (1), 26- 40.
- [4] Pekkaya, M. (2011). Arfima ve Figarch Yöntemlerinin Markowitz Ortalama Varyans Portföy Optimizasyonunda Kullanılması: İMKB-30 Endeks Hisseleri Üzerine Bir Uygulama. Doktora Tezi, Zonguldak Bülent Ecevit Üniversitesi, Türkiye.
- [5] Çevik, E. İ. (2012). İstanbul Menkul Kıymetler Borsası’nda Etkin Piyasa Hipotezinin Uzun Hafıza Modelleri İle Analizi: Sektörel Bazda Bir İnceleme. *Yaşar Üniversitesi E-Dergisi*, 7(26), 4437-4454.
- [6] Karia, A. A., Bujanga, I. ve Ahmad, İ. (2013). Fractionally Integrated ARMA for Crude Palm Oil Prices Prediction: Case of Potentially Overdifference. *Journal of Applied Statistics*, 40(12), 2735–2748.
- [7] Pekkaya, M., ve Albayrak, A. S. (2013). Arfima ve Figarch Yöntemlerinin Markowitz Ortalama Varyans Portföy Optimizasyonunda Kullanılması: İMKB-30 Endeks Hisseleri Üzerine Bir Uygulama. *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 42(1), 1303-1732.
- [8] Yılmaz, Ç. (2015). Hisse Senedi Getirilerindeki Uzun Hafıza Etkisinin ve Volatilitenin Belirlenmesi: BRIC Ülkeleri Üzerine Bir Uygulama. Yüksek Lisans Tezi, İstanbul Üniversitesi, Türkiye.
- [9] Boateng, A., Gil-Alana, L. A., Maseka, L., Siweya, H. ve Belete, A. (2016). Long Memory and Arfima Modelling: The Case of CPI Inflation Rate in Ghana. *The Journal of Developing Areas*, 50(3), 287-304.
- [10] Omekara C. O., Okereke O. E. ve Ukaeue L. U. (2016). Forecasting Liquidity Ratio of Commercial Banks in Nigeria. *Microeconomics and Macroeconomics*, 4(1), 28-36.
- [11] Fahreddinoğlu, H. G. (2017). Enflasyon’daki Hareketliliğin ARIMA ve ARFIMA Modelleriyle Belirlenmesi ve Tahmin Edilmesi: Azerbaycan Örneğinde. 3. Uluslararası Öğrenciler Sosyal Bilimler Kongresi, İstanbul, Türkiye.
- [12] Joshi, P. (2016). What is Long Memory in Time Series Analysis. <https://prateekvjoshi.com/2016/08/27/what-is-long-memory-in-time-series-analysis/>.
- [13] Palma, W. (2007). Long Memory Time Series Theory and Methods. John Wiley & Sons Inc., Canada.
- [14] Granger, C. W. J. ve Joyeux, R. (1980). An Introduction to Long-Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis*, 1(1), 15-29.
- [15] Hosking, J. R. M. (1981). Fractional Differencing. *Biometrika*, 68(1), 165-176.
- [16] Maddala, G. S. ve Kim, In-M. (2003). Unit Roots, Cointegration and Structural Changes. Cambridge University Press, New York.
- [17] Sowell, F. (1992). Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models. *Journal of Econometrics*, 53(1-3), 165-188.
- [18] Geweke, J. ve Porter-Hudak, S. (1983). The Estimation and Application of Long Memory Time Series Models. *Journal of Time Series Analysis*, 4(4), 221-238.
- [19] Beran, J., Feng, Y., Ghosh, S. ve Kulik, R. (2010). Long-Memory Processes Probabilistic Properties and Statistical Methods, Springer, London.
- [20] Bisaglia, L. (2002). Model Selection for Long-Memory Models. *Quaderni di Statistica*, 4.

-
- [21] Hosking, J. R. M. (1984). Modeling Persistence in Hydrological Time Series Using Fractional Differencing. *Water Resources Research*, 20(12), 1898-1908.
- [22] Cheung, Y. W. (1993). Long-Memory in Foreign Exchange Rates. *Journal of Business and Economic Statistics*, 11(1), 93-101.
- [23] Crato, N. ve Ray, B. K. (1996). Model Selection and Forecasting for Long-Range Dependent Processes. *Journal of Forecasting*, 15(2), 107-125.
- [24] Schmidt, C. M. ve Tschernig, R. (1993). Identification of Fractional ARIMA Models in The Presence of Long Memory. Discussion Paper, University of Munich.
- [25] World Bank Databank, World Development Indicators, Oil rents. (2019). <https://databank.worldbank.org/data/reports.aspx?dsid=2&series=NY.GDP.PETR.RT.ZS>.
- [26] Caballero, R., Jewson, S. ve Brix, A. (2002). Long Memory in Surface Air Temperature: Detection, Modeling, and Application to Weather Derivative Valuation. *Climate Research*, 21(2), 127-140.

Performance of Deep Learning Models on Imputed Time Series Data: A Simulation Study and Application to Leading Airline Companies' Stock Price

İsmail YENİLMEZ¹, Kürşat ATMACA²

¹ Eskişehir Technical University, Science Faculty, Department of Statistics, 26470, Eskişehir, Türkiye

² Turkish Revenue Administration, 06110, Ankara, Türkiye

Abstract

In this study, the validity of imputation techniques for deep learning methods in time series analysis is investigated using datasets based on daily closing data in the stock market. Datasets of daily closing stock prices for Turkish Airlines, Deutsche Lufthansa AG, and Delta Airlines, as well as a simulated dataset, are used. LSTM, GRU, RNN, and Transformer models, which are deep learning models, are employed. The original dataset and datasets with 5%, 15% and 25% missing data are analyzed imputing linear, spline, Stineman, mean and random imputation techniques. The results show that model performance varies depending on the imputation technique and the rate of missing data. GRU and Transformer models are favored for their robustness and excellent performance. For handling missing data, using spline and Stineman imputations is advisable to maintain high model accuracy. This study emphasizes the usability of various imputation techniques and deep learning models in time series analysis. It assesses model performance using both MAPE and RMSE to gain a comprehensive understanding of predictive accuracy and reliability, aiming to guide future research by comparing these methods.

Keywords: Missing Data, Synthetic Data, Transformer Model, RNN, Simulation, Airline Stocks.

I. INTRODUCTION

Financial markets are characterized by their complex and dynamic nature. Time series analysis is a tool for identifying historical trends and patterns in financial data, aiding in the prediction of future movements. In financial markets, time series analysis and panel analysis are used for a variety of data types, including exchange rates, stock prices, etc. These analyses provide insights to make investment decisions, overcome risk management problems, and develop market strategies [1]. On the other hand, panel data analysis, which includes both cross-sectional and time dimensions, can also be applied [2].

Time series analysis employs a range of methods, from traditional statistical techniques, such as ARIMA, SARIMA, and Exponential Smoothing, to modern deep learning models (DLMs) like long short-term memory (LSTM), Gated recurrent unit (GRU), recurrent neural network (RNN), and Transformer model (TM). In particular, the widespread use of TM [3], which is a relatively new method, shows the speed of development in this field. On the other hand, the Box-Jenkins (BJ) method, which is traditionally used in time series analysis, and the Artificial Neural Networks (ANNs) method are compared for demand forecasting [4]. The time series data on Rwanda's and Brazil's GDP per capita are modeled using the conventional BJ approach and the innovative ANNs, respectively [5-6]. While ANNs and DLMs have advantages over traditional methods, it should not be overlooked that they also pose challenges such as hyperparameter tuning [7].

One of the significant challenges in time series analysis is the problem of missing data. Missing data can arise for various reasons and negatively impact analysis accuracy, reducing model prediction reliability and potentially leading to incorrect decisions. Various imputation techniques, such as linear, spline, Stineman, mean, and random imputation, have been used to estimate missing values and reconstruct the time series. The performance of these techniques varies with the nature and the proportion of missing values. The choice of imputation method is crucial to improve the accuracy and reliability of time series analysis [8-9].

In literature, the performance of three machine learning models (ARIMA, LSTM, GRU) is compared for time series forecasting using a Bitcoin price dataset, finding ARIMA to be superior to deep learning-based regression models [10]. The effectiveness of ARIMA and GRU models is assessed in predicting high-frequency stock prices,

demonstrating that the GRU model outperformed the ARIMA model in accuracy [11]. In a comprehensive literature review comparing ARIMA and machine learning algorithms for time series forecasting, as well as their integration in hybrid models, artificial intelligence algorithms exhibit superior predictive performance in most applications, hybrid statistical-artificial intelligence models outperform individual methods by leveraging the best features of both [12]. On the other hand, providing an overview of utilizing Transformer architecture in time series analysis, the study details core components such as self-attention mechanism, positional encoding, multi-head, and encoder/decoder, along with various enhancements and best practices for addressing time series tasks, showcasing the effectiveness of TM [13]. Systematically investigating the usage of TM in time series analysis, the study examines adaptations from both network structure and application perspectives, emphasizing empirical analyses, model size evaluations, and seasonal-trend decomposition to showcase TM's performance and potential for future research [14]. The importance of imputation techniques in handling missing data alongside time series analysis is a separate topic, and has been addressed in [15], investigating imputation of missing values in time series data using deep learning methods [15]. The effects of various imputation methods—including linear, spline, and Stineman interpolation, as well as mean and random sample imputation—on the goodness of fit of statistical models, using synthetic data to control for missing data rate and dataset size has been investigated [16].

The rest of this article is organised as follows. The second section of the study presents the methodologies employed. The third section shares the findings obtained from the analysis. The final section comprehensively assesses and discusses the results.

II. METHOD

2.1. Estimation Techniques

In this section, the modeling techniques and imputation methods utilized in the research are presented. Time series modeling techniques includes the traditional approach of ARIMA, while deep learning methods such as LSTM, GRU, RNN, and Transformer models are discussed. As for imputation techniques, commonly used methods in the literature are discussed, and presented, including Linear, Spline, Stineman, Mean, and Random imputation techniques.

The ARIMA model, characterized by the parameters (p, d, q) , integrates the autoregressive (AR) component of order p , differencing of order d , and moving average (MA) component of order q . Specifically, the AR component involves lagged values of the time series up to p periods, with coefficients denoted as α_i for $i = 1, 2, \dots, p$. The MA component incorporates q lagged

forecast errors, with coefficients θ_j for $j = 1, 2, \dots, q$. The differencing parameter d indicates the number of times the time series is differenced to ensure stationarity. The ARIMA model is represented as in Equation (1):

$$\left(1 - \sum_{i=1}^p \alpha_i L^i\right) (1 - L)^d X_t = \left(1 - \sum_{j=1}^q \theta_j L^j\right) \varepsilon_t \quad (1)$$

where X_t represents the time series value at time t , ε_t is the error term at time t , capturing any residual variation not explained by the ARIMA model, and L denotes backshift (lag) operator such that $L^k X_t = X_{t-k}$.

While traditional methods like ARIMA are still useful, especially for simpler, linear time series problems and for their interpretability, the flexibility, scalability, and performance of deep learning models make them increasingly preferable for more complex and large-scale time series forecasting tasks. Deep learning approaches like RNN, LSTM, and TM are gaining popularity because they can handle complex, non-linear relationships and long-term dependencies in time series data more effectively than traditional methods like ARIMA. Additionally, they require less manual feature engineering, scale well with large datasets, and can adapt more easily to changes in data patterns.

RNNs update their hidden state based on the input at the current time step and the hidden state from the previous time step. The hidden state update and output calculation are presented in Eq. 2, respectively.

$$\begin{aligned} h_t &= \sigma_h(W_h x_t + U_h h_{t-1} + b_h) \\ y_t &= \sigma_y(W_y h_t + b_y) \end{aligned} \quad (2)$$

where h_t and x_t denote the hidden state and input at time step t , respectively, and h_{t-1} is the hidden state at the previous time step $t - 1$. Here W_h , U_h , b_h and σ_h represent the weight matrix for the input x_t , the weight matrix for the hidden state h_{t-1} , the bias term, and the activation function for the input layer, respectively, while y_t , W_y , b_y , and σ_y denote the output at time step t , the weight matrix for the output, the bias term for the output, and the activation function for the output layer, respectively.

The LSTM model can capture long-term dependencies in sequential data and is presented through five equations: the forget gate, input gate, cell state update, output gate and hidden state update. Among these five steps, the forget gate step, presented in Eq.3, is often considered the most significant.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (3)$$

where f_t and x_t represent the forget gate activation vector at time step t and the current input,

respectively; W_f and b_f denote the weight matrix and bias vector for the forget gate, respectively; σ is the sigmoid activation function; and $[h_{t-1}, x_t]$ denotes the concatenation of the previous hidden state h_{t-1} .

GRU networks simplify the LSTM architecture by combining the forget and input gates into a single update gate and merging the cell state and hidden state. The hidden state update equation is as presented in Eq.4.

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (4)$$

where h_t , z_t , and \tilde{h}_t represent the hidden state, the update gate, and the candidate hidden state at time step t , respectively; h_{t-1} is the hidden state at time at the previous time step $t - 1$, and \odot denotes matrix multiplication.

The TM is a type of neural network architecture that utilizes self-attention mechanisms to capture dependencies between different time steps in the data [3]. The key equation in the Transformer model is the self-attention mechanism, which computes the weighted sum of values based on the similarities with keys. This mechanism is as presented in Eq.5:

$$\text{output}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (5)$$

Q, K, V represent the query, key, and value matrices, respectively. d_k denotes the dimensionality of the key vectors.

2.2. Imputation Techniques

Information about the commonly used imputation techniques used in the study is briefly presented below:

- *Linear Imputation*: This method fills missing values with a linear interpolation between adjacent known data points. It assumes a linear relationship between consecutive observations [8].
- *Spline Imputation*: Spline interpolation is a non-linear method that fits a series of polynomial functions between data points to create a smooth curve, providing a smooth estimation of missing values [17].
- *Stineman Imputation*: Stineman imputation could theoretically refer to using the Stineman interpolation method to estimate missing values in a dataset. The Stineman interpolation method aims to preserve the data's monotonicity and convexity/concavity [18].
- *Mean Imputation*: In mean imputation, missing values are replaced with the mean of the available data. It is a simple and commonly used method but may not capture the underlying patterns in the data [19].

- *Random Imputation*: This technique fills missing values with randomly selected values from the observed data. While it is easy to implement, random imputation may introduce noise into the dataset and distort the original distribution [8].

Each imputation method has its advantages and limitations, and the choice depends on the nature of the data and the specific objectives of the analysis.

III. FINDINGS

3.1. Simulation

A synthetic dataset of 4,500 units is produced for the years 2006-2024. The stock price is modeled with a linear trend, starting from 50 Turkish Liras and reaching 150 Turkish Liras. Normally distributed random variables are used to create random noise in price changes, with the dataset representing end-of-day data and daily fluctuations in stock prices. The seasonal component of prices is modeled with a sine function [$f(\cdot) = \sin(\cdot)$] repeated at certain intervals. To increase the realism of the simulation, outliers with high volatility are added to randomly selected time periods to represent the impact of unexpected market events or news.

For time series analysis, deep learning approaches such as RNN, LSTM, GRU, and Transformer models are used. Various data imputation techniques, including linear, spline, Stineman, mean, and random imputation methods, are employed. The models are tested with 0% (complete data), 5%, 15%, and 25% missing data. The uniform distribution is used to reduce the data [20]. The Python programming language is used for the analysis, and the results are presented in Table 1.

For non-imputation (0%), TM performs the best in terms of the mean absolute percentage error (MAPE) and R^2 but has a higher the root mean square error (RMSE) compared to RNN, LSTM, and GRU. Linear, spline, and Stineman imputations maintain high accuracy with moderate increases in errors as imputation rates increase. Mean and random imputations result in significant errors and a decrease in prediction accuracy. TM generally provides the best accuracy (lowest MAPE and highest R^2), making it the preferred choice when imputation rates are low to moderate. GRU and TM perform better with higher imputation rates compared to LSTM and RNN. Linear, spline, and Stineman imputations are preferred for maintaining model performance across all estimators. Mean and random imputations result in significant errors and should be avoided if possible.

3.2. Analysis of stock market data

End-of-day data for the stock prices of Turkish Airlines (THYAO), Lufthansa (LHA.DE), and Delta Airlines (DAL) are analyzed. These three airlines, which are among the top 10 leading airlines in the world, are included in the study. The same stages applied in the

simulation are used in this analysis as well. Our goal is to compare the results from the synthetic data with those from real data applications [21]. The analysis of the stock prices of these three airline companies from 2006 to 2024 is presented in Tables 2-4 for THYAO, LHA.DE, and DAL, respectively.

For the THYAO dataset, TM and GRU exhibit the best performance with low MAPE and RMSE, and high R^2 , indicating accurate and reliable predictions without missing data under non-imputation (0%). Linear and spline imputations are more effective, with spline being slightly better. Stineman imputation is also effective, but mean and random imputations lead to poor performance. GRU and TM are the most robust and accurate, especially with the appropriate imputation methods. GRU and Transformer models also show the best overall performance, particularly when using spline and Stineman imputations.

For the LHA.DE dataset, GRU exhibits the best performance with the lowest RMSE and relatively low MAPE, indicating accurate and reliable predictions without missing data. TM also shows strong performance under non-imputation (0%) since it produces smallest MAPE. Linear and spline imputations are more effective, with spline being slightly better. Stineman imputation is also effective, but mean and random imputations lead to poor performance. GRU and Transformer models show the best overall performance, particularly when using spline and Stineman imputations.

For the DAL dataset, the results are similar to those obtained for other datasets when imputation is employed. GRU and TM outperform the other methods considered in this study when combined with the spline and Stineman imputation techniques, particularly as the imputation rate increases.

IV. CONCLUSION

The comparison of traditional and deep learning methods on both original and imputed data has not been extensively explored in the literature as far as we know. This study investigates the imputation of missing data and evaluates the performance of deep learning techniques, utilizing datasets from Turkish Airlines, Lufthansa, Delta Airlines, and simulated synthetic data. The study employs RNN, LSTM, GRU, and TM. The datasets are imputed to include 5%, 15%, and 25% missing data, which are then imputed using Linear, Spline, Stineman, Mean, and Random imputation methods. The analysis of the four tables for datasets THYAO, LHA.DE, DAL, and Synthetic datasets across various imputation techniques provides a comprehensive overview of model performance under different conditions.

Performance without missing data (0% imputation): Across all datasets, TM consistently shows the lowest MAPE and RMSE, indicating superior accuracy and reliability. GRU and LSTM models also perform well, with RNN exhibiting slightly higher error rates. The high R^2 values for all models across datasets indicate that the models explain a significant proportion of the variance in the data without any missing values.

Impact of imputation techniques: Linear and Spline Imputations generally result in moderate increases in MAPE and RMSE across all models and datasets, with spline imputation showing slightly better performance than linear imputation. RNN models exhibit a more significant increase in errors compared to LSTM, GRU, and Transformer models. Stineman imputation maintains low and stable MAPE and RMSE values, particularly for GRU and Transformer models, indicating effective handling of missing data. Mean and random imputations led to significant increases in MAPE and RMSE across all models and datasets, with random imputation resulting in the most drastic deterioration in performance. This indicates that these methods are ineffective for handling missing data in time series forecasting.

While performing adequately without missing data, RNNs struggle significantly with higher rates of imputation, particularly with mean and random methods. LSTM models demonstrate robust performance, handling linear and spline imputations well but showing moderate increases in errors with higher rates of mean and random imputations. GRU models consistently show the best overall performance across all imputation techniques and datasets, maintaining low error rates and high R^2 values. TM models excel particularly with spline and Stineman imputations, maintaining very low MAPE and RMSE values. However, they also exhibit increased sensitivity to mean and random imputations, similar to other models.

The comprehensive analysis of the THYAO, LHA.DE, DAL, and Synthetic datasets under various imputation scenarios reveals several key conclusions.

For model performance: TM and GRU models consistently outperform RNN and LSTM models, particularly in the presence of missing data. This is evident from their lower MAPE and RMSE values across different imputation techniques.

For effective imputation techniques: Spline and Stineman imputations are more effective in maintaining model performance, while mean and random imputations lead to significant performance degradation. This underscores the importance of selecting the appropriate right imputation technique in time series forecasting.

For robustness to missing data: GRU and TM exhibit greater robustness to missing data, maintaining accuracy and low error rates across different imputation methods and missing data rates.

For metric correlation: The correlation between MAPE and RMSE suggests that both metrics should be considered for a comprehensive evaluation of model performance. Although they generally align, their different sensitivities to outliers provide complementary insights into model accuracy.

For practitioners in the field of time series forecasting, GRU and Transformer models are recommended due to their robustness and superior performance. When dealing with missing data, spline and Stineman imputations should be preferred to ensure minimal degradation in model accuracy. Avoid using mean and random imputations, as they significantly impair model performance. Finally, always evaluate model performance using both MAPE and RMSE to capture a complete picture of the model's predictive accuracy and reliability.

Future research can further expand on the findings of this study by examining additional deep learning architectures and prediction techniques to assess their effectiveness in handling missing data. Extending this to include not only time series prediction but also other types of data, such as health and environmental science datasets, would increase the generalizability of the results.

ACKNOWLEDGMENTS

This study was supported by Eskişehir Technical University Scientific Research Project Commission under grant no: 23ADP172. This study was produced from Kürşat Atmaca's Master's Thesis, which was supervised by Dr. İsmail Yenilmez. An earlier version of this study was presented at UYİK-2024 (For details see [21]).

DISCLAIMER: The data utilized in this study are publicly accessible real-world datasets. This research focused on evaluating the performance of various deep learning methods and imputation techniques using these datasets. The study does not provide any commentary or recommendations regarding the buying, selling, or other actions related to companies' stocks. Therefore, we bear no responsibility for such actions. As the data used are publicly available, no permissions were required for this study.

REFERENCES

- [1] Chatfield, C. (2004). *The Analysis of Time Series: An Introduction*. Chapman and Hall/CRC.
- [2] Yenilmez, I., & Kantar, Y.M., 2019. An Analysis of Export Data with Panel Tobit Model. *ICONDATA19*, e-ISBN: 978-605-031-662-9, pp.92-97.
- [3] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All You Need. *Advances in Neural Information Processing Systems*, 30, 5998-6008.
- [4] Çalık, A. S., Yenilmez, İ., & Sevinçtekin, Ö. (2022). Comparison of Box-Jenkins and Artificial Neural Networks methods: Demand forecast for ceramic sanitary ware. *Proceeding Book of EGE 7th International Conference on Applied Sciences*, December 24-25, İzmir, Türkiye (pp.1024-29), ISBN: 978-605-72197-9-4.
- [5] Mugenzi, F. & Yenilmez, İ. (2023). Forecasting for GDP Per Capita Using Multiplier Perceptron and Gated Recurrent Unit. *Proceeding Book of Akdeniz 10th International Conference on Applied Sciences*, November 2-5, KYRENIA (pp.318-327), ISBN: 978-625-6830-49-3.
- [6] Yenilmez, I., & Mugenzi, F. (2023). Estimation of conventional and innovative models for Rwanda's GDP per capita: A comparative analysis of artificial neural networks and Box-Jenkins methodologies. *Scientific African*, 22, e01902. <https://doi.org/10.1016/j.sciaf.2023.e01902>.
- [7] Yenilmez, İ. & Akçay, E. K. (2023). Performance of Particle Swarm Optimization and Genetic Algorithm for Tuning of k-NN Hyperparameters. *Proceeding Book of Akdeniz 10th International Conference on Applied Sciences*, November 2-5, KYRENIA (pp.328-338), ISBN: 978-625-6830-49-3
- [8] Little, R. J. A., & Rubin, D. B. (2019). *Statistical Analysis with Missing Data*. John Wiley & Sons.
- [9] Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. Chapman and Hall/CRC.
- [10] Yamak, P. T., Yujian, L., & Gadosey, P. K. (2019). Comparison of ARIMA, LSTM, and GRU models for time series forecasting: Evidence from Bitcoin price data. *Journal of Financial Data Science*, 1(1), 45-60.
- [11] Ridwan, M., Sadik, K., & Afendi, F. M. (2023). Evaluating the effectiveness of ARIMA and GRU models in high-frequency stock price forecasting: A case study of HIMBARA bank stocks. *Journal of Financial Forecasting*, 5(2), 87-102.
- [12] Kontopoulou, V. I., Panagopoulos, A. D., Kakkos, I., & Matsopoulos, G. K. (2023). A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks. *Future Internet*, 15(8), 255. <https://doi.org/10.3390/fi15080255>.
- [13] Ahmed, S., Nielsen, I.E., Tripathi, A. et al. Transformers in Time-Series Analysis: A Tutorial. *Circuits Syst Signal Process* 42, 7433–7466 (2023). <https://doi.org/10.1007/s00034-023-02454-8>.
- [14] Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., & Sun, L. (2022). Transformers in Time Series: A Survey.

-
- [15] Fang, C., & Wang, C. (2020). Time Series Data Imputation: A Survey on Deep Learning Approaches. *ArXiv*
<https://arxiv.org/abs/2011.11347>
- [16] Yenilmez, İ. (2024). Imputation methods effect on the goodness of fit of the statistical model. In *Proceedings of the 9th International Conference on Business, Management and Economics*. Vienna, Austria. ISBN 978-609-485-514-6.
- [17] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- [18] Stineman, R. W. (1980). A consistently well-behaved method of interpolation. *Creative Computing*, 6(7), 54-57.
- [19] Donders, A. R. T., van der Heijden, G. J., Stijnen, T., & Moons, K. G. (2006). Review: A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology*, 59(10), 1087-1091.
<https://doi.org/10.1016/j.jclinepi.2006.01.014>
- [20] Yenilmez, İ., & Kantar, Y. M. (2023). New exponentiated generalized censored regression models: Monte Carlo simulation and application. *Concurrency and Computation: Practice and Experience*, 35(1), e7436.
<https://doi.org/10.1002/cpe.7436>
- [21] Atmaca, K., & Yenilmez, İ. (2024). RNNs and Transformer Model in case of Incomplete Time Series. *Conference paper presented at the Fifth International Congress of Applied Statistics (UYIK-2024)*, İstanbul, Türkiye

APPENDICES

Table 1. Results for synthetic dataset

SYNTHETIC		Lin. Imp.				Spl. Imp.			Sti. Imp.			Mea. Imp.			Ran. Imp.		
		0%	5%	15%	25%	5%	15%	25%	5%	15%	25%	5%	15%	25%	5%	15%	25%
RNN	MAPE	0.90	1.24	1.27	1.60	1.12	1.16	1.84	1.01	1.10	1.63	4.24	5.09	13.62	5.83	11.92	28.67
	R²	0.98	0.99	0.99	0.99	0.99	0.98	0.99	0.99	0.99	0.99	0.81	0.77	0.26	0.70	0.32	0.24
	RMSE	0.57	1.61	1.73	2.02	1.45	1.51	2.40	1.31	1.45	2.09	9.24	10.19	19.95	12.40	22.43	33.92
LSTM	MAPE	0.85	1.79	1.83	1.91	1.60	1.64	1.66	1.42	1.52	1.81	4.65	3.59	10.13	4.29	10.15	12.12
	R²	0.98	0.99	0.99	0.98	0.98	0.98	0.98	0.99	0.99	0.98	0.82	0.83	0.40	0.73	0.35	0.22
	RMSE	0.53	2.30	2.33	2.42	2.10	2.12	2.13	1.91	1.97	2.32	8.91	8.71	17.99	11.65	21.95	26.72
GRU	MAPE	0.86	0.91	0.94	0.99	0.90	0.91	1.10	0.88	0.90	0.96	3.74	3.55	13.12	4.31	11.08	14.91
	R²	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.82	0.83	0.39	0.73	0.33	0.23
	RMSE	0.53	1.19	1.24	1.31	1.15	1.18	1.45	1.15	1.17	1.23	8.82	8.73	18.11	11.71	22.21	26.53
TM	MAPE	0.15	0.17	0.21	0.27	0.19	0.20	0.22	0.18	0.19	0.21	8.90	3.13	10.66	4.10	10.90	16.10
	R²	0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.99	0.99	0.99	0.50	0.60	0.45	0.41	0.38	0.22
	RMSE	1.76	1.96	2.13	2.50	1.96	2.01	2.48	1.98	2.07	2.13	6.10	2.60	6.90	3.60	8.60	10.14

Table 2. Results for Turkish Airlines stocks' dataset

THYAO		Lin. Imp.				Spl. Imp.			Sti. Imp.			Mea. Imp.			Ran. Imp.		
		0%	5%	15%	25%	5%	15%	25%	5%	15%	25%	5%	15%	25%	5%	15%	25%
RNN	MAPE	2.71	7.18	8.24	10.09	6.11	7.85	9.52	6.35	7.86	12.55	24.30	52.82	34.52	46.83	96.7	176.24
	R²	0.99	0.94	0.90	0.87	0.95	0.95	0.92	0.94	0.93	0.81	0.51	0.28	0.16	0.43	0.31	0.11
	RMSE	3.64	17.41	21.38	25.74	14.04	15.48	17.91	17.01	17.72	31.37	48.96	57.53	59.13	89.17	87.53	84.14
LSTM	MAPE	3.24	4.10	4.40	4.44	2.86	3.41	3.95	2.92	3.47	4.60	24.61	33.54	42.84	49.74	81.76	122.00
	R²	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.69	0.70	0.51	0.35	0.26	0.22
	RMSE	3.77	4.74	6.09	6.78	3.35	3.98	4.02	5.18	6.23	5.80	38.60	36.79	45.17	85.57	86.04	88.21
GRU	MAPE	2.05	2.52	2.75	2.94	2.45	2.55	2.72	2.65	2.82	2.98	24.42	39.76	48.62	36.46	79.11	130.58
	R²	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.89	0.72	0.66	0.10	0.24	0.21
	RMSE	2.71	3.12	3.19	3.36	3.08	3.60	3.82	4.35	4.62	4.78	23.21	35.60	37.57	75.27	85.09	87.71
TM	MAPE	0.41	0.68	1.08	1.43	0.46	0.47	0.54	0.44	0.50	0.58	25.47	49.42	57.53	94.03	122.60	179.70
	R²	0.99	0.99	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.86	0.66	0.45	0.51	0.22	0.22
	RMSE	2.25	2.43	2.58	4.38	2.41	2.47	3.34	2.45	2.58	3.88	28.76	43.65	54.36	55.59	84.75	85.42

Table 3. Results for Deutsche Lufthansa AG stocks' dataset

LHA.DE		Lin. Imp.				Spl. Imp.			Sti. Imp.			Mea. Imp.			Ran. Imp.		
		0%	5%	15%	25%	5%	15%	25%	5%	15%	25%	5%	15%	25%	5%	15%	25%
RNN	MAPE	1.91	1.93	2.03	2.99	1.93	2.01	2.08	1.9	2.03	2.04	6.12	8.18	19.95	18.38	16.02	26.04
	R²	0.99	0.99	0.99	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.91	0.78	0.53	0.59	0.43	0.25
	RMSE	0.21	0.22	0.23	0.31	0.22	0.23	0.24	0.21	0.22	0.22	0.82	1.25	1.77	2.01	2.64	3.47
LSTM	MAPE	1.72	2.03	1.94	4.18	1.85	1.90	2.03	1.88	1.90	1.94	6.20	9.52	12.38	9.31	15.23	25.07
	R²	0.99	0.99	0.99	0.97	0.99	0.99	0.99	0.99	0.99	0.99	0.91	0.80	0.69	0.69	0.45	0.25
	RMSE	0.20	0.23	0.22	0.46	0.21	0.21	0.23	0.21	0.21	0.22	0.82	1.20	1.44	1.76	2.60	3.46
GRU	MAPE	1.65	1.88	1.89	2.38	1.80	1.78	1.86	1.78	1.86	1.87	5.57	9.84	11.95	9.49	12.96	22.28
	R²	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.91	0.80	0.68	0.69	0.45	0.29
	RMSE	0.19	0.20	0.21	0.27	0.20	0.21	0.22	0.21	0.21	0.21	0.83	1.20	1.46	1.75	2.59	3.38
TM	MAPE	0.38	0.39	0.44	0.51	0.39	0.40	0.45	0.48	0.45	0.50	9.30	8.71	10.88	7.00	10.85	17.97
	R²	0.99	0.99	0.99	0.91	0.99	0.99	0.97	0.97	0.96	0.97	0.42	0.53	0.34	0.47	0.23	0.22
	RMSE	0.34	0.35	0.43	0.72	0.35	0.39	0.62	0.53	0.47	0.53	0.92	0.83	0.99	0.99	1.51	85.43

Table 4. Results for Delta Airlines stocks' dataset

DAL		Lin. Imp.				Spl. Imp.			Sti. Imp.			Mea. Imp.			Ran. Imp.		
		0%	5%	15%	25%	5%	15%	25%	5%	15%	25%	5%	15%	25%	5%	15%	25%
RNN	MAPE	1.86	2.07	2.65	2.94	2.01	2.08	2.63	2.05	2.10	2.60	8.66	13.23	16.21	13.13	22.88	36.33
	R²	0.99	0.98	0.97	0.97	0.99	0.99	0.96	0.99	0.98	0.97	0.78	0.54	0.36	0.68	0.38	0.22
	RMSE	1.01	1.10	1.43	1.49	1.07	1.10	1.76	1.09	1.13	1.35	4.59	6.55	7.52	5.94	9.01	11.11
LSTM	MAPE	1.82	2.01	2.27	4.16	1.95	2.01	2.93	1.98	2.03	3.87	7.06	11.15	13.38	13.64	25.88	35.13
	R²	0.99	0.98	0.98	0.95	0.99	0.99	0.94	0.99	0.98	0.95	0.81	0.59	0.48	0.67	0.35	0.21
	RMSE	0.99	1.09	1.18	2.19	1.06	1.09	1.95	1.07	1.09	2.03	4.23	6.13	6.79	5.99	9.29	11.10
GRU	MAPE	1.78	1.96	2.16	2.24	1.92	2.00	2.73	1.95	1.98	2.36	6.41	13.41	17.68	14.31	23.59	35.20
	R²	0.99	0.99	0.99	0.98	0.99	0.99	0.97	0.99	0.99	0.99	0.92	0.77	0.66	0.82	0.61	0.46
	RMSE	0.99	1.09	1.14	1.17	1.05	1.07	1.49	1.06	1.08	0.62	2.24	3.38	3.96	3.28	4.91	5.96
TM	MAPE	0.37	0.44	0.47	0.48	0.39	0.4	0.44	0.41	0.44	0.46	7.83	13.59	18.32	9.06	20.49	29.94
	R²	0.99	0.98	0.98	0.97	0.99	0.98	0.97	0.97	0.97	0.97	0.92	0.79	0.67	0.89	0.71	0.50
	RMSE	0.98	1.28	1.44	1.47	1.16	1.21	1.41	1.24	1.27	1.44	2.30	3.67	4.34	2.76	4.52	5.75

Mekânsal Nokta Verilerinin Mekânsal İstatistiklerle Analizi: Yenilenebilir Enerji Kaynakları Uygulaması

Analysis of Spatial Point Data with Spatial Statistics: Renewable Energy Resources Application

Mehmet Burak BARTAN¹, Günseli AYTAÇ CANKURTARAN¹, Özlem TÜRKŞEN²

¹ Ankara Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik ABD, 06100, Ankara, Türkiye

² Ankara Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 06100, Ankara, Türkiye

Öz

Araştırma konusuna yönelik olarak elde edilen verinin konum bilgisi içermesi durumunda veri, mekânsal veri olarak adlandırılır. Nokta ve raster biçiminde tanımlanabilen mekânsal verilerin analizinde mekânsal istatistikler kullanılır. Savunma ve bilişim teknolojileri, sağlık, çevre bilimi, yer bilimi, enerji, şehir ve bölge planlama gibi alanlarda mekânsal verilerle sıklıkla karşılaşılır. Bu çalışmada, Türkiye’de yer alan rüzgâr türbinleri ve güneş panellerine ilişkin nokta verilerin mekânsal istatistiklerle analizi yapılmıştır. Enerji Piyasası Düzenleme Kurumu (EPDK) tarafından lisans verilen işletmelerden rüzgâr türbinleri için Çanakkale ili verileri ve güneş panellerine ilişkin ise ülkemizde lisans verilen tüm nokta verileri dikkate alınmıştır. Nokta verilere ilişkin öznitelik için kurulu güç değerleri ile ilgilenilmiştir. Mekânsal Betimsel İstatistikler, Mekânsal Örüntü Analizi, Mekânsal Otokorelasyon Analizi ve Mekânsal Enterpolasyon başlıklarında RStudio programı kullanılarak mekânsal istatistikler elde edilmiştir. Bu mekânsal istatistiklere göre, Türkiye’de yer alan rüzgâr enerji santrali (RES) ve güneş enerji santrali (GES) kurulu güç değerlerinin mekânsal dağılımının rastgele olmadığı görülmüştür. RES’ler ve GES’ler için 10-kat çapraz doğrulama ile Kriging Analizi uygulanarak nokta verilerin kurulu güç değerleri için öngörüler yapılmıştır.

Anahtar Kelimeler: Mekânsal Nokta Veri, Mekânsal İstatistiksel Analizler, Yenilenebilir Enerji

Abstract

Data is referred to as spatial data if it contains location information relevant to the research issue. The analysis of spatial data, which can be described in raster and point forms, uses spatial statistics. Spatial data is frequently encountered in areas such as defense and information technologies, health, environmental science, earth science, energy, urban and regional planning. In this study, spatial statistics were used to assess point data pertaining to solar panels and wind turbines in Turkey. All point data licensed in our country for solar panels and data from the province of Çanakkale were considered for wind turbines licensed by the Energy Market Regulatory Authority (EMRA). Installed power values were of interest for the attribute related to point data. Spatial statistics were obtained using the RStudio program under the titles of Spatial Descriptive Statistics, Spatial Pattern Analysis, Spatial Autocorrelation Analysis and Spatial Interpolation. According to these spatial statistics, it was observed that the spatial distribution of the installed power of wind power plant (WPP) and solar power plant (SPP) in Turkey was not random. Installed power predictions of point data were made by applying Kriging Analysis with 10-fold cross validation for the WPPs and the SPPs.

Keywords: Spatial Point Data, Spatial Statistical Analysis, Renewable Energy

I. GİRİŞ

Mekânsal kelimesi, verideki her bir birim, olay veya durumun ilgili harita üzerinde nerede meydana geldiğinin bilinmesini sağlayan coğrafi bir referansa sahip olduğu anlamına gelmektedir [1]. Mekânsal veri, mekânsal nesnelere ilişkin öznitelik değerleriyle birlikte nesnelere konum bilgilerinin de yer aldığı veridir. Öznitelikler farklı biçimlerde ele alınabilir. Bazıları doğası gereği fiziksel veya çevresel iken bazıları ise sosyal veya ekonomik niteliklerdir. Mekânsal veriler, özniteliklerin atıfta bulunduğu mekânsal nesne türüne (nokta, çizgi, alan) ve bu

özniteliklerin ölçüm düzeyine göre sınıflandırılır. Z_1, Z_2, \dots, Z_p 'nin p rastgele değişkeni ve S 'nin nokta konumunu ifade ettiğini varsayılır. Buna göre mekânsal veri genel olarak matris biçiminde Şekil 1'deki gibi temsil edilebilir [2].

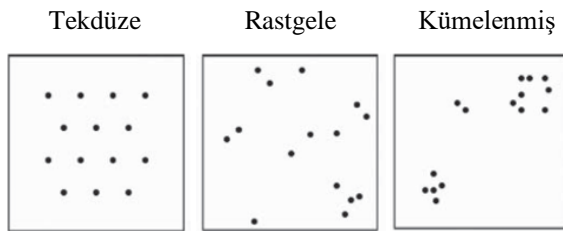
$$\begin{array}{c} \text{Öznitelikler} \\ Z_1 \quad Z_2 \quad \dots \quad Z_p \quad S \\ \left[\begin{array}{ccccc} z_1(1) & z_2(1) & \dots & z_p(1) & s(1) \\ z_1(2) & z_2(2) & \dots & z_p(2) & s(2) \\ \vdots & \vdots & & \vdots & \vdots \\ z_1(n) & z_2(n) & \dots & z_p(n) & s(n) \end{array} \right] \begin{array}{l} \text{Durum 1} \\ \text{Durum 2} \\ \vdots \\ \text{Durum } n \end{array} \end{array}$$

Şekil 1. Mekânsal veri matrisi

Mekânsal verinin aldığı değerlerin gösterimi, $\{z_1(i), z_2(i), \dots, z_p(i) | s(i)\}_{i=1,2,\dots,n}$ biçiminde yapılabilir. Burada, z_k , değişken Z_k ($k = 1, 2, \dots, p$)'nin bir gerçekleşmesini belirtirken, parantez içindeki sembol i , belirli bir durumu ifade eder. Her bir i durumuna $i = 1, 2, \dots, n$, mekânsal nesnenin konumunu temsil eden bir $s(i)$ konumu eklenir.

Mekânsal veri analizi, üzerinde çalışılan sistem içinde her bir nesne veya veri değeri ile ilişkili olan mekânsal referansları kullanan modeller ve yöntemler topluluğudur. Olaylar arasındaki mekânsal ilişkiler veya mekânsal etkileşimleri tanımlamak için mekânsal veri analizi yöntemlerinde verinin görsel olarak gösterimine ihtiyaç duyulmaktadır.

Mekânsal istatistiğin temel amacı, mekânsal dağılıma ilişkin betimsel istatistiklerin üretilmesi, hipotezlerin geliştirilmesi ve bu hipotezlerin test edilmesidir [3]. Mekânsal istatistikler kullanılarak veri setlerinin dağılım özellikleri belirlenir. Bunlar, tekdüze, rastgele ve kümelenmiş olmak üzere üç kategoride adlandırılır [4]. Mekânsal nokta verilerin örüntüleri Şekil 2'de görülmektedir.

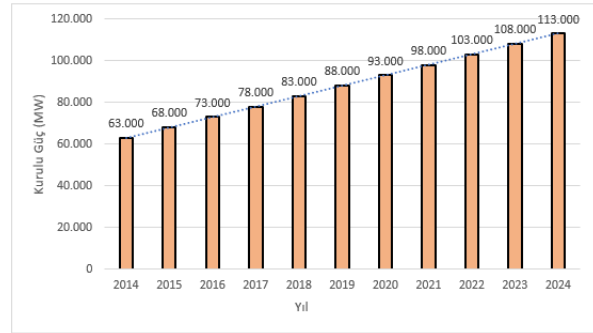


Şekil 2. Mekânsal nokta verilerin örüntü örnekleri [4]

Gelişen teknolojiye paralel olarak ortaya çıkan gereksinimlerin her geçen gün artmasından dolayı çoğalarak artan elektrik enerjisi tüketimi, üreticileri ve kullanıcıları yeni ve çevreye uyumlu enerji kaynakları aramaya zorlamaktadır [5]. Yenilenemeyen ya da yeraltı enerji kaynakları olarak da adlandırılan fosil enerji kaynakları olan kömür, petrol gibi enerji kaynaklarının çevreye olumsuz etkilerinin olması ve tükenebilir olması nedeniyle alternatif enerji

kaynaklarına yönelim hızla artmıştır. Güneş, rüzgâr, biyokütle, hidro, gel-git ve dalga enerjisi şeklinde sıralanabilen yenilenebilir enerji kaynakları, karbondioksit emisyonunu azaltarak doğanın korunmasına yardımcı olduğu gibi enerjide dışa bağımlılığı azaltmakta ve aynı zamanda istihdamın artmasına da katkı sağlamaktadır [6].

Yenilenebilir enerji kaynakları yatırımları, başta arz güvenliği olmak üzere, iklim değişikliği, çevresel faktörler ve sürdürülebilirlik bakımından artış göstermektedir. Ülkemizde son yıllarda artan nüfus ve artan sanayi üretimi ile birlikte enerji ihtiyacı da artmıştır. Yenilenebilir enerji kaynaklarından rüzgâr ve güneş enerjisi özellikle elektrik enerjisinin karşılanması konusunda öncü iki enerji kaynağı olarak karşımıza çıkmaktadır. Elektrik enerjisinin kurulu gücü son on yılda neredeyse iki katına çıkmıştır (Şekil 3).



Şekil 3. 2014-2024 yılları arasında Türkiye elektrik enerjisi kurulu güç değerleri [7]

Tablo 1. Yıllara göre Türkiye'nin elektrik üretiminde kullanılan yenilenebilir enerji kaynakları oranı [8]

Yıl	Hidrolik (%)	Doğal Gaz (%)	Kömür (%)	Rüzgâr (%)	Güneş (%)	Jeotermal (%)	Diğer (%)
2014	30,5	28,9	29,3	5,1	2,6	1,3	2,3
2015	31,0	28,5	28,0	5,5	3,0	1,5	2,5
2016	31,5	28,0	27,0	6,0	3,5	1,7	2,3
2017	32,0	27,5	26,0	6,5	4,0	1,8	2,2
2018	32,5	27,0	25,0	7,0	4,5	2,0	2,0
2019	33,0	26,5	24,0	7,5	5,0	2,2	2,0
2020	33,5	26,0	23,0	8,0	5,5	2,4	1,6
2021	34,0	25,5	22,0	8,5	6,0	2,6	1,4
2022	34,5	25,0	21,0	9,0	6,5	2,8	1,2
2023	19,6	21,4	36,3	10,4	5,7	3,4	3,2
2024	29,6	23,2	20,2	11,2	11,7	1,6	2,5

Elektrik enerjisi ihtiyacını karşılamak amacıyla birçok yatırım yapılmıştır. Yapılan yatırımların çoğunluğunun yenilenebilir enerji kaynaklarına yöneldiği görülmektedir (Tablo 1). Bu durumun başlıca nedeni; ülkemizdeki yenilenemeyen enerji kaynaklarının tamamının ithal edilmesidir. Bunun sonucunda, elektrik enerjisinde dışa bağımlılık ve maliyetler artmaktadır. Yenilenebilir enerji kaynakları arasında rüzgâr ve güneş enerjisi, temiz, tükenmez ve maliyet bakımından rekabetçi bir güç kaynağı sağlama potansiyeliyle öne çıkmaktadır. Özellikle düşük karbon

ayak izi ve diğer enerji kaynaklarıyla karşılaştırıldığında, emisyonları hızlı bir biçimde dengeleme yeteneği nedeniyle de son yıllarda yatırım miktarları oldukça artmıştır.

Bu çalışmada, yenilenebilir enerji kaynaklarından rüzgâr ve güneş enerjileri ile ilgilenilmiştir. Rüzgâr Enerji Santralleri (RES) ve Güneş Enerji Santralleri (GES) için EPDK'dan elde edilen mekânsal nokta veriler analiz edilmiştir. Mekânsal analizde mekânsal istatistiksel yöntemler uygulanmıştır. Çalışmanın ikinci bölümünde, mekânsal betimsel istatistikler, mekânsal örüntü analizleri, mekânsal global okorelasyon ve mekânsal enterpolasyon konuları hakkında açıklayıcı bilgiler verilmiştir. Çalışmanın üçüncü bölümünde, RES'ler ve GES'ler için iki ayrı uygulama yapılmıştır. Elde edilen sonuçlar ayrıntılı biçimde sunulmuştur.

II. MATERYAL VE YÖNTEM

Mekânsal istatistik yöntemlerinin hemen hemen tamamının çıkış noktası nokta ile ifade edilebilen gözlemlerin analizidir. Nokta veri için mekânsal istatistik, objelerin ya da olayların mekânsal konum bilgisine sahip nokta ile ifade edilebildiği ortamlarda, bu noktaların geometrik yapılarını ve desenlerini analiz ederek mekânsal düzenler ve ilişkiler hakkında derinlemesine bilgiler sağlamaktadır. Bu yöntemler, karar vericilerin ve araştırmacıların, mekânsal verileri anlamalarını ve etkili çözümler geliştirmelerini sağlamaktadır.

2.1. Mekânsal Betimsel İstatistikler

Mekânsal nokta veride, nokta dağılımının merkezi eğiliminin tespit edilmesi ve dağılımın orta noktasının bulunması son derece önemlidir. Mekânsal nokta verinin merkez noktasının tespiti için kullanılan birden fazla ölçüt bulunmaktadır [3]. Bunlar, ortalama merkez, ağırlıklı ortalama merkez ve ortanca merkez değerleridir.

2.1.1. Ortalama merkez

Mekânsal istatistikte ortalama merkez (mean center), olayların mekânsal dağılımlarının merkez noktasını göstermektedir [9]. Ortalama merkezin hesaplanmasında örneklemdaki elemanların koordinatları kullanılır. Ortalama merkez

$$(\bar{x}_{mc}, \bar{y}_{mc}) = \left(\frac{\sum_{i=1}^n x_i}{n}, \frac{\sum_{i=1}^n y_i}{n} \right) \quad (1)$$

biçiminde elde edilir. Burada, $(\bar{x}_{mc}, \bar{y}_{mc})$ ortalama merkezin koordinatları, (x_i, y_i) , $i = 1, 2, \dots, n$ gözleme ait koordinatları olup n ise örneklem büyüklüğünü göstermektedir.

2.1.2. Ağırlıklı ortalama merkez

Mekânsal istatistikte ağırlıklı ortalama merkez (weighted average center), ortalama merkezin bir değişken ile ilişkilendirilerek ağırlıklandırılmasıyla elde edilir [3]. Ağırlıklı ortalama merkez

$$(\bar{x}_{wmc}, \bar{y}_{wmc}) = \left(\frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}, \frac{\sum_{i=1}^n w_i y_i}{\sum_{i=1}^n w_i} \right) \quad (2)$$

olarak tanımlanır. Burada, $(\bar{x}_{wmc}, \bar{y}_{wmc})$ ağırlıklı ortalama merkez koordinatları (x_i, y_i) , $i = 1, 2, \dots, n$ gözleme ait koordinatlar, w_i , $i = 1, 2, \dots, n$ sayılı gözleme ait ağırlığı olup n ise örneklem büyüklüğünü göstermektedir.

2.1.3. Ortanca merkez

Mekânsal istatistikte ortanca merkez, istatistikteki ortancaya karşılık gelmektedir. Ortanca merkez noktası (median center) (x_{mdc}, y_{mdc}) , diğer tüm noktalara olan uzaklıkların toplamının minimum olduğu noktada konumlanmaktadır.

Literatürde ortanca merkezin diğer noktalara olan uzaklıklarının ölçülmesinde farklı metrikler kullanılmıştır. Bu metriklerin en yaygınları Öklid ve Manhattan uzaklıklarıdır. Öklid uzaklığında, ortanca merkez koordinatlara olan x -ekseni ve y -ekseni uzaklıklarının mutlak değeri esas alınmakta ve $\sum_{i=1}^n \sqrt{(x_j - x_{mdc})^2 + (y_i - y_{mdc})^2}$ ifadesi ile minimize edilmektedir. Manhattan uzaklığında, ortanca merkez koordinatlarına olan doğrusal uzaklıkların toplamı esas alınmakta ve $\sum_{i=1}^n |x_i - x_{mdc}| + |y_i - y_{mdc}|$ ifadesi ile minimize edilmektedir [3].

2.1.4. Standart uzaklık

Standart uzaklık gözlemlerin mekânsal dağılımının ortalama merkezden ne kadar uzaklaştığını gösteren bir ölçüttür. Standart uzaklık

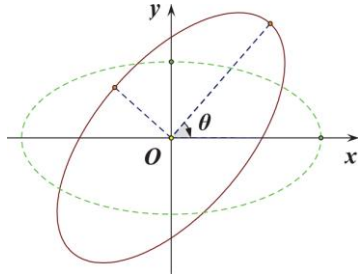
$$s_D = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x}_{mc})^2 + \sum_{i=1}^n (y_i - \bar{y}_{mc})^2}{n}} \quad (3)$$

olarak tanımlanmaktadır. Burada, s_D standart uzaklık, $(\bar{x}_{mc}, \bar{y}_{mc})$ ortalama merkez, (x_i, y_i) , $i = 1, 2, \dots, n$ sayılı gözleme ait koordinatlar olup n ise örneklem büyüklüğüdür [3].

2.1.5. Standart sapma elipsi

Standart sapma elipsi, standart uzaklık ile birlikte mekânsal istatistikte gözlemlerin mekânsal dağılımını betimlemek için kullanılan en yaygın ölçütlerden biridir. Bu elips, gözlemlere ait ortalama merkez çevresinde çizilir [3].

Mekânsal veri için standart sapma elipsinin tespitinde, x -ekseni ile y -ekseninden sapmalar ve y -ekseninden sapma açısı (θ) hesaplanmaktadır. Standart sapma elipsi, Şekil 4'te görülmektedir [10].



Şekil 4. Standart sapma açısı [10]

Şekil 4'te görülen θ açısının pozitif olması elipsin kuzeydoğu-güneybatı doğrultusunda, negatif olması ise elipsin kuzeybatı-güneydoğu doğrultusunda olduğu anlamına gelmektedir [11].

2.2. Mekânsal Örüntü Analizi

Mekânsal istatistikte en önemli konulardan biri, mekânsal dağılımın veya kümeleşmenin ölçülmesidir. Bu yöntemlerin amacı, nokta verilerin deseninin veya dağılımının tekdüze, rastgele veya kümelenmiş olup olmadığını belirlemektir.

Nokta verilerin örüntü analizinde kullanılan En Yakın Komşuluk Analizi ve Kuadrat Analizi yöntemleri, gözlemlenen nokta verinin mekânsal örüntüsünü ya da dağılımını analiz etmek için oldukça güçlü yöntemlerdir.

2.2.1. En Yakın Komşuluk Analizi

En Yakın Komşuluk Analizi, her birimin kendine en yakın birime olan uzaklık üzerine kurgulanmış bir yöntemdir. Öznitelik değerlerinden bağımsızdır ve sadece mekânsal olarak mesafeleri dikkate almaktadır.

En Yakın Komşuluk Analizi'nde gözlemlenen nokta verilerin mekânsal dağılımının rastgele olup olmadığı test edilir. Her gözlem için kendisine en yakın gözlem tespit edilerek tüm gözlemler için hesaplanan en kısa mesafelerin ortalaması alınarak gözlemlenen en yakın noktaya ortalama uzaklığı, r_{obs} [3]

$$r_{obs} = \frac{\sum_{i=1}^n d_i}{n} \quad (4)$$

olarak bulunur. Burada, d_i , $i = 1, 2, \dots, n$ noktasının kendisine en yakın uzaklıktaki noktaya olan doğrusal uzaklığı ve n , toplam nokta sayısını göstermektedir.

Beklenen en yakın komşuya uzaklık değeri ise, r_{exp}

$$r_{exp} = \frac{1}{2\sqrt{n/A}} \quad (5)$$

olarak hesaplanır. Burada, A , toplam çalışma alanı büyüklüğüdür.

r_{obs} , r_{exp} değerine bölünerek R -istatistiği

$$R - \text{istatistiği} = \frac{r_{obs}}{r_{exp}} \quad (6)$$

biçiminde elde edilir. Buna göre, $R - \text{istatistiği} > 1$ ise mekânsal dağılım tekdüze, $R - \text{istatistiği} < 1$ ise mekânsal dağılım kümelenme eğiliminde, $R - \text{istatistiği} \approx 1$ ise mekânsal veri, rastgele dağılımlıdır yorumu yapılır.

2.2.2. Kuadrat Analizi

Kuadrat Analizi nokta verinin mekânsal dağılımının incelenmesi amacı ile geliştirilmiş bir yöntemdir. Çalışma alanının eşit büyüklükteki alanlara (kuadrat) bölünmesine ve kuadrat içinde bulunan nokta sayısının dağılımının teorik ya da kurgusal dağılımlar ile karşılaştırılmasına dayanmaktadır [3].

Kuadrat Analizi'nde temel amaç, nokta verilerin mekânsal dağılımının tekdüze, rastgele veya kümelenmiş olup olmadığını belirlemektir. Kuadrat Analizi'nin ilk aşamasında, çalışma alanı eş büyüklükteki alanlara bölünür. Kuadrların kare ya da dikdörtgen olması durumunda tek bir kuadratin alanı

$$K = \frac{2A}{n} \quad (7)$$

formülü ile belirlenir. Burada, K , kuadrat alanı, A , toplam çalışma alanı büyüklüğünü, n , nokta gözlem sayısını ifade etmektedir.

Kuadrat formu kareye yakın ise karenin bir kenarı

$$k = \sqrt{K} = \sqrt{\frac{2A}{n}} \quad (8)$$

olarak bulunur.

Kuadrat Analizi'nin ikinci aşamasında, kuadrların içerisine düşen noktaların sayıları tespit edilerek frekans dağılım tablosu hazırlanır ve dağılım Kolmogorov-Simirnov ($K-S$) testi aracılığıyla hipotetik veya kurgusal dağılımlarla karşılaştırılır. Bunun için $K-S$ istatistiği

$$K - S = \sqrt{\frac{m_o \times m_H}{m_o + m_H}} \times \max |P_o - P_H| \quad (9)$$

biçiminde hesaplanır. Burada, m_o , gözlemlenen veriye ait kuadrat sayısını, m_H , hipotetik gözlemlere ait kuadrat sayısını, P_o , gözlemlenen veri kümülatif frekansını ve P_H , hipotetik veri kümülatif frekansını göstermektedir.

$K-S$ istatistiğinin, istatistiksel anlamlılığı için kurulan hipotez testinde incelenen iki mekânsal dağılımın aynı ya da farklı kitleden seçilip seçilmediği test edilir.

2.3. Global Mekânsal Otokorelasyon

Nokta verinin örüntü analizinde kullanılan yöntemlerde, gözlemlenen nokta verilerin sadece konumları dikkate alınmakta fakat, noktalara ait öznelik değerleri dikkate alınmamaktadır. Global mekânsal otokorelasyon ise, her mekânsal nokta verinin öznelik değerinin olması durumunda, gözlemler arasındaki ilişkiyi ortaya koymaktadır. Moran (1948) tarafından geliştirilen Moran'ın I indeksi, değişken değerlerine sahip nokta verinin dağılımının analizinde kullanılan en yaygın global mekânsal otokorelasyon yöntemidir [12]. Moran'ın I indeksinde gözlemlenen verinin mekânsal dağılımının rastgele olup olmadığı test edilir.

Moran'ın I indeksi, noktaların sahip olduğu değişken değerlerinin benzerlik ve yakınlık değerini

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_{i=1}^n (x_i - \bar{x})^2} \quad (10)$$

biçiminde tek bir değerde birleştirir. Burada, n , örnekleme nokta sayısını, x_i , $i = 1, 2, \dots, n$ noktasına ait öznelik değerini, x_j , $j = 1, 2, \dots, n$ noktasına ait öznelik değerini, \bar{x} , özneliğe ait ortalama değeri ve S_0 ise, örneklem için mekânsal ağırlıklar toplamını belirtir ve

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij} \quad (11)$$

olarak hesaplanır.

Beklenen indeks değeri $E(I) = (-1)/(n-1)$, Moran'ın I indeksi değeri ile karşılaştırılarak yorumlanır. $I > E(I)$ olduğu durumda gözlemlenen mekânsal nokta veride kümelenme eğiliminin bulunduğu, $I < E(I)$ olduğu durumda tekdüze dağılımın bulunduğu ve $I = E(I)$ durumunda ise noktaların rastgele dağıldığı ve özneliğe bağlı bir kümelenme eğiliminin olmadığı değerlendirilir [11].

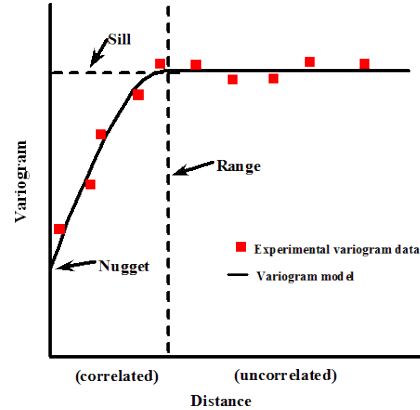
2.4. Mekânsal Enterpolasyon Yöntemleri

Mekânsal enterpolasyon, belirli bir coğrafi alandaki ölçüm noktaları arasındaki değerleri tahmin etmek için kullanılan bir yöntemdir [13]. Mekânsal enterpolasyon amacıyla Kriging analizi uygulanmıştır.

Kriging Analizi, ölçümü yapılmış konumlardan, ölçüm yapılmamış olan öznelik değerini öngörmek için son yıllarda jeostatistik alanında yaygın olarak kullanılan bir enterpolasyon yöntemidir [14], [15]. Kriging ile variogram oluşturularak mekânsal korelasyon modellenmekte ve bu modelden tahminler elde edilmektedir.

Variogram, mekânsal nokta verilerde mekânsal korelasyonu ölçmek için jeostatistikte kullanılan temel

bir araçtır. Veri noktaları arasındaki mesafe arttıkça veri benzerliğinin nasıl azaldığını açıklamaktadır. Variogram modeli, mekânsal enterpolasyon yöntemi olan Kriging'in gerçekleştirilmesi için çok önemlidir. Şekil 5, variogram modeli tahmini yapılmış deneysel bir variogramın grafiğini göstermektedir. Her kırmızı kare, deneysel variogramın bir gecikmesidir. x -ekseni, nokta çiftleri arasındaki mesafeyi ve y -ekseni de variogramın hesaplanan değerini temsil etmektedir. Grafikte görülen büyük bir değer, nokta çiftleri arasındaki korelasyonun daha düşük olduğunu göstermektedir. Şekil 5'teki grafik ayrıca variogram modelinin uyumunu kontrol eden Nugget, Range ve Sill olarak adlandırılan üç önemli parametreyi göstermektedir. Nugget, variogramın y -ekseni ile kesişimidir ve verilerin küçük ölçekli değişkenliğini diğer bir deyişle ölçüm hatalarını temsil etmektedir. Range, variogramın düzleştiği mesafedir. Range değerinin fiziksel anlamı, bu mesafede veya daha fazla uzaklıkta olan nokta çiftlerinin mekânsal olarak ilişkili olmamasıdır. Sill ise, toplam varyansı göstermektedir [16]. Literatürde kullanılan çeşitli variogram modelleri, Küresel Model, Üssel Model, Gauss Model, Doğrusal Model, Başlangıç Değeri Modeli biçiminde tanımlanabilir. Mekânsal veri analizlerinde en sık kullanılan ve kısa mesafelerde güçlü korelasyonları temsil eden yöntemler ise Küresel ve Üssel modellerdir.



Şekil 5. Variogram grafiği

Kriging Analizi, diğer tahmin yöntemlerine kıyasla daha yansız sonuçlar sağlamakta ve minimum varyans ile tahminin standart sapmasını hesaplama olanağı sunmaktadır. Kriging'i diğer enterpolasyon yöntemlerinden ayıran en önemli özellik, tahmin edilen her nokta veya alan için bir varyans değeri hesaplanabilmesidir. Bu yöntemin bir diğer avantajı ise, kriging varyansı sayesinde tahmin hatasının büyüklüğünü değerlendirme imkânı sunmasıdır [17]. Yaygın olarak kullanılan Kriging yöntemleri arasında Sıradan Kriging (Ordinary Kriging), Basit Kriging (Simple Kriging), Evrensel Kriging (Universal Kriging), Blok Kriging (Block Kriging), Göstergeli Kriging (Indicator Kriging), Ayırık Kriging (Disjunctive Kriging) ve Ko-Kriging (Cokriging) yöntemleri bulunmaktadır [1].

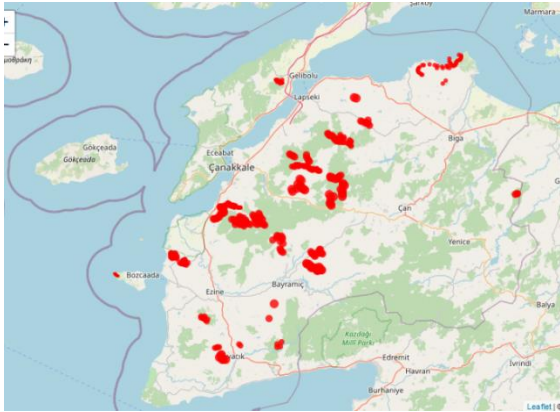
Ordinary Kriging yönteminin ilk adımı, enterpolasyon yapılacak noktalar kümesinden bir variogram oluşturmaktır. İkinci aşamada, deneysel variogramdaki trendi modelleyen basit bir matematiksel fonksiyon olan teorik variogram belirlenir. Ordinary Kriging yönteminde, bilinmeyen değerlerin belirlenmesi, değişkenlerin durağan olduğu ve ortalamasının sabit olduğu varsayımına dayanmaktadır. Variogram fonksiyonundan ağırlıkların belirlenmesi sırasında, tahmin ağırlıkları variogram modellerine göre hesaplanır.

III. UYGULAMA

Bu çalışmada, Türkiye’de yer alan RES’lere ve GES’lere ilişkin mekânsal nokta verilerin mekânsal istatistiklerle analizi yapılmıştır. EPDK tarafından lisans verilen işletmelerden RES’ler için Çanakkale ili verileri ve GES’ler için ise lisans verilen tüm nokta verileri ile öznitelik değerleri olarak kurulu güç değerleri dikkate alınmıştır. Mekânsal istatistikler, RES ve GES nokta verilerine dört temel başlıkta uygulanmıştır: (i) Mekânsal Betimsel İstatistikler (Mekânsal Ortalama, Standart Uzaklık, Standart Sapma Elipsi), (ii) Mekânsal Örüntü Analizi (Kuadrat Analizi, En Yakın Komşuluk Analizi), (iii) Mekânsal Otokorelasyon (Moran’ın I indeksi) ve Mekânsal Enterpolasyon (Kriging). Mekânsal istatistik analizlerinde RStudio (1.1.456 versiyon) programı kullanılmıştır.

3.1. Rüzgâr Enerji Santrali (RES)

EPDK Şeffaflık Platformundan RES verileri temin edilmiştir. 433 adet mekânsal nokta veriler koordinat ve kurulu güç özniteliklerini içermekte olup Çanakkale ili içindeki RES’leri temsil etmektedir. RES’lere ait veriler Şekil 6’da görselleştirilmiştir.

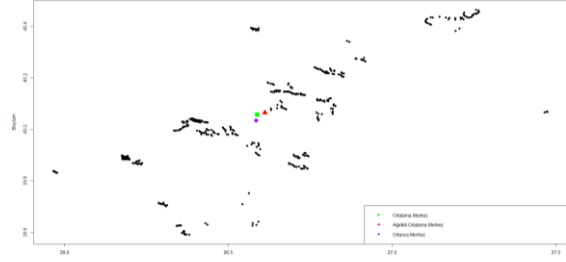


Şekil 6. EPDK Verilerine Göre Çanakkale’deki RES’ler

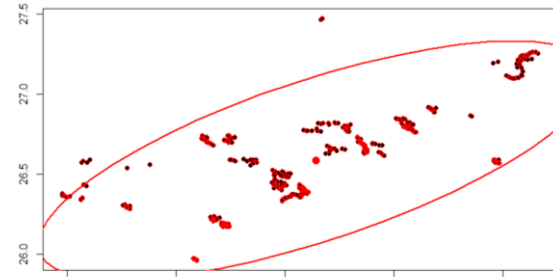
• RES için Mekânsal Betimsel İstatistikler

Şekil 7’de Rüzgâr türbinlerine ait koordinatların ortalama merkezi (40.057, 26.588) yeşil renk ile koordinatlara ait Kurulu Güç değerlerine ilişkin ağırlıklı ortalama merkezi (40.064, 26.611) kırmızı

renk ile ortanca merkez ise (40.033, 26.585) mor renk ile gösterilmiştir. Mekânsal ortalama ve mekânsal ağırlıklı ortalama birbirine yakın bulunmuştur.



Şekil 7. Çanakkale’deki RES’lerin mekânsal ortalama, mekânsal ağırlıklı ortalama ve mekânsal ortanca merkezi

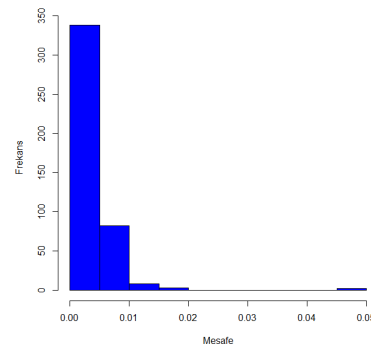


Şekil 8. RES’lere ait standart sapma elipsi

Mekânsal yayılım için Şekil 8’den standart sapma elipsine göre RES’lerin kuzeydoğu-güneybatı yönünde dağılım gösterdiği söylenir.

• RES için Mekânsal Örüntü Analizi

Şekil 9’da en yakın komşu mesafelerinin dağılımının sağa çarpık olduğu görülmektedir. Bu durum RES’lerin çoğunluğunun yakınlarında bir başka RES’in bulunduğu anlamına gelmektedir. Çok küçük mesafe (< 0.01) aralıklarında en yüksek frekanslar gözlemlenmiştir. Bu durum, RES’lerin çoğunlukla çok küçük mesafelerde kümelenmiş olduğunu göstermektedir. Orta ve Yüksek Mesafe (> 0.01) aralıklarında frekanslar hızla düşmektedir. Bu durum, daha uzak mesafelerdeki RES’lerin sayısının az olduğunu ve RES’lerin daha geniş alanlara yayılmadığını göstermektedir.



Şekil 9. RES’lerin en yakın komşuluk dağılımı

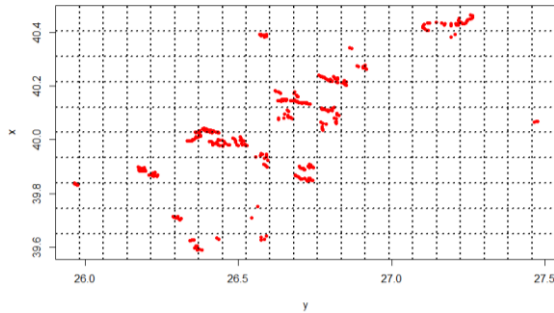
RES'ler için Kuadrat Analizi'nde öncelikle Çanakkale ilinin yüzölçümüne göre kuadrat kenar uzunluğu

$$k = \sqrt{K} = \sqrt{\frac{2(9817)}{433}} = 6.733804 \text{ km}^2$$

olarak elde edilmiştir.

Çanakkale ili için kısa kenar, 67 km ve uzun kenar ise 140 km olarak bulunmuştur. Buna göre, sırasıyla, kuadrat satır ve sütun sayıları $r_k = \frac{67}{6.733804} \cong 10$ ve $c_k = \frac{140}{6.733804} \cong 21$ olarak elde edilmiştir.

Kuadrat Analizi ile RES'lerin sıklığı baz alınarak nokta verilerin mekandaki dağılımı gösterilmektedir. 10×21'lik Kuadrat Analizi yapıldığında, Şekil 10'a göre RES rüzgâr hızlarının kuzey doğuda yoğunlaştığı gözlenmiştir. Şekil 10'dan, bazı bölgelerde RES'lerin kurulu olmadığı görülmektedir. Bu durumun olası nedenleri, rüzgâr verimliliği, arazi ve zemin yapısı, çevresel alanlar ve yasal düzenlemelerin uygun olmaması olarak değerlendirilebilir.

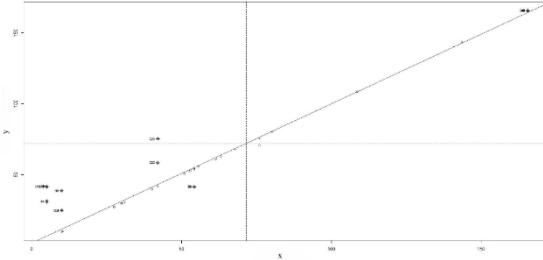


Şekil 10. RES'ler için kuadrat grafiği

• RES için Global Mekânsal Otokorelasyon

Global mekânsal otokorelasyon analizi sonucunda Moran'ın I indeksi değeri, $I = 0.9796$ bulunmuştur. $I > E(I)$ olduğundan Çanakkale'de yer alan RES'lerin kullanım gücünün mekânsal dağılımının rastgele olmadığı söylenebilir. Değer +1'e yakın olduğu için pozitif mekânsal otokorelasyonun olduğu yani benzer değerlerin kümelenmiş olduğu söylenebilir. Buna göre, gözlemlenen mekânsal nokta veride kümelenme eğiliminin olduğu yani benzer değişken değerlerine sahip noktaların birbirlerine mekânsal olarak yakın konumlandığı sonucuna ulaşılmıştır.

Şekil 11'deki Moran saçılım grafiğinde, her bir gözlemin standartlaştırılmış değeri (x -ekseni) ve bu değerlerin mekânsal ağırlıklı ortalaması (y -ekseni) görülmektedir. Çizgi, pozitif bir eğim göstermektedir. Gözlemlerin büyük çoğunluğunun çizgiye yakın olması pozitif mekânsal otokorelasyonun olduğunu göstermektedir.

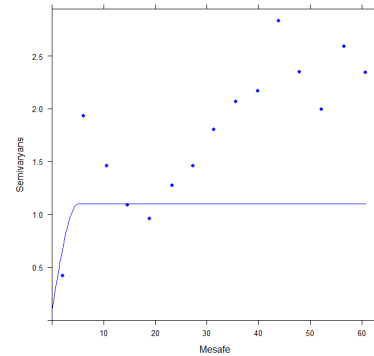


Şekil 11. RES'ler için Moran saçılım grafiği

• RES için Mekânsal Enterpolasyon

Variogram analizinde Nugget = 0.1, Sill = 1.1 ve Range = 1 bulunmuştur. Nugget değerine göre küçük mesafelerde gözlemlenen varyans ve ölçüm hataları oldukça düşüktür. Sill değerine göre RES'lerin toplam varyansı 1.1'dir. Bu değer veri setinin genelindeki değişimi göstermektedir. Mekânsal otokorelasyonun etkili olduğu mesafe yaklaşık olarak 1 km'dir. Bu, RES'lerin yoğunlukla 1 km içinde mekânsal olarak ve ayrıca kurulu güç ile ilişkili olduğunu göstermektedir.

Küçük mesafede düşük semivaryans, RES'lerin birbirine yakın olanların benzer kurulu güçlere sahip olduğu ve güçlü bir mekânsal otokorelasyonun olduğu Şekil 12'deki variogram grafiğinde görülmektedir.



Şekil 12. RES variogram grafiği

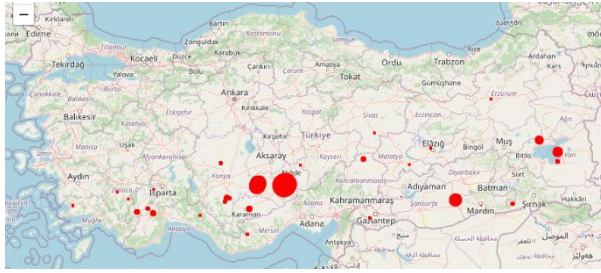
Kriging Analizi ile öngörü yapılabileceğini göstermek için 10 katlı çapraz doğrulama (Cross Validation-CV) uygulanmıştır. 433 adet mekânsal nokta veri incelenmiştir. 433 adet mekânsal nokta veri için kurulu güç değerleri öngörülmüştür. Bilinen gerçek kurulu güç değerleri ile öngörü değerleri karşılaştırılarak hatalar hesaplanmıştır. Hata Kareler Ortalamasının Karekökü (Root Mean Squared Error - RMSE) hesaplanmıştır. Her yinelemede elde edilen bu değerler Tablo 2'de gösterilmiştir. Tablo 2'deki RMSE değerleri incelendiğinde test verisinde Kriging modelinin iyi performans gösterdiği söylenebilir. Ortalama RMSE değeri, 0.1863 olarak hesaplanmıştır.

Tablo 2. RES'ler için CV iterasyonu *RMSE* değerleri

<i>İterasyon</i>	<i>RMSE</i>
1	0.1016
2	0.1773
3	0.1979
4	0.2028
5	0.3207
6	0.1299
7	0.0875
8	0.1650
9	0.2260
10	0.1412

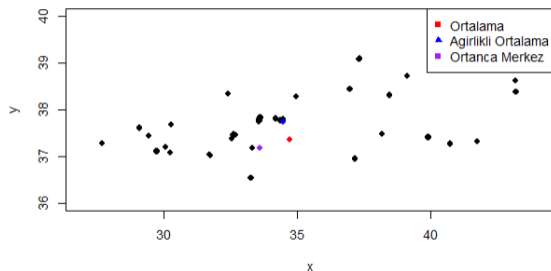
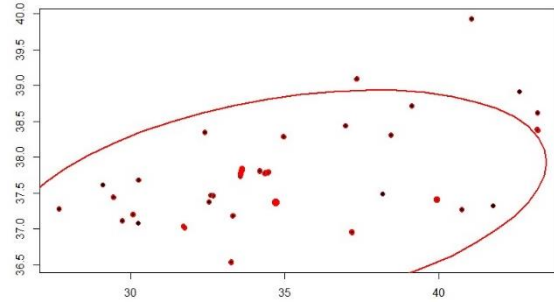
3.2. Güneş Enerji Santrali (GES)

EPDK Şeffaflık Platformundan GES verileri temin edilmiştir. Temin edilen veriler, Türkiye'de EPDK tarafından lisans verilen kurulu GES'leri temsil etmektedir. 1306 adet nokta veri olarak elde edilen verilerin koordinat ve kurulu güç değerleri mevcuttur. Türkiye'deki GES'lere ait veriler kurulu güç değerlerine göre boyutlandırılarak Şekil 13'de görselleştirilmiştir.

**Şekil 13.** Türkiye'deki GES'lere ait koordinatlar

• *GES için Mekânsal Betimsel İstatistikler*

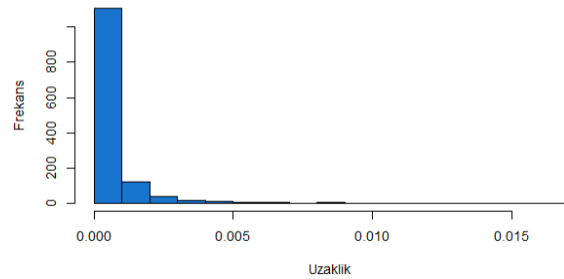
Şekil 14'te Güneş Panellerine ait koordinatların ortalama merkezi (37.3660, 34.7092) (Adana, Pozantı) kırmızı renk ile koordinatlara ait kurulu güç değerlerine ilişkin ağırlıklı ortalama merkezi (37.7318, 34.4665) (Niğde-Bor) mavi renk ile ortanca merkez ise (37.1889, 33.6009) (Karaman-Yeşildere) mor renk ile gösterilmiştir. Ortalama ve ağırlıklı ortalama değerlerinin birbirine yakın olduğu Şekil 14'ten görülmektedir.

**Şekil 14.** GES'lerin mekânsal ortalama, mekânsal ağırlıklı ortalama ve mekânsal ortanca merkezi**Şekil 15.** GES'lere ait standart sapma elipsi

Şekil 15'te yer alan standart sapma elipsine göre güneş panelleri kuzeydoğu-güneybatı yönünde dağılım göstermiştir.

• *GES için Mekânsal Örüntü Analizi*

Şekil 16'da, en yakın komşu mesafelerinin dağılımının sağa çarpık olduğu görülmektedir. Bu durum, GES'lerin yoğunluğunun yakınlarında bir başka GES'in bulunduğu anlamına gelmektedir. Çok küçük mesafe (< 0.01) aralıklarında en yüksek frekanslar gözlemlenmiştir. Bu durum, GES'lerin yoğunlukla çok küçük mesafelerde kümelenmiş olduğunu göstermektedir. Orta ve Yüksek Mesafe (> 0.01) aralıklarında frekanslar hızla düşmektedir. Bu durum, daha uzak mesafelerdeki GES'lerin sayısının az olduğunu ve GES'lerin daha geniş alanlara yayılmadığını göstermektedir.

**Şekil 16.** GES'lerin en yakın komşuluk dağılımı

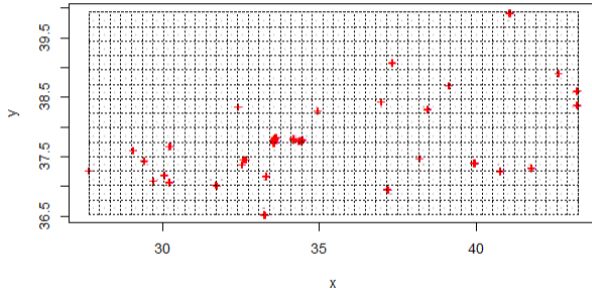
GES'ler için Kuadrat analizinde öncelikle Türkiye'nin yüzölçümüne göre kuadrat kenar uzunluğu (k)

$$k = \sqrt{K} = \sqrt{\frac{2(783562)}{1306}} = 34.64 \text{ km}^2$$

olarak elde edilmiştir. Türkiye için kısa kenar 490 km ve uzun kenar ise 1600 km olarak bulunmuştur. Buna göre, sırasıyla, kuadrat satır ve sütun sayıları $r_k = \frac{490}{34.64} = 14.145 \cong 14$ ve $c_k = \frac{1600}{34.64} = 46.189 \cong 46$ olarak elde edilmiştir.

Kuadrat Analizi ile GES'lerin sıklığı baz alınarak mekandaki dağılımı gösterilmektedir. 14×16 'lık Kuadrat analizi yapıldığında Şekil 17'de, GES'lerin,

sıcaklığın fazla olduğu Konya, Niğde ve Antalya illerinde daha çok olmak üzere güneyde yoğunlaştığı görülmektedir.

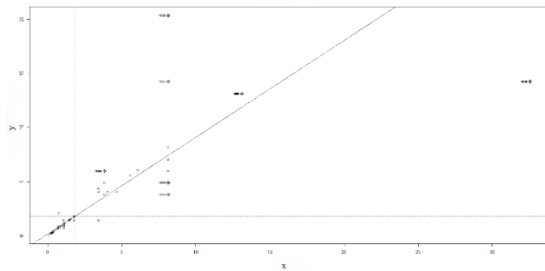


Şekil 17. GES'ler için kuadrat grafiği

• GES için Global Mekânsal Otokorelasyon

Global mekânsal otokorelasyon analizi sonucunda Moran'ın I indeksi değeri, $I = 0.8944$ olarak bulunmuştur. $I > E(I)$ olduğundan Türkiye'de yer alan lisanslı güneş panellerinin kullanım gücünün mekânsal dağılımının rastgele olmadığı söylenebilir. Değer +1'e yakın olduğu için pozitif mekânsal otokorelasyonun olduğu yani benzer değerlerin kümelenmiş olduğu söylenebilir. Buna göre, gözlemlenen mekânsal nokta veride kümelenme eğiliminin olduğu yani benzer değişken değerlerine sahip noktaların birbirlerine mekânsal olarak yakın konumlandığı sonucuna ulaşılmıştır.

Şekil 18'de görülen Moran saçılım grafiğinde, her bir gözlemin standartlaştırılmış değeri (x-ekseni) ve bu değerlerin mekânsal ağırlıklı ortalaması (y-ekseni) görülmektedir. Çizgi, pozitif bir eğim göstermektedir. Gözlemlerin büyük çoğunluğunun çizgiye yakın olması pozitif mekânsal otokorelasyonun olduğunu göstermektedir. Ayrıca noktaların orijine yakın olmasının bu noktaların kapasitelerinin benzer ve nispeten düşük kapasitelere sahip olduğu söylenebilir.



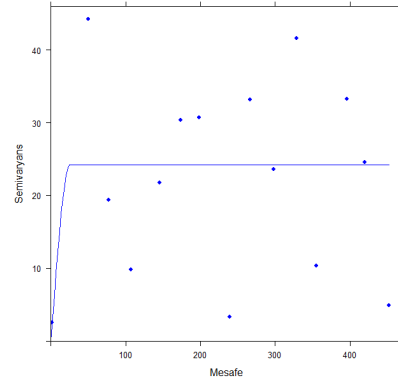
Şekil 18. GES'ler için Moran saçılım grafiği

• GES için Mekânsal Enterpolasyon

Variogram analizinde, Nugget = 0, Sill = 24.2189 ve Range = 25.7299 bulunmuştur. Nugget değerine küçük mesafelerde gözlemlenen varyans ve ölçüm hataları yoktur. Sill değerine göre GES'lerin toplam varyansı yaklaşık olarak 24.22'dir. Bu değer veri setinin genelindeki değişimi göstermektedir. Mekânsal

otokorelasyonun etkili olduğu mesafe yaklaşık olarak 25.73 km'dir. Bu, GES'lerin çoğunlukla 25.73 km içinde mekânsal olarak ayrıca kurulu güç ile ilişkili olduğunu göstermektedir.

Küçük mesafede düşük semivaryans, GES'lerin birbirine yakın olanların benzer kurulu güç değerlerine sahip olduğunu ve güçlü bir mekânsal otokorelasyonun olduğu Şekil 19'da variogram grafiğinde görülmektedir.



Şekil 19. GES variogram grafiği

Kriging Analizi ile öngörü yapılabileceğini göstermek için 10 katlı CV uygulanmıştır. 1306 adet mekânsal nokta veri incelenmiştir ve bu mekânsal nokta veriler için kurulu güç değerleri öngörülmüştür. Bilinen gerçek kurulu güç değerleri ile öngörü değerleri karşılaştırılarak hatalar hesaplanmıştır. Her yinelemede $RMSE$ değerleri hesaplanarak Tablo 3'te gösterilmiştir. Tablo 3'teki $RMSE$ değerleri incelendiğinde, test verisinde Kriging modelinin iyi performans gösterdiği söylenebilir. Ortalama $RMSE$ değeri ise 1.2406 olarak hesaplanmıştır.

Tablo 3. GES'ler için CV iterasyonu $RMSE$ değerleri

İterasyon	$RMSE$
1	2.6157
2	0.0819
3	0.1211
4	1.0477
5	0.2599
6	0.0888
7	0.2513
8	1.6281
9	2.1425
10	0.24335

IV. SONUÇ

Bu çalışmada, ülkemizde yer alan RES'lerin en yoğun olduğu illerden biri olan Çanakkale'deki RES mekânsal nokta verileri ile ülkemizdeki mevcut GES'lere ilişkin mekânsal nokta verilerin mekânsal istatistiksel yöntemler ile analizi amaçlanmıştır.

RES'lere ilişkin elde edilen bulgular doğrultusunda, Çanakkale ili nokta verisinde mekânsal ortalama ve mekânsal ağırlıklı ortalamasının birbirine yakın olduğu, standart sapma elipsine göre kuzeydoğu-güneybatı yönünde dağılım gösterdiği, En Yakın Komşuluk Analizi'ne göre ise en yakın komşu mesafesinin dağılımının sağa çarpık yani RES'lerin çoğunluğunun yakınlıklarında RES'lerin olduğu görülmüştür. Kuadrat analizi sonucunda ise, RES rüzgâr hızlarının yoğun olduğu yerin kuzey doğuda olduğu gözlemlenmiştir. Moran'ın *I* indeksine göre gözlemlenen mekânsal nokta veride kümelenme eğiliminin olduğu yani benzer değerlere sahip noktaların birbirlerine mekânsal olarak yakın konumlandığı sonucuna ulaşılmıştır. Kriging Analizi'nde ise 10-kat CV uygulanarak her iterasyonda *RMSE* değerleri elde edilmiştir. Kriging Analizi sonucu Çanakkale bölgesinde 10-kat CV'de yer alan test verileri için RES'lerin gerçek kurulu güçleri ve tahmini kurulu güçleri karşılaştırıldığında yakın sonuçlar elde edildiği görülmüştür.

GES'lere ilişkin elde edilen bulgular doğrultusunda, lisanslı tüm nokta verilerin mekânsal ortalama ve mekânsal ağırlıklı ortalamasının birbirine yakın olduğu, standart sapma elipsine göre kuzey doğu/güney batı yönünde dağılım gösterdiği, En Yakın Komşuluk Analizi'ne göre ise en yakın komşu mesafesinin dağılımının sağa çarpık yani GES'lerin çoğunluğunun yakınlıklarında GES'lerin olduğu görülmüştür. Kuadrat analizi sonucunda ise, GES'lerin sıcaklığın fazla olduğu güneyde yoğunlaştığı gözlemlenmiştir. Moran'ın *I* indeksine göre gözlemlenen mekânsal nokta veride kümelenme eğiliminin olduğu yani benzer değerlere sahip noktaların birbirlerine mekânsal olarak yakın konumlandığı sonucuna ulaşılmıştır. Kriging Analizi'nde ise 10-kat CV uygulanarak her iterasyonda *RMSE* değerleri elde edilmiştir. Elde edilen *RMSE* değerleri incelendiğinde Kriging modelin bazı test veri kümelerinde daha iyi performans gösterdiği görülürken bazılarında ise performansın beklenildiği kadar iyi olmadığı söylenebilir.

TEŞEKKÜR

Bu çalışmada yer verilen uygulamaların bir kısmı V. Uluslararası Uygulamalı İstatistik Kongresi (UYİK-2024)'nde sözlü olarak sunulmuştur. Değerli katkıları için Editörlere ve Hakemlere çok teşekkür ederiz.

KAYNAKLAR

[1] Eryılmaz, H. (2010). Mekânsal İstatistikte Nokta Örüntü Teknikleri ve Bir Uygulama, Doktora tezi, Anadolu Üniversitesi, Türkiye, s. 6-20.

- [2] Fisher, M. M. ve Wang, J. (2011). Spatial Data Analysis Models, Methods and Techniques, Springer, New York, s. 1-10.
- [3] Çubukçu, M. (2020). Planlamada ve Coğrafyada Temel İstatistik ve Mekânsal İstatistik. Nobel Yayınevi, Ankara, s. 1-295. H. Jafari, A.K. Gupta, Impact strength and dynamic mechanical properties correlation in elastomer-modified polypropylene, *J. Appl. Polym. Sci.* 78 (2000).
- [4] Crawley, M. J. (2013). The R Book, Wiley, West Sussex, United Kingdom, s.825-867.
- [5] Erkurtulmuş, S. A. (2022). Açık deniz üzeri yüzer rüzgâr türbini platform konseptleri. Karadeniz Teknik Üniversitesi, Türkiye, s.1-295.
- [6] Akalın, S. (2018). Açık Deniz rüzgâr enerjisi türbinlerinin kurulum yeri seçimi için bir model önerisi. Y.lisans tezi, Gazi Üniversitesi, Türkiye, s. 1-295.
- [7] 2014-2024 yılları arasında Türkiye elektrik enerjisi kurulu güç değerleri. <https://enerji.gov.tr/bilgi-merkezi-enerji-elektrik>, (Erişim Tarihi: Haziran 2024).
- [8] Yıllara göre Türkiye'nin elektrik üretiminde kullanılan yenilenebilir enerji kaynakları oranı. <https://www.epdk.gov.tr/detay/icerik/3-0-167/resmi-istatistikler>, (Erişim Tarihi: Temmuz 2024)
- [9] Türkşen, Ö. (2019). An application of spatial statistics: spatial analysis of simulated fault plane geodetic points, *BAUN Fen Bil. Enst. Dergisi*, 21(1), 81-93.
- [10] Wang, B. and Shi, W. ve Miao, Z. (2015). Confidence analysis of standard deviational ellipse and its extension into higher dimensional euclidean space. *PLoS ONE*, 10(3), 1-17.
- [11] Wong, W.S.W. ve Lee, J. (2005). Statistical analysis of geographic information with ArcView GIS and ARCGIS. John Wiley and Sons, s.1-464.
- [12] Moran, P. A. P. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society, Series B (Methodological)*, 10(2), 243-251.
- [13] Kınalıoğlu, İ. (2024). Coğrafi bilgi sistemleri destekli mekânsal istatistiksel yöntemler kullanılarak konya şehrinde trafik kaza kara noktalarının belirlenmesi. Y.lisans Tezi, Uşak Üniversitesi, Türkiye, s. 48-59.
- [14] Caruso, C. ve Quarta, F. (1998) Interpolation methods comparison. *Computers and Mathematics with Applications*, 35, s.109-126.
- [15] Kalkhan, M. A. (2011). Spatial statistics geospatial information modelling and thematic mapping, CRC Press, USA.
- [16] Pacific Northwest National Laboratory (PNNL). Kriging Variogram Model, https://vsp.pnnl.gov/help/Vsample/Kriging_Variogram_Model.htm (Erişim Tarihi: Haziran, 2024)
- [17] Armstrong, M. (1998). Basic linear geostatistics, Springer, Berlin.

Türkiye'deki Ekonomik Büyüme Üzerine Gecikmesi Dağıtılmış Otoregresif Sınır Testi Yaklaşımı

Autoregressive Distributed Lag Bound Test Approach On Economical Growth In Turkey

Sinem GEÇGEL KARAGÖL¹, Barış AŞIKGİL²

¹ Mimar Sinan Güzel Sanatlar Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Yüksek Lisans Programı,
İstanbul, Türkiye

² Mimar Sinan Güzel Sanatlar Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü, İstanbul, Türkiye

Öz

Bu çalışmanın amacı, Türkiye'de 1960-2014 dönemleri arasında, ekonomik büyüme, elektrik tüketimi ve enflasyon arasındaki ilişkinin kısa ve uzun dönemdeki etkilerinin araştırılmasıdır. Genişletilmiş Dickey-Fuller (ADF) ve Phillips-Perron (PP) birim kök testlerinden faydalanılarak değişkenlerin birinci düzeyde durağan oldukları tespit edilmiştir. Böylece değişkenler üzerinde uzun dönem ilişkinin belirlenmesi için Pesaran vd. (2001) tarafından geliştirilen Gecikmesi Dağıtılmış Otoregresif (ARDL) Model yaklaşımı yapılmıştır. Uzun dönemde ARDL sınır testi katsayılarının incelenmesinden sonra kısa dönem dinamiklerine ilişkin hata düzeltme modeline (HDM) bakılarak katsayılar üzerinden yorum yapılmıştır. Model üzerinde yapısal kırılma olup olmadığını araştırmak için CUSUM ve CUSUM Kare testleri kullanılmış ve modelde yapısal kırılma olmadığı sonucuna ulaşılmıştır. Bunun yanında, hata terimiyle ilgili varsayımların sağlanıp sağlanmadığı incelenmiştir. ARDL sınır testi yaklaşımı ile uzun dönemde Türkiye'deki enflasyon ve elektrik tüketiminin ekonomik büyümeyi olumlu veya olumsuz etkileyeceği üzerinde durulmuştur.

Anahtar Kelimeler: Ekonomik Büyüme, Enflasyon, Elektrik Tüketimi, ARDL Sınır Testi, Gecikmesi Dağıtılmış Otoregresif Model.

Abstract

The aim of this study is to explore long and short term effects of relations among consumption of electricity, inflation and economical growth by using the annual data between period of 1960-2014 in Turkey. It is observed by using Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) unit root tests that the first differences of the variables are stationary. Thus, the Autoregressive Distributed Lag (ARDL) Model approach given by Pesaran et al. (2001) is used in order to determine long term relationship. After examining the long-term ARDL bound test coefficients, the error correction model (ECM) for short-term dynamics is interpreted based on the coefficients. In order to explore whether there is a structural break in the model, CUSUM and CUSUM Square tests are used and in view of the results of these tests there is no structural break. Moreover, it is examined whether the assumptions related with the error term are satisfied or not. It is emphasized that inflation and electricity consumption in Turkey will affect the economic growth positively or negatively in the long term by using the ARDL bounds test approach.

Keywords: Economical Growth, Inflation, Consumption of Electricity, ARDL Bounds Test, Autoregressive Distributed Lag Model.

I. GİRİŞ

Küreselleşen dünyada, insanoğlunun gündelik gereksinimlerinin karşılanması ve sanayi devrimin gerçekleşmesiyle birlikte insan gücünün yerini makinelerin alması sonucu elektriğe karşı duyulan ihtiyaç her geçen gün artmaktadır. Elektrik tüketimine karşı artan talep sonucu ülkeler yatırımlarını enerji üzerine yönlendirmişlerdir. Uzun vadede bakıldığında, enerji üzerindeki yatırım maliyetlerinin yüksek olması sebebiyle sürdürülebilir ve izlenen politikanın gerçekler doğrultusunda olması ülkenin geleceği açısından önemli bir konuyu teşkil etmektedir.

Tarihte iktisatçıların hemfikir olamayıp tartışmalara sebep olduğu konulardan biri ekonomik büyüme ve enflasyon arasındaki ilişkidir. İkinci Dünya Savaşı'na kadarki sürede, fiyatların umumi düzeyinde değişme görülmemişken, savaş sonrasında yaşanan yüksek enflasyon, enflasyon ile ekonomik büyüme arasındaki ilişkiye karşı duyulan ilgiyi arttırmıştır. Bu doğrultuda ekonomik büyüme, enflasyon ve elektrik tüketimi arasındaki nedensellik ilişkinin incelenmesi için ARDL yaklaşımdan faydalanılmıştır.

Bu çalışma dört temel bölümü kapsamaktadır. İlk bölümde, ekonomik büyüme, enflasyon ile elektrik tüketimi konuları üzerinde durulmuş olup konuyla ilgili yakın zamandaki literatür taraması yapılmıştır. İkinci bölümde, materyal ve metod üzerinde durulmuştur. Üçüncü bölümde, konuyla ilgili analizler yapılmıştır. Dördüncü bölümde, elde edilen analiz sonuçları yorumlanmıştır.

1.1. Literatür Araştırması

Halıcıoğlu [1], 1960-2005 dönemleri arasında yıllık veri setinden faydalanarak Türkiye'de uzun dönemde karbondioksit salınımının enerji tüketimi, gelir ile dış ticaretten etkilendiğini ifade etmiştir. Öztürk ve Acaravcı [2], 1968-2005 dönemlerindeki yıllık verileri kullanarak ekonomik büyüme, karbondioksit salınımı, enerji tüketimi ile istihdam değişkenlerinin uzun dönemli ilişki varlığını ileri sürmüşlerdir. Polat vd. [3], 1950-2016 dönemleri arasında yıllık veri setinden yararlanarak Türkiye'de kısa dönemde sadece istihdamdan elektrik tüketimine olumlu yönde nedensellik olduğunu uzun döneme bakıldığında ise istihdam ile elektrik tüketimi değişkenlerinden reel gayri safi milli hasıla (GSMH)'ye doğru pozitif nedensellik olduğunu belirtmişlerdir. Fuinhas ve Marques [4], 1965-2009 dönemleri arasında yıllık veri seti kullanarak, Portekiz, İtalya, Yunanistan, İspanya ve Türkiye'de enerji tüketimi ile ekonomik büyüme arasında geri beslemeye dayalı nedensellik olduğunu ileri sürmüşlerdir. Öcal ve Aslan [5], 1990-2010 dönemlerinde Türkiye'de yenilenebilir enerji tüketiminin ekonomik büyümeyi negatif yönde etki ettiğini tespit etmişlerdir. Koçak [6], 1960-2010 dönemleri arasında yıllık veri setinden yararlanarak Türkiye'de uzun dönemde enerji kullanımının karbondioksit salınımını arttırdığını ifade etmiştir. Özçağ [7], 1960-2013 dönemleri arasında yıllık veri kullanarak Türkiye'de kısa ve uzun dönemde ekonomik büyümeye enerji tüketimi arasında pozitif nedensellik olduğunu belirtmiştir. Pata ve Terzi [8], 1964-2014 dönemleri arasında yıllık veri kullanarak Türkiye'nin uzun dönemde turizm ile ekonomik büyüme ilişkisinin bulunmadığını fakat kısa döneme bakıldığında turizm sektöründen gayri safi yurtiçi hasıla (GSYH)'ye doğru pozitif nedenselliğin varlığını ifade etmişlerdir. Türemez ve Göktaş [9], 1984-2017 dönemleri arasında yıllık veri seti kullanarak Türkiye'de turizm talebinin Avrupa Birliği toplam geliri ile Gümrük Birliği'ne katılımı anlamlı olarak etki ederken, reel döviz kuru ile Avrupa Birliği adaylık aşamasında ülkeye giriş yapan

turist sayısını anlamlı olarak etki etmediği sonucuna varmışlardır. Terzi ve Bekar [10], 1974-2014 dönemleri arasında yıllık verileri kullanarak Türkiye'deki doğrudan yabancı yatırımlar, turist sayısı ile dışa açıklığın birbirini etkilediğini ifade etmişlerdir.

II. MATERYAL VE METOD

Çalışmada, 1960-2014 dönemleri arasındaki 55 gözleme sahip yıllık veriler World Bank (<https://data.worldbank.org/>) [11]'tan elde edilmiş olup EViews10 paket programından yararlanılmıştır. Bağımlı değişken olan gayri safi milli hasıla ile bağımsız değişkenler olan elektrik tüketimi ile enflasyon arasındaki hem kısa hem de uzun dönem ilişkilerinin incelenmesi amaçlanmıştır.

Öncelikle, değişkenlere logaritmik dönüşüm uygulanmıştır. Logaritma dönüşümü uygulanan değişkenlerin gösterimi, gayri safi yurt içi hasıla (GSYH), elektrik tüketimi (EL), enflasyon (ENF) şeklindedir. Ardından, normallik varsayımının sağlanıp sağlanmadığı incelenmiştir. Sonrasında, durağanlıklarının incelenmesi için birim kök testinden yararlanılarak durağanlıklarının birinci farkta olduğu sonucuna varılmıştır. Uzun dönemde değişkenler arasındaki denge ilişkisine bakılmak için, diğer bir deyişle değişkenler arasındaki eşbütünlüşme ilişkisinin olup olmadığı ARDL yaklaşımı aracılığıyla incelenmiştir.

2.1. Gecikmesi Dağıtılmış ve Otoresif Modeller

y değişkeni, x değişkenine belirli bir zaman boşluğu sonrasında cevap verdiği takdirde, bu boşluğa gecikme, ilgili modele ise gecikmeli ilişki denilmektedir. Dinamik modeller gecikmeli ve otoresif modeller olarak adlandırılmaktadır. Bağımlı değişken bağımsız değişkenlerin gecikmeleri tarafından açıklanıyorsa gecikmesi dağıtılmış model olarak ifade edilmektedir. Gecikmesi dağıtılmış model aşağıda gösterilmektedir:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_j x_{t-j} + \varepsilon_t; \quad t = 1, 2, \dots, n \quad (1)$$

En küçük kareler (EKK) yönteminden yararlanılarak model tahmini yapılmaktadır. Fakat bağımsız değişkenler kendi gecikmelerinden meydana geldiğinden çoklu doğrusal bağlantı problemi oluşması ihtimali vardır. Ayrıca gecikme uzunluğundan dolayı tahmin edilen modelde gözlem kaybı yaşanması ihtimali mevcuttur.

Gecikmesi dağıtılmış modeller beşe ayrılmaktadır. Bunlar; Koyck modeli, Almon (Polinomial) gecikme modeli, Nerlove kısmi iyileştirme modeli, Cagan'ın uyumcu beklenti modeli ve gecikmesi dağıtılmış otoresif (ARDL) modelidir.

2.1.1. ARDL modeller

Uzun dönemdeki ilişkilerin sınanması maksadıyla yararlanılan farklı yaklaşımlar vardır. Bu yaklaşımlardan, Engle-Granger eşbütünleşme testi, Johansen eşbütünleşme testi ve sınır testi (Bounds Test) en çok kullanılanlarıdır.

Pesaran ve Shin [12], Pesaran [13] ve Pesaran vd. [14]'a göre ARDL yaklaşımı, durağanlık dereceleri aynı olmayan zaman serilerinde, eşbütünleşme analizinin yapılamamasından kaynaklanan problemi ortadan kaldırmak için geliştirilmiştir. Modelde, serilerin durağanlık derecelerinde meydana gelen farklılıkları yok ederek, hem uzun hemde kısa dönemde ilişkinin olup olmadığını incelenmesinde yararlanılmaktadır. Eğer çalışmadaki serilerin, bir ve daha fazlasının düzey halinde durağan olduğu görülürse yani $I(0)$ söz konusu ise eşbütünleşme ilişkisinde inceleme yapılamamaktadır. Ancak, ARDL yaklaşımı ile serilerin düzeyde ya da birinci farkta durağan olmaları göz ardı edilerek seriler arasındaki eşbütünleşme ilişkisine bakılabilir. Buna ek olarak, ARDL yaklaşımı, az gözlemlili veri kümeleri için kullanışlıdır [15].

ARDL yaklaşımı iki adımdan meydana gelmektedir. İlk adım, çalışmadaki seriler arasında uzun dönemde bir ilişkinin varlığını test edilmesidir. Sonraki adımsa kısa ve uzun dönemdeki parametrelerin üretilmesi ile tahminlerin gerçekleştirilmesidir. ARDL sınır testi yaklaşımı uygulanırken uyulması gereken en önemli hususlardan biri, ilgili değişkenlerin ikinci veya daha yüksek farklarda durağanlığının söz konusu olmamasıdır.

ARDL yaklaşımında, ilgili değişkenler arasında eşbütünleşmenin olup olmadığını araştırılması için ilk olarak değişkenler arasındaki ilişkilerin incelenmesi gerekmektedir. Bunun için kullanılacak modellerin, kısıtsız hata düzeltme modeli (KHDM) biçiminde kurulması ve EKK yöntemine dayanmakta olup, küçük hacimli örneklemedeki çalışmalarda kullanılabilme olanağı sağlamaktadır. ARDL sınır testi yaklaşımı, KHDM kullanıldığında, Engle-Granger nedensellik analizinden daha iyi istatistiki niteliktedir. Ayrıca, düşük sayıdaki gözlemlerde, Johansen ile Engle-Granger nedensellik analizlerinden daha güvenilir sonuç elde etmektedir [16].

KHDM, incelenilen modelin kısa ve uzun dönemli bileşenlerinin otokorelasyon problemlerine sahip olup olmadığının tahmin edilmesine olanak sağlar. Breusch-Godfrey LM testiyle kestirilmiş modelin yapısında mevcut bir otokorelasyon problemi olup olmadığı belirlenir. Wald testi ile eşbütünleşik bir ilişkinin varlığına karar verilir. Testin yokluk hipotezi "ele alınan değişkenler arasında eşbütünleşik ilişki yoktur" şeklindedir ve hem uzun hem de kısa döneme sahip parametreler arasında ilişki olup olmadığı sınanır.

Gecikmesi dağıtılmış modellerde bağımlı değişken ile bağımsız değişkenlerin gecikmeli değerlerinin modelde bağımsız değişken şeklinde bulunması durumu ARDL model olarak adlandırılmaktadır. ARDL model genel olarak aşağıdaki gibi gösterilmektedir:

$$y_t = \mu + \sum_{i=1}^p \alpha_i y_{t-i} + \sum_{j=0}^q \gamma_j x_{t-j} + \varepsilon_t; \quad t = 1, 2, \dots, n \quad (2)$$

Burada, p'ler y_t üzerindeki gecikmeleri, q'lar ise x_t üzerindeki gecikmeleri verir. Modelde y ve x sırasıyla, bağımlı ve bağımsız değişkeni temsil etmektedir. Eşitlik (2) açılarak yazıldığında,

$$y = \mu + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + \gamma_0 x_t + \gamma_1 x_{t-1} + \dots + \gamma_q x_{t-q} + \varepsilon_t \quad (3)$$

elde edilir. y_t ve x_t 'nin durağanlık düzeyleri y^* ve x^* olarak ifade edilirse,

$$y^* = \beta_0 + \beta_1 x^* \quad (4)$$

elde edilir. Burada, $x^* = x_t = x_{t-1} = \dots = x_{t-q}$ 'dir. Hata terimleri olmadan modelin uzun dönemdeki çözümü, Eşitlik (3)'de yerine konulduğunda,

$$y^* = \frac{\mu}{1 - \sum \alpha_i} + \frac{\sum \gamma_i}{1 - \sum \alpha_i} x^* \quad (5)$$

biçiminde verilir. Eşitlik (5) daha açık şekilde yazıldığında,

$$y^* = \frac{\mu}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_n} + \frac{(\gamma_1 + \gamma_2 + \dots + \gamma_m)}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_n} x^* \quad (6)$$

veya

$$y^* = B_0 + B_1 x^* \quad (7)$$

elde edilir. Modelin uzun dönemdeki sonucuna varılması $\sum_{i=1}^p \alpha_i < 1$ şartının sağlanmasıyla mevcuttur.

Pesaran vd. [14]'nin belirtildiği gibi, ARDL yaklaşımında uzun dönem ilişkisinin tespit edilip bağımsız değişkenlerin katsayı tahmini yapıldıktan sonra değişkenler arasında kısa dönem ilişkisini incelemek için KHDM'nin kurulması gerekmektedir. Sınır testi, temel olarak KHDM'nin EKK yöntemi ile tahmin edilmesine bağlıdır. Söz konusu KHDM aşağıda gösterilmektedir:

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \pi_{yx} x_{t-1} + \sum_{i=1}^{p-1} \varphi_i' \Delta z_{t-i} + \omega' \Delta x_t + \varepsilon_t \quad (8)$$

Modelde yer alan Δy_t ve Δx_t kısa dönem dinamiklerinin, π_{yy} ve π_{yxx} uzun dönemdeki parametresini, t trendi, c_0 sabit parametreyi ifade etmektedir. $c_0 = -(\pi_{yy}, \pi_{yxx})\mu + [\gamma_{yx} + (\pi_{yy}, \pi_{yxx})]\gamma$, $c_1 = (\pi_{yy}, \pi_{yxx})\gamma$ ve $\pi_{yxx} \equiv \pi_{yx} - \omega'\pi_{xx}$ olarak tanımlanmaktadır. ARDL sınır testinde, Eşitlik (8) kullanılarak beş farklı model elde edilmiştir.

Model 1: Sabitsiz ve trendsiz ($c_0 = 0$ ve $c_1 = 0$) ise;

$$\Delta y_t = \pi_{yy}y_{t-1} + \pi_{yxx}x_{t-1} + \sum_{i=1}^{p-1} \phi_i' \Delta z_{t-i} + \omega' \Delta x_t + \varepsilon_t \quad (9)$$

Model 2: Kısıtlı sabitli ve trendsiz ($c_0 = -(\pi_{yy}, \pi_{yxx})\mu$ ve $c_1=0$) ise;

$$\Delta y_t = \pi_{yy}(y_{t-1} - \mu_y) + \pi_{yxx}(x_{t-1} - \mu_x) + \sum_{i=1}^{p-1} \phi_i' \Delta z_{t-i} + \varepsilon_t \quad (10)$$

Model 3: Kısıtsız sabitli ve trendsiz ($c_0 \neq 0$ ve $c_1=0$) ise;

$$\Delta y_t = c_0 + \pi_{yy}y_{t-1} + \pi_{yxx}x_{t-1} + \sum_{i=1}^{p-1} \phi_i' \Delta z_{t-i} + \omega' \Delta x_t + \varepsilon_t \quad (11)$$

Model 4: Kısıtsız sabitli ve kısıtlı trendli ise;

$$\Delta y_t = c_0 + \pi_{yy}(y_{t-1} - \gamma_y t) + \pi_{yxx}(x_{t-1} - \gamma_x t) + \sum_{i=1}^{p-1} \phi_i' \Delta z_{t-i} + \omega' \Delta x_t + \varepsilon_t \quad (12)$$

Model 5: Kısıtsız sabitli ve kısıtsız trendli ise;

$$\Delta y_t = c_0 + c_1 t + \pi_{yy}y_{t-1} + \pi_{yxx}x_{t-1} + \sum_{i=1}^{p-1} \phi_i' \Delta z_{t-i} + \omega' \Delta x_t + \varepsilon_t \quad (13)$$

Denklemlerin temel hipotezi, $H_0^{\pi_{yy}}: \pi_{yy} = 0$ ve $H_0^{\pi_{yxx}}: \pi_{yxx} = 0'$ olup alternatif hipotezleri, $H_1^{\pi_{yy}}: \pi_{yy} \neq 0$ ve $H_1^{\pi_{yxx}}: \pi_{yxx} \neq 0'$ şeklinde ifade edilir.

Sınır testinin uygulanabilmesi için öncelikle, uygun gecikme uzunluğunun belirlenmesini sağlayan, Akaike bilgi kriteri (AIC), Schwarz bilgi kriteri (SIC) ve Hannan-Quinn (HQ) gibi bilgi kriterlerinden yararlanılmaktadır. En küçük değere sahip gecikme uzunluğu ele alınarak oluşturulan modelde otokorelasyon problemi olması durumunda bu model ele alınmayarak ikinci en küçük değerdeki gecikme uzunluğu seçilir. Buna rağmen otokorelasyon problemi

aynı şekilde devam ediyorsa, problem yok olana dek işlem sürdürülür. Eşbütünlüşme ilişkisinin olup olmadığı, bağımlı ile bağımsız değişkenlerin birinci gecikmesinin denklemdaki anlamlılığına tabidir.

Hipotezlerinin sınanması F testi veya Wald testi yardımıyla yapılır. Parametrelere ilişkin F istatistiği, tablo kritik değerleriyle karşılaştırılırken ilk olarak serilerin durağanlık derecelerinin benzer olmasına önem verilir [17].

Modeldeki değişkenlerden birisi düzeyde durağan, diğerlerinin birinci farkının durağan olmasıyla, tablo alt ve üst kritik değerleri aracılığıyla şayet değişkenlerin tümü düzeyde durağan olması durumdaysa, tabloda bulunan alt kritik değerleriyle kıyaslama yapılır. Aynı şekilde, değişkenler birinci farkta durağan ise üst kritik değerleriyle kıyaslama yapılır. Değişkenlerden birinin düzeyde, diğerlerinin birinci farkta durağan olması söz konusu olduğunda hesaplanan F istatistik değeri, üst kritik değer üzerindeyse seriler arasında eşbütünlüşme ilişkisinin olduğuna, alt kritik değer altındaysa eşbütünlüşme ilişkisinin olmadığına işaret eder. F istatistiğinin alt ile üst kritik değerlerin arasında bulunmasıyla, net karara varılmamakla birlikte alternatif yöntemlere başvurmak zorunda kalınır. Değişkenlerin düzeyde durağan olması ve F istatistiğinin tablo alt kritik değerinin üstünde kalması durumunda değişkenler arasında eşbütünlüşme ilişkisinin var olduğu, tersi halindeyse eşbütünlüşme ilişkisinin bulunmadığı sonucuna varılır. Ayrıca, değişkenlerin hepsinin birinci farkları alınıp durağan olmaları durumunda, F istatistiğinin yalnızca tablo üst kritik değerinden büyük olması ile değişkenler arasında eşbütünlüşme ilişkisinin olduğunu işaret eder [18].

2.1.2. Eşbütünlüşme testi

Eşbütünlüşme analizi, fark alma yoluyla değişkenler arasında kısa ve uzun dönem ilişkilerinde bilgi kaybının yaşanmaması açısından avantajlı bir yöntemdir. Bununla birlikte, her bir eşbütünlüşme serinin hata düzeltme modelinin (HDM) kurulabilmesiyle, uzun ve kısa dönem ilişkilerini fark etme olanağı sağlar.

Eşbütünlüşme analizi ikiye ayrılmaktadır. Engle ve Granger [19]'a göre, düzeyde durağan olmayıp birinci farkı alındığında durağan olan zaman serilerinin, düzey hallerinde modellenebilmesi ve bilgi kaybının yaşanmaması sağlanır. Fakat geliştirilen bu yaklaşım ile iki değişkene sahip bir seride, değişkenlerden birinin eşitliğinde eşbütünlüşme ilişkisi görülürken, diğer değişkenin eşitliğinde aynı şekilde bir ilişki gözlenmeyebilir. Bununla birlikte, değişken sayısının artmasıyla birden fazla uzun dönemli ilişkinin ortaya çıkabilmesi olasılığına karşılık Engle-Granger testi çoklu eşbütünlüşme yapısını ayırt edememektedir. Diğer eşbütünlüşme analizi ise, Johansen [20] ve Johansen ve Juselius [21]'nin vektör otoregresif (VAR) modellerine dayanarak birden fazla eşbütünlüşme

ilişkisinin araştırılması için elverişli olan bir eşbütünleşme testidir. Johansen eşbütünleşme yöntemi, olabilirlik oranlarını elde etmek için HDM'nin hesaplanmasını gerektirir.

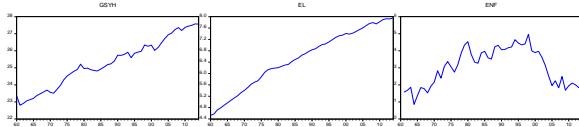
HDM aşağıdaki şekilde verilebilir:

$$\Delta y_t = \theta_0 + \sum_{i=1}^p \theta_i \Delta y_{t-i} + \sum_{j=0}^q \delta_j \Delta x_{t-j} + \lambda \hat{u}_{t-1} + \varepsilon_t \quad (14)$$

Burada θ_0 , sabit parametreyi; λ , hata düzeltme parametresini ifade etmekte olup $-1 < \lambda < 0$ 'dır. \hat{u}_{t-1} ise denge hata terimi veya hata düzeltme terimi olup, $\hat{u}_{t-1} = y_{t-1} - \beta_0 - \beta_1 x_{t-1}$ biçiminde tanımlanır. İncelenen bir modelde, deterministik bileşenlerde gözlemlenen tek bir farklılığın söz konusu olması durumunda Johansen yaklaşımıyla elde edilen sonuçlar farklılık göstermektedir [22].

III. SONUÇLAR VE TARTIŞMA

Bağımlı ve bağımsız değişkenlere logaritmik dönüşüm uygulandıktan sonra elde edilen çizgi grafikleri Şekil 1'de değişkenlere ait tanımlayıcı istatistik sonuçları Tablo 1'de verildiği gibidir.



Şekil 1. Değişkenlere ait çizgi grafik incelemesi

Tablo 1. Değişkenlerin tanımlayıcı istatistikleri ve normallik incelemesi

	GSYH	EL	ENF
Minimum	22,8054	4,5251	0,8554
Maksimum	27,5803	7,9566	4,9676
Ortalama	25,2761	6,5196	3,0460
Ortanca	25,2163	6,6478	3,1655
Standart Sapma	1,4118	1,0255	1,0846
Jarque Bera	2,5631	3,6309	4,0268
p değeri	0,2776	0,1627	0,1335

Tablo 1 incelendiğinde, tüm değişkenlere ait Jarque-Bera test istatistiklerine karşılık gelen p değerlerinin 0,05'ten büyük olmasıyla birlikte değişkenlerin normal dağıldığı söylenebilmektedir.

3.1. Birim Kök Testi

Çalışmada serilerin durağanlık düzeyleri, Dickey-Fuller [23] ile Phillips-Perron [24]'da verilen birim kök testleri ile analiz edilmiştir. Serilerin fark alınmamış halleri için test sonuçları Tablo 2'de verildiği şekildedir.

Tablo 2. Düzeyde ADF ve PP birim kök test sonuçları

		ADF Birim Kök Testi		PP Birim Kök Testi	
		Sabitli	Sabitli ve Trendli	Sabitli	Sabitli ve Trendli
GSYH	t değeri	0,0004	-3,4894	0,0004	-3,8320
	p değeri	0,9542	0,0507*	0,9542	0,0222**
EL	t değeri	-4,1339	-0,9404	-4,2394	-0,9267
	p değeri	0,0019***	0,9434	0,0014***	0,9452
ENF	t değeri	-1,6502	-1,3693	-1,6141	-1,1749
	p değeri	0,4503	0,8587	0,4684	0,9053

Not: ***, **, * sırasıyla %1, %5 ve %10 düzeyinde anlamlılığı ifade etmektedir.

Şekil 1 göz önünde bulundurularak Tablo 2 incelendiğinde, bağımlı değişken olan GSYH için test sonuçlarında serinin durağanlığının olmadığı söylenebilmektedir. Bağımsız değişkenler olan EL ve ENF için ADF ve PP test sonuçlarına bakıldığında ise sabitli ve trendli durumda değişkenlerin durağanlıklarının söz konusu olmadığı kanısına varılmaktadır. Serilerinin birinci farkı alınmış biçimleri için birim kök sonuçları Tablo 3'te verildiği gibidir.

Tablo 3. Birinci farkta ADF ve PP birim kök test sonuçları

		ADF Birim Kök Testi		PP Birim Kök Testi	
		Sabitli	Sabitli ve Trendli	Sabitli	Sabitli ve Trendli
GSYH	t değeri	-8,5267	-8,3846	-8,5267	-8,3847
	p değeri	0,0000***	0,0000***	0,0000***	0,0000***
EL	t değeri	-4,8929	-6,2512	-4,9290	-6,1993
	p değeri	0,0002***	0,0000***	0,0002***	0,0000***
ENF	t değeri	-7,6932	-7,9439	-7,7037	-10,2635
	p değeri	0,0000***	0,0000***	0,0000***	0,0000***

Not: ***, **, * sırasıyla %1, %5 ve %10 düzeyinde anlamlılığı ifade etmektedir.

Tablo 3'e göre, bağımlı ve bağımsız değişkenlerin sabitli ile sabitli ve trendli durumlarının p değeri sonuçlarına bakıldığında, tüm değerlerin 0,05'ten küçük olduğu gözlemlenmektedir. Böylece değişkenlerin birinci farkta durağan olduğu söylenebilmektedir.

3.2. Eşbütünleşme Testi

Eşbütünleşme testinde, gecikme sayısının belirlenebilmesi için yapılan analiz Tablo 4'te verildiği gibidir.

Tablo 4. Gecikme sayısının tespit edilmesi

Lag	LogL	LR	FPE	AIC	SIC	HQ
0	108,2033	NA	0,0228	4,7321	4,8501	4,7765
1	101,8439	384,3415	4,39e-06	-3,8231	-3,3508*	-3,6454*
2	105,6155	6,4198	5,52e-06	-3,6007	-2,7740	-3,2896
3	118,6097	20,4589	4,72e-06	-3,7706	-2,5897	-3,3262
4	128,0745	13,6937	4,74e-06	-3,7904	-2,2552	-3,2127
5	143,3829	20,1941*	3,77e-06*	-4,0588*	-2,1693	-3,3478
6	147,6934	5,1359	4,90e-06	-3,8593	-1,6155	-3,0149
7	159,5620	12,6262	4,75e-06	-3,9814	-1,3833	-3,0037

Tablo 4'e göre Likelihood Ratio (LR), Final Prediction Error (FPE), AIC ölçütleri incelendiğinde gecikme sayısı 5 olarak belirlenmiştir. SIC ve HQ bilgi kriterleri gecikme sayısını 1 olarak saptamasına rağmen daha kapsamlı olan 5 gecikme modeli ele alınmıştır. Sonrasında, sınır testi yaklaşımıyla seriler arasında eşbütünleşme ilişkisinin varlığına ilişkin sonuçlar Tablo 5'de verildiği gibidir.

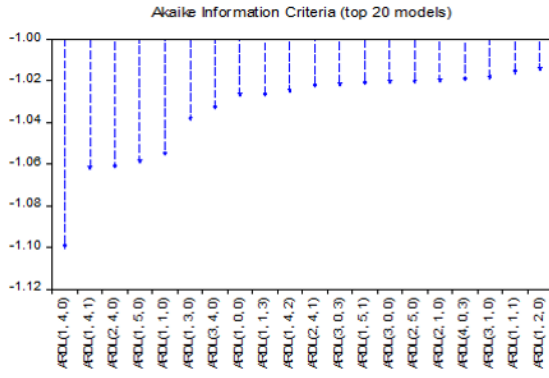
Tablo 5. Sınır testi sonuçları

Bağımsız Değişken Sayısı	F İstatistiği	%5 Anlamlılık Düzeyinde Kritik Değerler	
		Alt Sınır	Üst Sınır
2	5,1808	3,79	4,85

Değişkenler arasında eşbütünlük ilişkisini tespit etmek amacıyla F istatistik değeri ile alt ve üst sınır değerlerinin bulunduğu Tablo 5'e bakıldığında, hesaplanan F istatistik değeri, üst sınır değerini aştığından dolayı seriler arasında eşbütünlük ilişkisinin varlığı belirlenmiştir. Bu sebepten kısa ve uzun dönem ilişkilerinin saptanmasında ARDL modeli oluşturulabilir.

3.3. ARDL Modeli

Çalışmada serilerin öncelikle sabitli daha sonrasında sabitli ve trendli modeli incelenerek en iyi model ele alınmıştır. Şekil 2'de verilen grafiğe göre denenen 20 alternatif model arasından en iyisinin en küçük AIC değerine sahip ARDL(1,4,0) olduğu saptanmıştır.

**Şekil 2.** AIC için 20 en iyi model

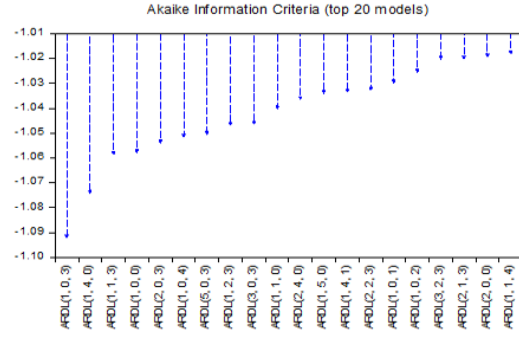
Şekil 2'ye göre, ARDL(1,4,1) modeli ile ARDL(2,4,0) modelinin birbirine yakın olduğu ama ARDL(1,4,0) modelinin diğer modellere nazaran daha büyük bir farkla en iyisi olduğu görülmektedir. ARDL(1,4,0) için sabitli model sonuçları Tablo 6'da sunulduğu gibidir.

Tablo 6. ARDL(1,4,0) sabitli model sonuçları

Değişkenler	Katsayı	t istatistiği	p değeri
GSYH(-1)	0,6769	6,3100	0,0000***
EL	1,8594	2,9860	0,0047***
EL(-1)	-1,4378	-1,4417	0,1566
EL(-2)	0,8110	0,8569	0,3962
EL(-3)	0,4977	0,5567	0,5806
EL(-4)	-1,2098	-2,1447	0,0377**
ENF	-0,0459	-2,0349	0,0481**
C	4,7526	2,6383	0,0116**
Durbin-Watson	1,9134		

Not: ***, **, * sırasıyla %1, %5 ve %10 düzeyinde anlamlılığı ifade etmektedir.

Tablo 6'daki ARDL modelinin sonuçları incelendiğinde, GSYH(-1), EL, EL(-4), ENF değişkenleri ile sabit p değerlerinin 0,05'ten küçük ve anlamlı olduğu gözlenmektedir. Durbin-Watson değerinin 2'ye yakın olmasından dolayı, hata terimi için birinci dereceden otokorelasyon probleminin olmadığı kanısına varılmıştır. Şekil 2'de sabitli ve trendli model sonuçları verilmektedir.

**Şekil 3.** AIC için 20 en iyi model

Şekil 3'e göre, AIC için 20 en iyi model arasından ARDL(1,0,3) modelinin en iyi sonucu verdiği gözlenmektedir. ARDL(1,0,3) için sabitli ve trendli model sonuçları Tablo 7'de sunulduğu gibidir.

Tablo 7. ARDL(1,0,3) sabitli ve trendli model sonuçları

Değişkenler	Katsayı	t istatistiği	p değeri
GSYH(-1)	0,5637	6,0018	0,0000***
EL	1,3175	4,1091	0,0002***
ENF	-0,0726	-1,5274	0,1340
ENF(-1)	0,0066	0,1209	0,9043
ENF(-2)	0,0178	0,3227	0,7485
ENF(-3)	-0,0942	-2,1539	0,0369**
C	3,9869	1,9878	0,0532*
TREND	-0,0404	-2,3245	0,0249**
Durbin-Watson	1,8803		

Not: ***, **, * sırasıyla %1, %5 ve %10 düzeyinde anlamlılığı ifade etmektedir.

Tablo 7'de sunulan sabitli ve trendli model sonuçları incelendiğinde, GSYH(-1), EL, ENF(-3) ve trendin p değerlerinin 0,05'ten küçük olduğu gözlenmektedir. Durbin-Watson değeri 2'ye yakın olduğundan, hata terimi için birinci dereceden otokorelasyon problemi olmadığı kanısına varılmıştır. Şekil 2 ve Şekil 3'e bakıldığında, en küçük AIC değerine sahip olan sabitli ARDL(1,4,0) modelinin uzun dönem tahmini incelenmiştir.

3.4. ARDL Uzun Dönem Modeli

İki en iyi model arasından AIC'si daha küçük olan ARDL(1,4,0) modelinin tahmin sonuçlarına göre hesaplanan eşbütünlük ilişkisi ile uzun dönem katsayıları Tablo 8'de verildiği gibidir.

Tablo 8. ARDL(1,4,0) eşbütünlük inceleme

Değişkenler	Katsayı	t istatistiği	p değeri
C	4,7526	2,6383	0,0116**
GSYH(-1)	-0,3231	-3,0116	0,0043***
EL(-1)	0,5205	3,4215	0,0014**
ENF*	-0,0459	-2,0349	0,0481**
D(EL)	1,8994	2,9860	0,0047***
D(EL(-1))	-0,0989	-0,1674	0,8678
D(EL(-2))	0,7120	1,2123	0,2320
D(EL(-3))	1,2097	2,1447	0,0377**

Not: ***, **, * sırasıyla %1, %5 ve %10 düzeyinde anlamlılığı ifade etmektedir.

*: Değişken, Z=Z(-1) + D(Z) olarak yorumlanabilir.

Tablo 8'e bakıldığında, D(EL(-1)) ve D(EL(-2)) değişkenleri haricindeki değişkenlerin p değerleri 0,05'ten küçük ve anlamlı oldukları görülmektedir.

ARDL(1,4,0) için uzun dönem katsayıları Tablo 9’da verildiği gibidir.

Tablo 9. ARDL(1,4,0)’ın uzun dönem katsayıları

Değişkenler	Katsayı	t istatistiği	p değeri
EL	1,6112	14,4542	0,0000***
ENF	-0,1422	-2,6387	0,0115***

EC=GSYH – (1,6112*EL-0,1422*ENF)

Tablo 9 incelendiğinde, GSYH ile EL arasında ve GSYH ile ENF arasında istatistiksel açıdan anlamlı ilişkide olduğu söylenebilmektedir. Böylelikle, uzun dönemde EL’de oluşacak 1 birimlik artışın GSYH’yi ortalama 1,6112 birim artırması beklenmektedir. Ayrıca, uzun dönemde ENF’de oluşacak 1 birimlik artışın GSYH’i ortalama 0,1422 birim azaltması beklenmektedir.

3.5. Hata Düzeltme Modeli

Modeldeki kısa dönem dinamiklerini gözlemlemenin diğer gerekli yolu HDM’nin katsayısının bulunmasıdır. Bu amaçla elde edilen HDM sonuçları Tablo 10’da sunulduğu gibidir.

Tablo 10. ARDL(1,4,0) yaklaşımına dayalı hata düzeltme modeli sonuçları

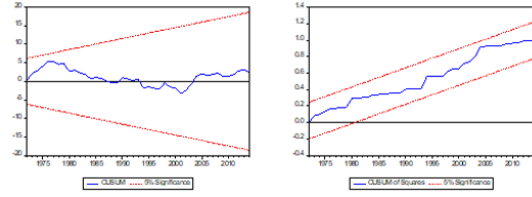
Değişkenler	Katsayı	t istatistiği	p değeri
C	4,7526	3,9872	0,0003***
D(EL)	1,8594	3,5393	0,0010***
D(EL(-1))	-0,0989	-0,1739	0,8628
D(EL(-2))	0,7120	1,3065	0,1983
D(EL(-3))	1,2098	2,2183	0,0319**
CointEq(-1)	-0,3231	-4,0330	0,0002***

Not: ***, **, * sırasıyla %1, %5 ve %10 düzeyinde anlamlılığı ifade etmektedir.

Narayan ile Smyth [25], hata düzeltme (CointEq(-1)) değişken katsayısının 1’den büyük olması durumunda sistemin dalgalanarak dengeye geleceğini ve dalgalanmanın her defasında azalarak uzun dönemde dengeye dönüşeceğini belirtmişlerdir. Bundan dolayı, hata terimi katsayısının kısa dönem ilişki için -1 ile 0 arasında olması beklenmektedir. Tablo 10’a bakıldığında, hata teriminin katsayısı -0,3231 olarak bulunmuştur. Böylece hata terimi katsayısının beklendiği gibi negatif ve istatistiksel olarak anlamlı çıktığı gözlenmektedir. Bunun yanında, istatistiksel olarak D(EL(-1)) ve D(EL(-2)) hariç diğer değişkenlerin anlamlı olduğu söylenebilmektedir.

3.6. CUSUM ve CUSUM Kare Testleri

Modelde kısa dönem dinamiklerine dair hata düzeltme terimini elde etmek için kullanılan ve uzun dönem katsayılarının kararlılığın ölçülmesi için Brown vd. [26] aracılığıyla öne sürülen CUSUM testleri kararlılık durumuna karşın etkin sonuçlar vermektedir [27]. Bu amaçla elde edilen CUSUM ve CUSUM Kare grafikleri Şekil 4’te verildiği gibidir.



Şekil 4. CUSUM ve CUSUM kare grafikleri

Şekil 4’e göre, CUSUM test istatistiklerinden elde edilen eğri, %5 anlamlılığı işaret eden kritik doğrular içinde yer almaktadır. Bundan dolayı, tahmin edilen katsayıların uzun dönemde kararlı olduğu gözlenmektedir. CUSUM kare testinin CUSUM testinden farkı daha duyarlı olması olup ardışık artıkların karelerinin hesaplanmasına dayanır. CUSUM kare grafiği incelendiğinde, %5 anlamlılık düzeyinde eğrinin kritik doğrular içerisinde kaldığı ayrıca uzun dönemde yapısal kırılmaya uğramadığı söylenebilmektedir.

3.7. Kurulan Model İçin Varsayımların İncelenmesi

Kurulan modelin etkin parametre tahminlerine sahip olup olmadığını belirlenmesi ve kullanılan test istatistiklerinin geçerliliği için yapılan artık analizi sonuçları Tablo 11’de verildiği gibidir.

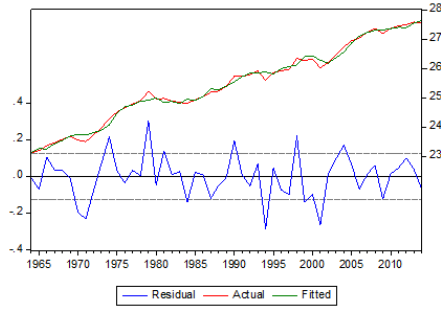
Tablo 11. Model varsayımlarının incelenmesi

Uygulanan Testler	Test İstatistiği (p değeri)
Jarque-Bera	0,5421 (0,7626)
Breusch-Godfrey LM	9,4062 (0,0939)
Breusch-Pagan-Godfrey	11,6734 (0,1118)
Ramsey Reset	0,0985 (0,7552)
Pearson Korelasyon	r=0,2949 (0,0304)

Tablo 11’e göre, hata teriminin normalliğinin incelendiği Jarque-Bera testi için p değeri 0,7626 olup 0,05’ten büyük olduğundan normal dağılım varsayımının sağlandığı söylenebilmektedir. Tanısal test sonuçlarından olan Breusch-Godfrey LM testi ile hatalar arasında otokorelasyon probleminin varlığı incelenmiştir. Testin p değeri 0,0939 olup 0,05’ten büyük olduğu görülerek, hatalar arasında otokorelasyon yoktur hipotezinin reddedilemeyeceği söylenebilmektedir. Hatalar için değişen varyanslılıkla ilgili bir test olan Breusch-Pagan-Godfrey incelendiğinde, p değeri 0,1118 olup 0,05’ten büyük olduğu görülmektedir. Böylece, değişen varyanslılık yoktur ifadesinin reddedilemeyeceği söylenebilmektedir. Model tanımlama hatasının varlığını test etmek için Ramsey Reset testi kullanılmıştır. Testin p değeri 0,7552 olup 0,05’ten büyük olduğu görülmektedir. Böylece, modelin tanımlanması ile ilgili bir hata olmadığı söylenebilir. Ayrıca, bağımsız değişkenler olan EL ve ENF arasındaki ilişki incelenmiş ve bu ilişkinin zayıf bir ilişki (r=0,2949) olduğu saptanmıştır. Dolayısıyla bu modelde çoklu doğrusal bağlantı problemine rastlanılmadığı söylenebilir.

3.8. Artık Grafiğinin İncelenmesi

Kurulan modelden elde edilen tahmin değerlerinin gözlenen değerlerle uyumunun ve artıklarının sunulduğu grafik Şekil 5'te verildiği gibidir.



Şekil 5. Artık grafiği

Şekil 5'e göre, gözlenen değerler ile tahmin edilen değerlerin birbirlerine çok yakın olduğu ve her iki serinin çizgi grafiklerinin de birbirleriyle örtüştüğü görülerek başarılı bir model kurulmasının gerçekleştiği söylenebilmektedir.

IV. SONUÇ

Bu çalışmada, 1960-2014 dönemlerinde Türkiye'de GSYH, EL ve ENF arasındaki ilişkiler ARDL yaklaşımı yardımı ile analiz edilmiştir. Bu çerçevede, değişkenlerin durağanlık incelenmesi için ADF ve PP testlerinden faydalanılmış olup test sonuçlarına bakıldığında değişkenlerin birinci farklarında durağanlığının söz konusu olduğu, ayrıca hiçbir değişkenin iki veya daha yüksek farkında durağan olmadığı sonucuna varılmıştır. Bundan dolayı, çalışmadaki değişkenler arasında eşbütünlüşme ilişkisinin varlığı araştırılmış ve eşbütünlüşme ilişkisinin olduğu gözlemlenmiştir. Eşbütünlüşme ilişkisinin varlığı sebebiyle, değişkenler arasında kısa ve uzun dönem ilişkilerinin bulunması amacıyla Pesaran vd. [14]'nin geliştirdikleri ARDL yaklaşımı uygulanmıştır. Uzun dönemde GSYH ile EL değişkenleri arasında pozitif ve istatistiksel açıdan anlamlı bir ilişkide olduğu görülmektedir. Bu bakımdan, uzun dönemde EL'de olası bir artışın GSYH'yi artıracakını işaret etmektedir. Ayrıca, uzun dönemde GSYH ile ENF değişkenleri arasında negatif ve istatistiksel açıdan anlamlı bir ilişkide olduğu görülmektedir. Böylece uzun dönemde ENF'de olası bir artışın GSYH'yi azaltacağı görülmektedir. Uzun dönem ARDL katsayılarının incelenmesinden sonra, HDM'nin katsayısının negatif ve anlamlı çıktığı sonucuna ulaşılmıştır. HDM'nin bulunmasıyla yapısal kırılma olup olmadığı incelenmiştir. Bu kapsamda, CUSUM ve CUSUM kare grafiklerine bakıldığında yapısal kırılmanın olmadığı görülmüştür. Son kısımda, model varsayımları ve artık grafiği incelenmiştir. Bu kapsamda, model ile ilgili herhangi bir varsayım bozulmasının görülmediği ve gözlenen ile tahmin edilen değerlerin tam bir uyum içinde olduğu sonucuna varılmıştır.

Sonuç olarak, uzun dönemde elektrik tüketiminde meydana gelecek artışın ekonomik büyümeyi artırarak olumlu yönde etkileyeceği ve enflasyonda meydana gelecek artışın ekonomik büyümeyi olumsuz etkileyeceği söylenebilir.

TEŞEKKÜR

Bu çalışma, Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü İstatistik Bölümünde hazırlanmış "Gecikmesi Dağıtılmış Otoresgresif Model (ARDL) Yaklaşımı İle Türkiye'de Ekonomik Büyüme, Elektrik Tüketimi ve Enflasyon Arasındaki İlişkinin İncelenmesi" adlı yüksek lisans tezinden üretilmiştir.

KAYNAKLAR

- [1] Halıcıoğlu, F. (2009). An econometric study of CO2 emissions, energy consumption, income and foreign trade in Turkey. *Energy Policy*, 37, 1156-1164.
- [2] Öztürk, I. ve Acaravcı, A. (2010). CO2 Emissions, energy consumption and economic growth in Turkey. *Renewable and Sustainable Energy Reviews*, 14(9), 3220-3225.
- [3] Polat, Ö., Enes, E. ve San, A. G. S. (2011). Türkiye'de elektrik tüketimi, istihdam ve ekonomik büyüme ilişkisi. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 16(1), 349-362.
- [4] Fuinhas, J. A. ve Marques, A. C. (2012). Energy consumption and economic growth nexus in Portugal, Italy, Greece, Spain and Turkey: An ARDL bounds test approach(1965-2009). *Energy Economics*, 34(2), 511-517.
- [5] Öcal, O. ve Aslan, A. (2013). Renewable energy consumption and economic growth Nexus in Turkey. *Renewable and Sustainable Energy Reviews*, 28, 494-499.
- [6] Koçak, E. (2014). Türkiye'de çevresel Kuznets eğrisi hipotezinin geçerliliği: ARDL sınır testi yaklaşımı. *İşletme ve İktisat Çalışmaları Dergisi*, 2(3), 62-73.
- [7] Özçağ, M. (2015). Türkiye'de enerji tüketimi, ekonomik büyüme ve dışa açıklık ilişkisi: ARDL modeli. *Finans Politik & Ekonomik Yorumlar*, 52(605), 7-16.
- [8] Pata, U. K. ve Terzi, H. (2017). Türkiye'nin iktisadi büyümesinde turizm sektörünün katkısı. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 48, 45-64.
- [9] Türemez, Y. ve Göktaş, D. (2018). Türkiye'ye yönelik Avrupa Birliği ülkeleri turist talebinin eşbütünlüşme analizi. *ABMYO Dergisi*, 51, 51-66.
- [10] Terzi, H. ve Bekar, S. (2019). Türkiye'de doğrudan yabancı yatırımlar, turizm ve dışa açıklık arasındaki ilişki: 1974-2014 dönemi. *Doğuş Üniversitesi Dergisi*, 20 (1), 15-30.
- [11] <https://data.worldbank.org/> (Erişim yılı:2019)
- [12] Pesaran, M. H. ve Shin, Y. (1995). Long-run structural modelling. *Unpublished manuscript*, University of Cambridge.

- [13] Pesaran, M. H. (1996). Cointegration and speed of convergence equilibrium. *Journal of Econometrics*, 71(1-2), 117-143.
- [14] Pesaran, M. H., Shin, Y. ve Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, 16, 289- 326.
- [15] Narayan, S. ve Narayan, P. K. (2004). Determinants of demand of Fiji's exports: An empirical investigation. *The Developing Economics*, 17(1), 95-112.
- [16] Narayan, P. K. ve Narayan, S. (2005). Estimating income and price elasticities of imports for Fiji in a cointegration framework. *Economic Modelling*, 22, 423 - 438.
- [17] Karagöl, E., Erbaykal, E. ve Ertuğrul, H. M. (2007). Türkiye'de ekonomik büyüme ile elektrik tüketimi ilişkisi: Sınır testi yaklaşımı. *Doğuş Üniversitesi Dergisi*, 8(1), 72-80.
- [18] Yamak, N. ve Tanrıöver, B. (2007). Türkiye'de nominal faiz oranı-genel fiyat düzeyi ilişkisi: Gibson paradoksu, 8. *Türkiye Ekonometri ve İstatistik Kongresi*, 1-13.
- [19] Engle, R. F. ve Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation and testing. *Econometrica*, 55(2), 251-276.
- [20] Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2-3), 231-254.
- [21] Johansen, S. ve Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, 52(2), 169-210.
- [22] Ahking, F. W. (2002). Model mis-specification and Johansen's co-integration analysis: an application to the US money demand. *J Macroecon*, 24, 51-66.
- [23] Dickey, D. ve Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of American Statistical Association*, 74, 427-431.
- [24] Phillips, P. C. B. ve Perron, P. (1988). Testing for A unit root in time series regression. *Biomètrika*, 75(2), 336-346.
- [25] Narayan, P. K. ve Smyth, R. (2006). What determines migration flows from lowincome to high-income countries? An empirical investigation of Fiji-US migration 1972-2001. *Contemporary Economic Policy*, 24(2), 332- 342.
- [26] Brown, R. L., Durbin, J. ve Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society*, 37(2), 149-192.
- [27] Altıntaş, H. (2008). Türkiye'de para talebinin istikrarı ve sınır testi yaklaşımıyla öngörülmesi: 1985-2006. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 30, 16-46.

Türkiye’de Zeytin Fiyatı ile Zeytin Üretim Alanı ve Üretim Miktarı Arasındaki İlişkinin Toda-Yamamoto Testi ile Belirlenmesi

Determination of the Relationship between Olive Price and Olive Production Area and Production Amount in Turkey by Toda-Yamamoto Test

Kaan KAPLAN¹, Halil KIZILASLAN¹

¹Tokat Gaziosmanpaşa Üniversitesi, Ziraat Fakültesi, Tarım Ekonomisi Bölümü, Tokat, Türkiye

Öz

Dünyada yaklaşık 30 farklı ülkede yetiştirilen zeytin, ülkelerin dış ticaretleri açısından önem arz etmektedir. Bu ticaretin etkileri Akdeniz havzasındaki ülkelerin zeytin üretimini tetiklemektedir. Bu sebeple uluslararası zeytin fiyatları ticarete önemli rol oynamaktadır ve Türkiye zeytin üretiminde bu havzada önemli bir konumdadır. Bu amaçla 1991-2022 yılları arasındaki Türkiye’deki zeytin ürün fiyatı (\$/ton), zeytin üretim alanı (ha) ve zeytin üretim miktarı (ton) arasındaki ilişkinin tespit edilmesi amaçlanmıştır. Çalışmada öncelikle serilerin durağanlığının test edilmesi amacıyla birim kök testlerine bakılmış ve serilerin hem düzeyde hem birinci farkta durağan olduğu tespit edilmiştir. Farklı seviyelerde durağanlığın olması sebebiyle uygun yöntem olarak Toda-Yamamoto nedensellik testi belirlenmiştir. Toda-Yamamoto nedensellik testinin yapılabilmesi için uygun VAR modeli Akaike Bilgi Kriteri (AIC) göz önünde bulundurularak oluşturulmuştur. VAR Modeli gecikme uzunluğu Schwarz Bilgi Kriteri (SC) hariç diğer bilgi kriterleri göz önünde bulundurularak 2 gecikme uzunluğu olarak yeniden belirlenmiştir. Modelin geçerliliğini sınamak adına öncelikle AR Karakteristik Polinomlarının ters köklerine bakılmış, daha sonrasında ise LM Otokorelasyon Testi ve White Değişen Varyans Testleri uygulanmıştır. Buna göre oluşturulan VAR Modelinin geçerlilik sınamalarına göre LM Otokorelasyon Test İstatistiği olasılık değeri 0.0829, White Değişen Varyans olasılık değeri ise 0.0035 olarak tespit edilmiştir. Toda-Yamamoto Nedensellik testi sonuçlarına göre ise ürün fiyatı ile üretim alanı arasındaki wald test 10.29723, olasılık değeri 0.0162 olarak tespit edilmiş, üretim fiyatı ile ekili alan arasındaki wald test 10.69738 ve olasılık değeri 0.0135 olarak bulunmuştur. Tüm sonuçlara göre zeytin ürün fiyatı (\$/ton) ile zeytin üretim alanı (ha) ve zeytin üretim miktarı (ton) arasında bir nedensellik ilişkisinden söz edilmektedir.

Anahtar Kelimeler: Zeytin Fiyatı, Toda-Yamamoto, VAR Modeli, Türkiye

Abstract

Olives, which are grown in about 30 different countries around the world, are important for the foreign trade of countries. The effects of this trade trigger the olive production of the countries in the Mediterranean basin. For this reason, international olive prices play an important role in trade and Turkey has an important position in olive production in this basin. For this purpose, it is aimed to determine the relationship between olive product price (\$/ton), olive production area (ha) and olive production amount (tons) in Turkey between 1991-2022. In the study, firstly, unit root tests were used to test the stationarity of the series and it was found that the series were stationary both at level and at first difference. Due to the different levels of stationarity, Toda-Yamamoto causality test was determined as the appropriate method. In order to conduct the Toda-Yamamoto causality test, the appropriate VAR model was constructed by considering the Akaike Information Criterion (AIC). The lag length of the VAR model was re-determined as 2 lags by taking into account the information criteria other than the Schwarz Information Criterion (SC). In order to test the validity of the model, firstly, the inverse roots of the AR Characteristic Polynomials are examined, and then the LM Autocorrelation Test and White Variance Tests are applied. According to the validity tests of the VAR Model, the probability value of the LM Autocorrelation Test Statistic is 0.0829 and the probability value of the White Changing Variance is 0.0035.

Keywords: Olive Price, Toda-Yamamoto, VAR Model, Türkiye

I. GİRİŞ

Zeytin Ağacı, Akdeniz Havzası’nda ve dünya genelinde bilinen en eski ekili ağaçlardan biridir [1]. Zeytin ağacı, Oleaceae familyasına ait bir bitkidir ve yaklaşık olarak 20-29 ana cins içerisinde yer alır. Olea cinsi içindeki tek yenilebilir meyve olan zeytin ağacının tarımının Akdeniz havzasında başladığı bilinmektedir [2]. Türkiye’nin

güneyi, Lübnan, Suriye, Filistin ve İsrail’de eski yazılı tabletlerde zeytin çukurlarından bahsedilmekte ve antik mezarlarda bulunan odun parçaları, zeytin meyvesinin kökeninin bu bölgede olduğunu göstermektedir [3].

Zeytin, dünyada yaklaşık 30 ülkede ticari olarak yetiştirilmektedir. Özellikle Akdeniz havzası en fazla zeytin ağacına sahip bölge konumundadır ve bu sebeple bu havzadaki ülkeler için öneme sahiptir. Örneğin yarım milyon kadar çiftçi İspanya’da zeytin yetiştiriciliği yapmaktadır. Zeytin ağacı geçmişten günümüze tüm Akdeniz havzasına yayılmış durumdadır ve günümüzde Tunus’ta da üretilmektedir. Tunus’ta zeytin yetiştiriciliğinden gelir elde eden nüfus, ülke nüfusunun onda birinden fazlasını oluşturmaktadır. Küresel zeytinyağı pazarı 2021 yılında 13.8 milyar Dolar değerindeyken, 2029 yılında bu değer 17.8 milyar ABD Dolarına ulaşması beklenmektedir [4].

Dünyada 2022 yılı verilerine göre 10.948.521 hektarlık alanda zeytin üretimi yapılmaktadır. Yine 2022 yılı verilerine göre 21.449.867 ton üretim gerçekleştirilmiştir [5]. Zeytin üretimi haricinde zeytinyağı da önemli bir konudur. Genel olarak zeytinyağı, Akdeniz ülkelerinde üretilmekte ve tüketilmektedir. Ancak son yıllarda oluşan trendler, sağlıklı ve dengeli beslenme konuları ile birlikte geleneksel olmayan bir pazarda Avrupa dışında büyümektedir. Güney Amerika’da Şili, Arjantin, Uruguay gibi ülkelerde, Okyanusya’da Avustralya, Yeni Zelanda gibi ülkelerde ve Kuzey Amerika’da Amerika Birleşik Devletleri gibi ülkelere yayılmıştır.

2021/2022 döneminde dünya sofralık zeytin ithalatı 767.500 ton olarak gerçekleşmiştir. Sofralık zeytinde dünya ithalatının yaklaşık %60’dan fazlası ABD, Brezilya, AB, Kanada ve Rusya tarafından talep edilmektedir. Türkiye, sofralık zeytin ithalatı olmayan ülke konumundadır.

Türkiye, zeytin üretimi için uygun iklim koşullarına sahip olması nedeniyle dünya genelinde önemli bir konuma sahiptir hem sofralık zeytin hem de zeytinyağı üretimi açısından. Food and Agricultural Organization of the United Nations (FAOSTAT) verilerine göre, zeytin üretim alanı 1991 yılında 543.467 hektardan 2022 yılında 901.126 hektara yükselmiştir.

Yapılan literatür taramasında çalışma konusu ile ilgili birtakım araştırmalar bulunmaktadır. Şengül, “Türkiye’de Sofralık Zeytin Fiyatlarındaki Dalgalanmalar: ARIMA-GARCH Yaklaşımıyla Volatilité Araştırması” isimli çalışmada Ocak 2008 – Aralık 2022 dönemine ait verileri kullanmış ve 9 aylık dönemde 2 aylık tahminler orta vadeli planlamalar için güvenilir sonuçlar verdiğini ortaya koymuştur [6]. Kılıç

ve Turhan ise Türkiye’de sofralık zeytin yağı ihracatının ve ihracat değerinin 1982-2019 yılları arasındaki uzun ve kısa dönem etkilerini incelemek amacıyla ARDL modeli uygulamışlardır. Çalışma sonuçlarına göre hata düzeltme katsayısı anlamlı bulunmuş ve ihracat miktarının ihracat değeri üzerinde etkisi olduğu sonucuna ulaşılmıştır [7]. Bulut ve arkadaşları elektrik tüketiminin ekonomik büyüme üzerindeki etkilerini araştırmak amacıyla Toda-Yamamoto nedensellik testi kullanmışlardır ve test sonuçlarına göre elektrik tüketiminden büyümeye doğru bir nedensellik sonucuna ulaşılmıştır [8]. Özçelik, Rusya-Ukrayna Savaşı süresince dünyadaki gıda fiyatlarının durumunu analiz etmek amacıyla Fourier Bootstrap ARDL ve Fourier Bootstrap Toda-Yamamoto testlerinden yararlanmışlardır [9]. Okur ve Çiçek, Türkiye’de kırmızı et fiyatları ile besi yemi fiyatları arasındaki nedensellik ilişkisinin belirlenmesi amacıyla Toda-Yamamoto nedensellik ilişkisinden faydalanmışlardır. Çalışmada Ocak 2008 – Aralık 2022 dönemine ait veri setiyle çalışılmış ve besi yemi fiyatlarından kırmızı et fiyatlarına doğru %1 anlamlılık düzeyinde nedensellik ilişkisi tespit edilmiş ancak kırmızı et fiyatlarından besi fiyatlarına doğru anlamlı bir ilişki tespit edilememiştir [10].

Bu çalışmanın sonraki bölümleri şu şekilde düzenlenmiştir. Öncelikle 2. Başlıkta çalışmanın ana materyali ve çalışmada kullanılan yöntemlerin detaylı açıklamasına yer verilmiştir. 3. Başlıkta ise araştırma bulgularına yer verilmiş ve öncelikle serilerin durağanlığının test edildiği birim kök testlerine yer verilmiştir. Daha sonrasında Toda-Yamamoto nedensellik testinin yapılabilmesi için uygun VAR modeli oluşturulmuştur. VAR modelinin gecikme uzunluğu belirlenmiş, uygun gecikme uzunluğu ile model yeniden tespit edilmiştir ve gecikme uzunluğu Tablo 2’de verilmiştir. Modelin geçerliliği için gerekli sınamalar yapıldıktan sonra Toda-Yamamoto nedensellik testi için gerekli olan Wald Testinin yapılabilmesi adına Görünürde İlişkisiz Regresyon Modeli oluşturulmuş ve Toda-Yamamoto nedensellik analizine yer verilmiştir. Zeytin fiyatları ve üretimine yönelik daha önce böyle bir çalışmanın yapılmamış olması sebebiyle literatüre ışık tutacağı düşünülmektedir

II. MATERYAL VE METOD

2.1. Materyal

Çalışmada zeytin fiyatları ile zeytin üretim miktarı ve zeytin üretim alanı arasındaki ilişkiyi belirlemek amacıyla 1991 – 2022 yılları arasındaki veri seti kullanılmıştır. Veriler FAOSTAT [11] kayıtlarından elde edilmiştir. Ürün fiyatlarında belirli aralıklarla eksik olması sebebiyle 1991 sonrası tercih edilmiştir.

2.2. Yöntem

Çalışmanın yöntem kısmında öncelikle uygun model seçiminin yapılması amacıyla belirlenen değişkenlerin analizleri yapılmıştır. Buna göre zeytin fiyatları, zeytin üretim alanı ve zeytin üretim miktarı değişkenlerinin durağanlık yapısı incelenmiştir. Seriler farklı seviyelerde durağanlık içermektedirler ancak çalışma için seçilen model seriler arasında aynı seviyede durağanlık koşulu aramamaktadır. Durağanlık yapısının ortaya koymak için Genişletilmiş Dickey-Fuller (ADF) [12] ve Phillips-Perron (PP) [13] birim kök testlerinden yararlanılmıştır.

Toda ve Yamamoto'nun [14] nedensellik testi, klasik nedensellik testlerinden farklı olarak birkaç önemli özellik taşır. İlk olarak, seriler arasında aynı seviyede durağanlık şartı veya değişkenler arasında eşbütünleşme ilişkisi gerekliliği bulunmamaktadır. Bu, testin daha esnek ve geniş bir uygulama alanına sahip olmasını sağlar. İkinci olarak, bu nedensellik testinde değişkenlerin düzey değerleri kullanılır; böylece seriler üzerinde herhangi bir bilgi kaybı yaşanmaz. Toda ve Yamamoto yöntemi için (k+dmax) gecikmeli bir VAR modeli geliştirilir. Burada 'k', klasik VAR modelinde kullanılan gecikme sayısını ifade ederken, 'dmax' değişkenlerin maksimum bütünleşme derecesini gösterir. Bu model, nedensellik ilişkilerini daha doğru bir şekilde belirlemek için önemlidir.

Daha sonra, Toda ve Yamamoto tarafından geliştirilen ileri bir Wald (MWALD) test istatistiği, k gecikmeli VAR modeli parametrelerine uygulanarak nedensellik varlığını test eder. Wald testinin uygulanabilmesi için ise değişkenler arasındaki nedensellik modeli için Görünürde İlişkisiz Regresyon (GİR) modeli kullanılır. Bu yöntem, Zellner tarafından ilk olarak geliştirilmiş olup, regresyon sistemlerinde denklemleri ayrı ayrı tahmin etmek yerine Genelleştirilmiş En Küçük Kareler Yöntemi (GEKK) ile bir arada tahmin etmenin daha etkin sonuçlar verdiğini göstermiştir.

MWALD test istatistiği χ^2 dağılım sergiler. Test istatistiğinin anlamlı olması durumunda değişkenler

arasında bir nedensellik ilişkisinin varlığı sonucuna ulaşılmaktadır [15]. Kurulan VAR modeline ait denklemler eşitlik (1) ve eşitlik (2) de belirtilmiştir.

$$Y_{t-v_0} + \sum_{i=1}^{k+dmax} \alpha 1i y_{t-1} + \sum_{i=1}^{k+dmax} \beta 1i x_{t-1} + e_{1t} \quad (1)$$

$$X_{t-v_0} + \sum_{i=1}^{k+dmax} \alpha 2i y_{t-1} + \sum_{i=1}^{k+dmax} \beta 2i x_{t-1} + e_{2t} \quad (2)$$

Toda ve Yamamoto'nun (1995) çalışmasında öne sürülen nedensellik testi, Granger nedenselliğini Wald testi kullanarak değerlendirir. İki değişken arasındaki nedensellik ilişkisini test etmek için şu adımlar izlenir: Denklem 1 (x değişkeni ile y değişkeni): Sıfır hipotez, "x değişkeninin y değişkeninin Granger nedeni olmadığıdır." şeklinde formüle edilir. Yani, $H_0: \beta_1 = 0$, burada β_1 denklem 1'in x değişkenine ait gecikme terimi katsayısıdır. Denklem 2 (y değişkeni ile x değişkeni): Sıfır hipotez, "y değişkeninin x değişkeninin Granger nedeni olmadığıdır." şeklinde ifade edilir. Yani, $H_0: \beta_1 = 0$, burada β_1 denklem 2'nin y değişkenine ait gecikme terimi katsayısıdır.

Her iki durumdada, Wald testi kullanılarak ilgili gecikme terimi katsayısının sıfır olup olmadığı test edilir. Eğer gecikme terimi katsayısı istatistiksel olarak anlamlı bulunursa, Granger nedenselliği varlığı kabul edilir. Bu yöntem, değişkenler arasındaki nedensellik ilişkilerini belirlemek için kullanılan yaygın bir yöntem olup, Toda ve Yamamoto'nun önerisi doğrultusunda ileri bir istatistiksel analiz sunar.

III. SONUÇLAR VE TARTIŞMA

3.1. Bulgular

Serilerin durağanlığının test edildiği Genişletilmiş Dickey-Fuller testi (ADF) ve Phillips-Perron (PP) Birim Kök sınamaları Tablo 1'de verilmiştir. Buna göre ele alınan değişkenlerden sadece üretim miktarı değişkeni PP birim kök testinde düzeyde sabit ve düzeyde sabit+trend'de durağanlaşmıştır. Diğer iki değişken düzeyde birim kök içermekte ancak birinci farkları alındığında tüm değişkenler durağanlık göstermektedir.

Tablo 1. ADF ve PP birim kök sınaması sonuçları

	Değişkenler	Düzye		Birimci Farklar	Düzye		
		ADF	PP		ADF	PP	
Düzye	Sabit	Ekili Alan (ha)	-0.8154	-6.6405	Ekili Alan (da)	-6.5108***	-6.7569***
		Ürün Fiyatı (\$/ton)	-1.3542	-1.3323	Üretim Miktarı (ton)	-2.8317**	-5.5383***
		Üretim Miktarı (ton)	-1.1834	-5.8716***	Ürün Fiyatı (TL/kg)	-3.7523***	-57.8949***
	Sabit+Trend	Ekili Alan (ha)	-2.1099	-2.5019	Ekili Alan (da)	-6.4210***	-6.6652***
		Ürün Fiyatı (\$/ton)	-1.1708	-1.0869	Üretim Miktarı (ton)	-5.5055***	-5.5528***
		Üretim Miktarı (ton)	-3.5095*	-12.9306***	Ürün Fiyatı (TL/kg)	-3.9767***	-57.9563***

*, **, *** sırasıyla %10, %5 ve %1 düzeylerinde anlamlılığı ifade etmektedir.

Tablo 2'de VAR modeli için optimum gecikme uzunluklarına yer verilmiştir. Yalnızca Schwarz Bilgi

Kriterine göre bir gecikme uzunluğu varken, LR Test İstatistiği, Son Tahmin Hatası (FPE), Akaike Bilgi

Kriteri ve Hannan-Quinn Bilgi Kriterine göre iki gecikme uzunluğu belirlenmiştir. Bu sebeple yeniden

VAR modeli oluşturulmuş ve model oluşturulurken iki gecikme uzunluğu alınmıştır.

Tablo 2. VAR modelinde optimum gecikme uzunluğu sonuçları

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-12.80429	NA	0.0000576	1.053619	1.193739	1.098445
1	74.13823	150.7004	3.20e-06	-4.142549	-3.582070*	-3.963247
2	86.33922	18.70818*	2.64e-06*	-4.355948*	-3.375110	-4.042169*

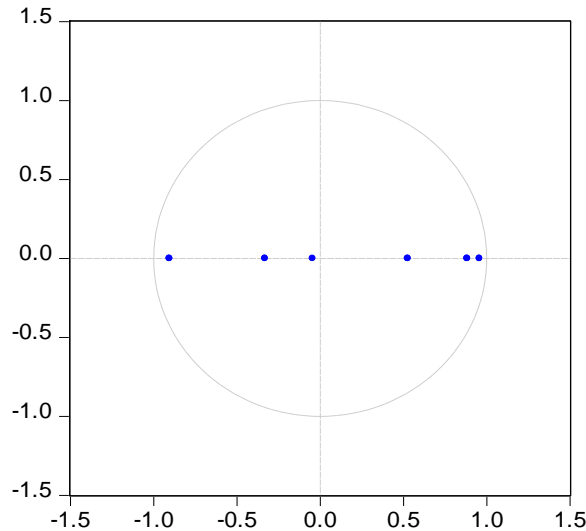
Oluşturulan VAR modelinin geçerliliğinin sınanması için yapılmış olan testler Tablo 3'te verilmiştir. Modelin geçerliliği için Lagrange Çarpanı (LM) Otokorelasyon Testi ve White Değişen Varyans Testi (With Cross) yapılmıştır. Toda-Yamamoto nedensellik testinde gecikme uzunluğunu $k+d_{max}$ olarak alacağımız için bu sınamalarda da gecikme uzunluğu 3 olarak alınmıştır. Buna göre LM Test İstatistiği

15.35607 ve olasılık değeri 0.0829 bulunmuştur. White Değişen Varyans testi ise 50.93871 ve olasılık değeri 0.0035 olarak bulunmuştur. Her iki sınamada olasılık değerleri 0.05'ten büyüktür ve her iki sınamada da %5 anlamlılık düzeyine göre ters hipotez geçerli olduğu için serilerde otokorelasyon ve değişen varyans sorununa rastlanmamıştır. Seriler analizler için uygundur.

Tablo 3. Otokorelasyon ve değişen varyans testi sonuçları

Lagrange Çarpanı (LM) Otokorelasyon Testi		
Gecikme Uzunluğu	LM-Test İstatistiği	Olasılık Değeri
3	15.35607	0.0829
White Değişen Varyans Testi (With Cross)		
Gecikme Uzunluğu	F-Test İstatistiği	Olasılık Değeri
3	50.93871	0.0035

Inverse Roots of AR Characteristic Polynomial



Şekil 1. AR karakteristik polinomialinin ters kökleri

Şekil 1'de AR karakteristik polinomlarının ters köklerine yer verilmiştir. Burada önemli olan nokta Modulus değerlerinin çember içerisinde kalması veya tüm değerlerin 1.00'dan düşük olması beklenir.

Şekil'de görüldüğü gibi tüm Modulus değerleri çember içerisindedir ve seriler analiz için uygundur.

Tablo 4. Görünürde ilişkisiz regresyon modeli sonuçları

Equation: $LUF = C(1)*LUF(-1) + C(2)*LUF(-2) + C(3)*LUF(-3) + C(4)*LUM(-1) + C(5)*LUM(-2) + C(6)*LUM(-3) + C(7)*LEA(-1) + C(8)*LEA(-2) + C(9)*LEA(-3) + C10$	
R-squared	0.887863
Adjusted R-squared	0.834745
S.E. of regression	0.177912
Equation: $LUM = C(11)*LUF(-1) + C(12)*LUF(-2) + C(13)*LUF(-3) + C(14)*LUM(-1) + C(15)*LUM(-2) + C(16)*LUM(-3) + C(17)*LEA(-1) + C(18)*LEA(-2) + C(19)*LEA(-3) + C(20)$	
R-squared	0.843006
Adjusted R-squared	0.768640
S.E. of regression	0.206624
Equation: $LEA = C(21)*LUF(-1) + C(22)*LUF(-2) + C(23)*LUF(-3) + C(24)*LUM(-1) + C(25)*LUM(-2) + C(26)*LUM(-3) + C(27)*LEA(-1) + C(28)*LEA(-2) + C(29)*LEA(-3) + C(30)$	
R-squared	0.979202
Adjusted R-squared	0.969351
S.E. of regression	0.028651

Tablo 4'te görünürde ilişkisiz regresyon modeline yer verilmiştir. Toda-Yamamoto Nedensellik testini uygulayabilmek amacıyla Wald testine ihtiyaç duyulmaktadır. Wald testinin uygulanabilmesi amacıyla görünürde ilişkisiz regresyon analizinin C katsayılarına ihtiyaç duyulmaktadır. LUF; üretim fiyatını, LUM; üretim miktarını ve LEA; ekili alanı vermektedir.

Tablo 5. Toda-Yamamoto nedensellik testi sonuçları

Model	Gecikme Uzunluğu	Wald Test	Olasılık Değeri
UM=f(UF)	$k = 2$	10.29723	0.0162
	$d_{max} = 1$		
EA=f(UF)	$k = 2$	10.69738	0.0135
	$d_{max} = 1$		

Tablo 5'da Toda-Yamamoto nedensellik testine yer verilmiştir. Ürün fiyatını ile üretim miktarı ve ekili alan arasındaki nedensellik ilişkileri $UM=f(UF)$ ve $EA=f(UF)$ şeklinde iki model oluşturulmuştur. Gecikme uzunluğu uygun wald testi için $k+d_{max}$ olduğu için 3 olarak belirlenmiştir. Buna göre ürün fiyatı ile üretim alanı arasındaki wald test 10.29723, olasılık değeri 0.0162 olarak tespit edilmiş, üretim fiyatı ile ekili alan arasındaki wald test 10.69738 ve olasılık değeri 0.0135 olarak bulunmuştur. Buna göre %5 anlamlılık düzeyine göre ürün fiyatı ile üretim miktarı ve ekili alan arasında bir nedensellik ilişkisinin varlığından söz edilebilir. Yani ürün fiyatı ve üretim miktarını hem de üretim alanını ele alınan dönem içerisinde etkilemiştir.

IV. SONUÇ

Bu çalışmada, Türkiye'de 1991-2022 yılları arasında zeytin ürün fiyatı, zeytin üretim miktarı ve zeytin üretim alanı arasındaki nedensellik ilişkisi test edilmiştir. Serilerin uygunluğu birim kök sınaması ile test edilmiş, gecikme uzunluğu belirlenmiş ve uygun VAR modeli tespit edilmiştir. Toda-Yamamoto nedensellik testinin belirlenebilmesi amacıyla gerekli sınamalar yapılmış ve nedensellik ilişkisi ortaya konmuştur. Ürün fiyatındaki yıllar itibarıyla oluşan değişiklik hem ürün miktarını hem de üretim alanını %5 anlamlılık düzeyinde etkilemektedir. Buna göre Toda Yamamoto nedensellik testine göre ürün fiyatı ile üretim miktarı ve ekili alan arasında bir nedensellik ilişkisinin varlığından söz edilebilir.

Türkiye zeytin üretimi açısından Akdeniz Havzasında önemli bir ülke konumundadır. Türkiye dünyada zeytin üretim alanında beşinci sırada yer almasına rağmen üretiminde İspanya'dan sonra ikinci sırada gelmektedir. Dünyada ihracatçı bir konumda bulunan Türkiye açısından zeytin üzerine yapılan politikalar önem arz etmektedir. Marmara bölgesinde birçok zeytin ağacı finansal sürdürülebilirlik sağlayamadığı gerekçesiyle sökülüştür. Zeytin ve zeytin yağındaki iç ve dış pazar potansiyeli dikkate alındığında zeytin yetiştiriciliğinin finansal açıdan sürdürülebilirliğinin artırılması Türkiye ekonomisi açısından önemli katkılar sağlayabilir.

Son yıllarda Türkiye'nin Uluslararası Zeytin Konseyi'nde aktif olması, zeytin ve zeytinyağının tanıtımında Ulusal Zeytin ve Zeytinyağı Konseyi (UZZK) ile İzmir Ticaret Borsası'nın iş birliği yapması dış ticaret açısından önemlidir.

Sonuç olarak yıllar itibarıyla üretim alanı ve üretim miktarı artış göstermiş ancak ürün fiyatı yıllar itibarıyla değişkenlik göstermektedir. Fiyat dengesizlikleri üreticilerin zeytin üretiminde kararsız kalması, üretime devam etmeme isteği olarak geri dönüş sağlamaktadır. Dünyada en önemli ihracatçılardan birisi konumunda olan Türkiye'nin bu konuda yeni tarımsal politikalar geliştirerek konumu koruma ve birinci sıraya yükselme şansı bulunmaktadır.

Çalışmanın gerek tarım sektöründe gerekse bu ürün özelinde birbirini etkileyen değişkenlerin belirlenip nedenselliğinin belirlenmiş olması önemlidir. Gelecek çalışmalarda bu nedensellik ilişkisi yanında geleceğe yönelik projeksiyonlarla sonraki yıllar için tahminleme yapılması önem arz etmektedir. Bu sayede zeytin üretimi ve fiyatları üzerine gelecek yıllara yönelik bir fikir oluşması açısından önemlidir.

TEŞEKKÜR

Bu çalışma 21-23 Mayıs 2024 tarihleri arasında Marmara Üniversitesi'nde düzenlenen Uluslararası Uygulamalı İstatistik Kongresi (UYİK-2024)'de sözlü bildiri olarak kabul edilmiş ve sözlü bildiri olarak sunulmuştur.

KAYNAKLAR

- [1] Liphshitz, N., Gophna, R., Hartman, M. ve Biger, G. (1991). The Beginning of Olive (*Olea europaea*) Cultivation in the Old World: A reassessment. *Journal of Archaeological Science*, 18, 441-453.
- [2] Kestelenos, G., Kiritsakis, A. (2017). Olive Tree History and Evolution. Kiritsakis A ve Shahidi F. *Olives and Olive Oil as Functional Foods: Bioactivity, Chemistry and Processing*, 1-12. John Wiley & Sons.
- [3] Vossen, P. (2007). Olive Oil: History, Production and Characteristics of the World's Classic Oils, *Hort Sciences*, 42(5): 1093-1100.
- [4] Fortune Business Insights, (2022). Olive Oil Market, Global Market Analysis, *Insights and Forecast*.
- [5] FAOSTAT, 2024. Erişim Adresi: <https://www.fao.org/faostat/en/#data>
- [6] Şengül, Z. (2023). Türkiye'de Sofralık Zeytin Fiyatlarındaki Dalgalanmalar: ARIMA-GARCH Yaklaşımıyla Volatilité Araştırması. *ANADOLU Ege Tarımsal Araştırma Enstitüsü Dergisi*, 33(2), 281-295.
- [7] Kılıç, T.M., Turhan, Ş. (2023). Türkiye'nin Sofralık Zeytin İhracatı ve Değeri Arasındaki İlişki: ARDL Sınır Testi Yaklaşımı. *Türk Tarım ve Doğa Bilimleri Dergisi*, 10(2), 206-213.
- [8] Bulut, Ö.U., Aykırı M., Balcı, Ö. (2022). Elektrik Tüketiminin Ekonomik Büyüme Üzerindeki Asimetrik Etkisi: NARDL ve Toda-Yamamoto Yöntemlerinden Kanıtlar. *Uluslararası Ekonomi ve Yenilik Dergisi*, 8(2), 99-119.
- [9] Özçelik, O. (2023). Rusya-Ukrayna Savaşı Gölgesinde Dünyadaki Gıda Fiyatlarının Belirleyicileri: Fourier Bootstrap ARDL ve Fourier Bootstrap Toda-Yamamoto Yaklaşımlarından Kanıtlar. *Tarım Ekonomisi Dergisi*, 29(1), 29-47.
- [10] Okur, M.G., Çiçek, A. (2023). Türkiye'de Kırmızı Et Fiyatları ile Besi Yemi Fiyatları Arasındaki Nedensellik İlişkisinin Toda-Yamamoto Testi ile Belirlenmesi. *ÇOMÜ Ziraat Fakültesi Dergisi*, 11(2), 368-376.
- [11] Dickey, D.A., Fuller, W.A. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica*, 49(4): 1057-1072.
- [12] Perron, P. (1997). Further Evidence on Breaking Trend Functions in Macroeconomic Variables. *Journal of Econometrics*, 80 (2): 355-385.
- [13] Toda, H.Y., Yamamoto, T. (1995). Statistical Inferences in Vector Autoregression with Possible Integrates Processes. *Journal of Econometrics*, 66(1): 225-250.
- [14] Çetin, M., Saygın, S., Demir, H. (2020). Tarım Sektörünün Çevre Kirliliği Üzerine Etkisi: Türkiye Ekonomisi için Bir Eşbütünleşme ve Nedensellik Analizi. *Tekirdağ Ziraat Fakültesi Dergisi*, 17(3), 329-345.

Analysis of Artificial Intelligence Methods in Classifying Heart Attack Risk: Black-Box Models vs. Glass-Box Models

Kalp Krizi Riskinin Sınıflandırılmasında Yapay Zekâ Yöntemlerinin Analizi: Kara Kutu Modelleri ve Cam Kutu Modelleri

Ebru GEÇİCİ¹, Eyüp Ensar IŞIK¹, Mısra ŞİMSİR¹, Mehmet GÜNEŞ¹

¹ Yildiz Technical University, Faculty of Mechanical Engineering, Industrial Engineering Department, 34349, İstanbul, Türkiye

Abstract

Artificial Intelligence (AI) is becoming more and more involved in human life day by day. Healthcare is one of the areas where AI is widely used, such as in the diagnosis prediction, and/or classification of diseases. Techniques such as machine learning provide high-accuracy results, but many algorithms have black-box structures, where the reasoning behind the predictions is not known. Explainable AI emerges to address this by providing explanations for complex models. While interpretable ("glass-box") models are desirable, they may have lower accuracy than complex ("black-box") models. Finding the right balance is crucial, especially in critical areas such as healthcare. It is also important to provide individual explanations for the predictions. This study uses patient data to explore a model to predict heart attack risk. Therefore, we compare glass-box models (logistic regression, naive Bayes, decision tree, and explainable boosting) with black-box models (random forest, support vector machine, multi-layer perceptron, gradient boosting, and stochastic gradient boosting). The results show that explainable boosting achieves the highest accuracy. To delve into individual explanations on a patient basis, the explainable boosting algorithm is compared with the random forest algorithm, which gives the best results among the black-box models. Here, LIME and SHAP are used to provide interpretability of random forests. As a result, it is concluded that the random forest algorithm has differences in the importance weights of the variables compared to the explainable boosting algorithm. Both results provide valuable tools for healthcare stakeholders to choose the most appropriate model.

Keywords: Artificial Learning, Explainable Artificial Intelligence, Classification, Healthcare Industry, Heart Attack

I. INTRODUCTION

Artificial Intelligence (AI) and AI-focused applications have increased frequently in recent years. This growth continues to be seen in AI's supporting applications that assist daily life, as well as its supporting role at critical decision points. At these points, AI's only result-oriented support to the decision maker can lead to problems explaining the reasons for the decision. With current studies in AI, complex models with improved prediction accuracy are being used, making explanations even more difficult. Such problems have led to the emergence of terms such as *understandability*, *comprehensibility*, *interpretability*, *explainability*, and *transparency* in AI [1]. The expected starting point of these terms is the need to explain how machine learning (ML) models make their outputs or decisions, which are becoming increasingly complex and cannot be explained by themselves. With the increase in this need, the study of eXplainable Artificial Intelligence (XAI) has increased significantly in recent years [2].

XAI emphasizes the explainability of what features may result from decisions while preserving the predictive accuracy of the methods used. Therefore, decision-makers need to choose a model that ensures prediction accuracy in the field under study but is also high in interpretability of the decision made [3]. In this context, ML models can be examined under two main headings: *glass-box* and *black-box* models. While glass-box (or white-box) models include self-interpretable methods, black-box models represent methods that cannot be interpreted automatically due to the high number of layers and parameters but can be provided with local explanations using different XAI techniques [4]. Although the prediction accuracy of black-box models is relatively high through their complex structure, they require additional techniques to explain the decisions made. Rudin [5] discusses using glass-box models instead of black-box models in the first stage to avoid spending too much effort explaining the black-box models.

The increasing use of AI applications in areas that directly affect social and human life increases the importance of using the suitable model in the right area. Although the accuracy of the decision is of utmost importance, especially in health-related studies, healthcare professionals who are decision-makers need to explain the circumstances under which this decision was made [6-7]. As a result of models used in areas such as disease detection, when deciding whether an individual has a disease, an explanation needs to be given to the individual, such as what factors caused the disease. For this reason, XAI studies in the healthcare sector have increased in recent years [8].

One of the areas where AI applications are frequently used in the healthcare sector is in detecting patients' heart attack risk [9]. Many studies have been made in this field, and models that increase prediction accuracy to the highest level have been proposed. However, the number of studies examining the explainability of the proposed models is quite limited. In this study, we compare the glass-box methods and black-box methods. In predicting heart attack risk, the differences between using a complex model that is difficult to explain due to its structure and a model that may not have deficient performance but is easily explainable are being discussed.

The rest of the paper is organized as follows: The literature review section presents related studies in this field. Then, the methodology is given in section 3. Section 4 addresses the application and its results. Last, we conclude in Section 5.

II. LITERATURE REVIEW

For many years, ML algorithms have been used in a wide range of applications in various fields, such as recommendation systems, cybersecurity, image processing, industrial applications, education, and

healthcare. The literature section of this study provides an overview of studies on ML applications in healthcare. ML applications in the healthcare sector can be found in many areas, from disease diagnosis to personalized treatment, drug discovery to radiology, etc. The article reviews the applications of ML and AI algorithms in healthcare [10]. It states that support vector machines (SVM), decision trees (DT), random forests (RF), and artificial neural networks (ANNs) are widely used algorithms in this field [11-15].

Heart attack/stroke is one of the most critical and focused problems in healthcare, and it causes many deaths all over the world. The latest advances in the application of ML have shown that it is possible to detect heart disease at an early stage using electrocardiograms and patient data [16]. By analyzing large amounts of patient data, ML algorithms can more accurately and quickly identify risk factors for heart attack. In contrast to traditional methods, ML algorithms can use a patient's medical history, genetic information, and lifestyle to build more complex and predictive models. This enables physicians to monitor patients more effectively, identify high-risk individuals in advance, and take the necessary preventive measures.

Sahu et al. [17] compare conventional ML algorithms (SVM, Naïve Bayes (NB), DT, RF, Logistic Regression (LR), k-nearest neighborhood (KNN)) and deep learning algorithms for using two different data sets (taken from the UC Irvine (UCI) and Kaggle repositories) to predict the heart attack and death rates related to heart attack [17]. They conclude that the one-dimensional convolutional neural network (1D-CNN) algorithm predicts heart attack and death rates with 99% accuracy, outperforming conventional methods. Rao et al. [18] attempt to predict whether a patient will have a heart disease. LR and ANN models are used to classify them. They compare the accuracy rates of the two models, and LR outperforms ANN by 90%. Mahmud et al. [19] try to predict heart failure in patients based on clinical data. They use one of the well-known data sets (taken from Kaggle repositories), which combines five different cardiac datasets, making it the most comprehensive resource available for heart disease research. RF, NB, KNN, and DT methods are used to build a combined meta-model—the results of the evaluation show that the meta-model outperforms other state-of-the-art models. The accuracy of the meta-model is 87%. Mamun and Elfouly [20] introduce a hybrid 1D-CNN model utilizing features selected by feature selection algorithms as well as a substantial data set derived from online survey data. The 1D-CNN has shown superior accuracy compared to contemporary ML algorithms and ANNs. The model's performance is compared with ANN, RF, AdaBoost, and SVM, and 1D-CNN outperforms these methods in terms of accuracy, false negative rates, and false positive rates. Ozcan and Peker [21] introduce a classification and

regression tree chart, a supervised ML method to predict whether a patient will have a heart disease. They try to explain the relationship between the input variables and the response, so they rank the features that affect heart disease by importance. The accuracy of the proposed algorithm, which is 87%, shows the reliability of the model. Yu [22] uses ML algorithms to predict the likelihood of occurrence of heart diseases in patients. The data set taken from the UCI repository is used to analyze eight different ML classifiers. LR, SVM, KNN, NB, DT, RF, gradient boosting, and AdaBoost algorithms are compared. As a result, the gradient boosting classifier achieves the highest accuracy with 95.08%.

XAI studies have increased in recent years due to the need to learn which inputs result from the decisions obtained from increasingly complex ML models [2,15]. Although the studies in this field use different terms representing similar needs, each term does not have the same meaning. Although there is no clear definition for XAI due to different terminologies, to provide a consensus, Barredo Arrieta et al. [1] define XAI as "a set of practices that produce details or reasons to make its functioning clear or understandable, given an audience." The authors state that this definition indirectly includes causality, transferability, informativeness, fairness, and reliability, which are seen as missing in other definitions but are covered by XAI. As can be understood from this definition and the topics it should include, XAI provides explanations for AI applications that serve many purposes and enable us to understand AI models better.

AI models can be divided into two categories in terms of explainability. The former is models that can be explained independently without using additional techniques. These models can be found in the literature under "transparent models, glass-box models, intrinsic explainability, ante-hoc approaches, and inherently interpretable ML models" [4]. This study uses the term "glass-box models" for self-explanatory models. Although their explainability is at different levels, linear regression/LR, DT, KNN, rule-based methods, general additive models, and Bayesian models are considered glass-box models [1]. These models can be explained after the prediction without any post-hoc analysis. On the other hand, models that cannot be explained by themselves due to their complex structures and where the obtained predictions can only be explained by post-hoc analyses are called "black-box models" in the literature. RF, SVM, multi-layer perceptron (MLP), and ANNs are examples of black-box models. It is crucial to be able to explain these models that provide high prediction accuracy. Therefore, new techniques have been developed to explain black-box models, and they can be divided into two categories: (i) model-agnostic and (ii) model-specific. The most well-known of these techniques are SHapley Additive exPlanations (SHAP) [24] and Local

Interpretable Model-Agnostic Explanations (LIME) [25], which are classified as model-agnostic.

Literature reviews conducted in recent years clearly reveal how popular XAI is. In addition to providing information about the current terminology in the field, these studies also present newly developed methods and application areas in detail [1-3], [23], [26]. Although it has applications in many fields, such as finance, education, environmental science, and agriculture, XAI stands out, especially with its uses in healthcare.

III. METHODOLOGY

This part of the study presents information on the glass-box and black-box models used in the analysis. LIME and SHAP, which allow local interpretation of black-box models, are examined. Lastly, performance metrics that allow the evaluation of the generated models are discussed.

3.1. Artificial Learning Algorithms

The ML and ANN models used in the analyses are presented in this section. In this context, the glass-box and black-box methods are first considered, and then the methods that enable the black-box models to be explained locally are mentioned. This section also provides information about the performance metrics used to compare the models created.

ML techniques are divided into two categories: supervised learning and unsupervised learning [27]. This distinction is related to the presence or absence of the output value in the data set: (i) if the dependent variable (response), y , is present in the data set, it is called supervised learning, and (ii) if there is no dependent variable in the data set, it is called unsupervised learning. Furthermore, supervised learning is divided into prediction and classification according to the structure of the dependent variable in the data set. While regression is used to predict the dependent variable, which has a continuous structure, classification involves classifying data using the output, which has a discrete structure. Unsupervised learning algorithms, on the other hand, are preferred for purposes such as making inferences about the data or organizing the data set (such as dimensionality reduction), and the well-known applied methods in this field are clustering algorithms. Clustering, which falls under unsupervised learning, involves grouping processes by bringing together independent variables with similar characteristics [27]. In this study, the aim is to evaluate whether patients are at risk of having a heart attack. The response value has a binary structure (0: lower heart attack risk and 1: higher heart attack risk), i.e., the response is a discrete variable. In this context, the glass-box and black-box classification methods are examined below.

LR measures the statistical significance of each independent variable relative to probability. It is highly probabilistic and a powerful ML method that models binomial output [28]. The *NB* method is one of the methods based on Bayes theorem and is called the probabilistic classification method [29]. The naive assumption is an assumption of conditional independence between each pair of features given the value of the class variable. Moreover, it is a preferred algorithm because it is easier to use and understandable and gives faster results than other complex methods [30]. In the *DT classification*, each node represents a feature, each branch represents a rule, and each leaf represents a result [31]. DTs have a hierarchical structure developed by dividing the data set into smaller structures. The *Explainable boosting machine (EBM) classifier* is a cyclic gradient boosting generalized additive model, and similar to the DTs, it is also a tree-based method. Furthermore, these models have as high prediction accuracy as black-box models, but their interpretation is inherently easier than black-box models.

Many DTs work together to create an RF algorithm, and then the average of all these trees is used [28]. This structure allows more consistent results to be obtained compared to DT algorithm results [27]. *SVM classification* creates hyper-planes to separate data into multiple classes [32]. Unlike other classification algorithms, it tries to maximize the distance between the created clusters. It considers the separation of points by a line or plane and the resulting distance. *MLP* is one of the well-known ANN algorithms that can indirectly detect complex nonlinear relationships between dependent and independent variables [33]. This method is inspired by the working structure of the human brain; inputs pass through layers respectively, and output is created. Each layer consists of neurons, and the values obtained from here are obtained by passing through the activation function. *Gradient boosting* is one of the ML ensemble methods that create more than one model and then combine them to produce improved results. These well-known ML models reunite several weak learners into strong learners, in which each new model is trained to minimize the loss function appropriate to the structure of the problem. At each step, the algorithm calculates the direction of improvement for the ensemble's predictions and then trains a new weak model to move in that direction. The new model's predictions are then added to the ensemble, and this process is repeated until a stopping condition is met. As in the gradient descent method, the stochastic gradient descent model tries to minimize the loss function value defined iteratively. The reason for using the concept of stochastic in its name is based on using structures such as applying mini-batches of differentiation in iterations and creating random subsets. In this way, while trying to reach the highest efficiency value in calculations, the randomness value also increases.

The local interpretation of the black-box models given above is not straightforward due to the complex structure of the models. Therefore, intermediary stages are needed to evaluate these models with respect to observations. In this context, *LIME* and *SHAP* are used in this study. These methods are visualization techniques applied to ML algorithms and are recommended to explain the model by bringing the model predictions closer to an interpretable model. To explain individual predictions, LIME creates new data points that resemble the instance of interest. These points are generated based on a statistical model learned from the features of the dataset, treating them as independent variables. Note that it is considered that the features are independent of the other and follow a normal distribution, whose parameters are inferred from the data set [34]. SHAP combines game theory concepts with local explanation techniques [24]. SHAP transforms the original input data into a more straightforward form using a specific function. In this model, using a reduced data set, the original model can be approximated with a linear function of binary variables [35].

3.2. Evaluation of the Models: Performance Metrics

The explanation of the algorithms used in the analyses provides information about the output of whether a heart attack has occurred. The output value in the data set is considered as a 0 – 1 binary structure, and classification algorithms are used to select the model. Similarly, the definition of the performance metrics used to compare algorithms should also be appropriately chosen for the output structure. In this context, the performance metrics are accuracy, precision, recall, and f1-score. Moreover, the receiver operating characteristic (ROC) curve and the area under the ROC curve value (AUC) are also reported.

A confusion matrix must be created to use the above performance metrics in classification problems. The values in the confusion matrix are used to visualize and summarize the results. In this context, the confusion matrix is created for problems with binary output values, as shown in Table 1.

Table 1. Confusion matrix for a problem which has a binary structure

		Predicted Values	
		Positive (1)	Negative (0)
Actual Values	Positive (1)	True Positive (TP)	False Negative (FN)
	Negative (0)	False Positive (FP)	True Negative (TN)

Accuracy, one of the performance metrics, shows the percentage of samples classified correctly and is mathematically expressed as $\frac{TP+TN}{TP+TN+FP+FN}$. *Precision*, another performance metrics, is used to calculate how many of the values predicted as positive are actually true positives. Its mathematical expression is presented

as $\frac{TP}{TP+FP}$. The *recall* metric, like precision, deals with positive values, but unlike the previous one, recall is a performance metric that shows how much of the operations that should be predicted as positive are predicted as positive (Mathematically, $\frac{TP}{TP+FN}$). The *f1-score* value is calculated as the harmonic average of recall and precision values, i.e., $2 \times \frac{Precision \times Recall}{Precision + Recall}$.

ROC curve is a metric that allows visual evaluation, unlike the performance metrics given above. True positive rate and false positive rate are used to obtain this curve. While the false positive rate value on the x-axis is calculated using $\frac{FP}{FP+TN}$, the true positive rate corresponds to the recall value explained above. *AUC* is obtained by calculating the area under the ROC curve. Their expressions are shown in Figure 1. In Figure 1, the dark blue line corresponds to the ROC curve, while the gray area corresponds to the AUC value.

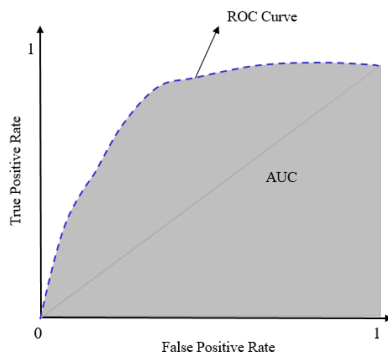


Figure 1. Representation of the ROC curve and AUC value in a graph

IV. APPLICATION AND RESULTS: PREDICTION OF HEART ATTACK RISK

In this part of the study, the methods specified are analyzed using the selected data set, and the results are presented. First, information about the data set is given, and then the application details are addressed.

4.1. Data Set

The data set "heart attack" to be used in the analysis has been shared with researchers and users as open access [36]. The data set contains information about whether the individuals whose information is included have had a heart attack. Moreover, it consists of a total of 303 observation values, 13 features, and one output value. Definitions of the features are included in Table 2.

4.2. Application and Results

This section provides information on the application and the results obtained. AI algorithms classified as glass-box and black-box in the Methodology section are used to select the model to be created to determine the risk of a heart attack. The created models are run using a computer with an Intel(R) Core(TM) i5-10210U CPU @ 1.60GHz 2.11 GHz processor and 8 GB RAM. Python programming language and existing libraries are preferred when creating the models. Accordingly, the libraries used for artificial learning are LR (*LogisticRegression*), NB (*GaussianNB*), DT classifier (*DecisionTreeClassifier*), SVM classifier (*SVC*), MLP (*MLPClassifier*), gradient boosting (*GradientBoostingClassifier*), stochastic gradient boosting (*SGDClassifier*) and explainable gradient boosting (*ExplainableBoostingClassifier*), whereas the libraries used to explain the black box models are LIME (*LimeTabular*) and SHAP (*shap*).

Table 1. Explanation of the features

Feature	Explanation	Structure
Age (Years)	The age of individuals	Integer
Sex	The gender of individuals <i>Categories:</i> 0: female, 1: male	Categorical
cp	The chest pain type <i>Categories:</i> 1: typical angina, 2: atypical angina, 3: non-anginal pain, 4: asymptomatic	Categorical
trestbps (mm Hg)	The resting blood pressure (on admission to the hospital)	Integer
chol (mg/dl)	The serum cholesterol level fetched via BMI sensor	Integer
fbs	The fasting blood sugar > 120 mg/dl <i>Categories:</i> 1: true, 0: false	Categorical
restecg	The resting electrocardiographic results <i>Categories:</i> 0: normal, 1: having ST-T wave abnormality, 2: definite left ventricular hypertrophy by Estes' criteria	Categorical
thalach	The maximum heart rate achieved	Integer
exang	The exercise induced angina <i>Categories:</i> 1: yes, 0: no	Categorical
oldpeak	The ST depression induced by exercise relative to rest, previous peak	Integer
slope	The slope of the peak exercise ST segment <i>Categories:</i> 1: unsloping, 2: flat, 3: downsloping	Categorical
caa	The number of major vessels (0-3) colored by flourosopy	Integer
thal	<i>Categories:</i> 3: normal; 6: fixed defect; 7: reversable defect	Categorical

According to the specified information, the data set is first analyzed, and then models are established. After the model-building phase is completed, the best model must be selected. Then, the results must be explained so that decision-makers can understand; for example, in this problem, the decision-makers are physicians. However, as mentioned, the artificial learning algorithms used in this study are presented as two pillars: (i) glass-box and (ii) black-box. Due to their structure, the models referred to as glass-box can be understood and interpreted by experts in the field who

do not know ML. On the other hand, suppose a model under the black-box heading is chosen. In that case, it becomes difficult for field experts to interpret the results and interpret the data from an individual perspective. For this reason, explanatory methods are used to help explain black-box models. The models obtained afterward are the best among the established models so that the outputs can be easily interpreted. This process to be followed during the implementation phase is visualized in Figure 2.

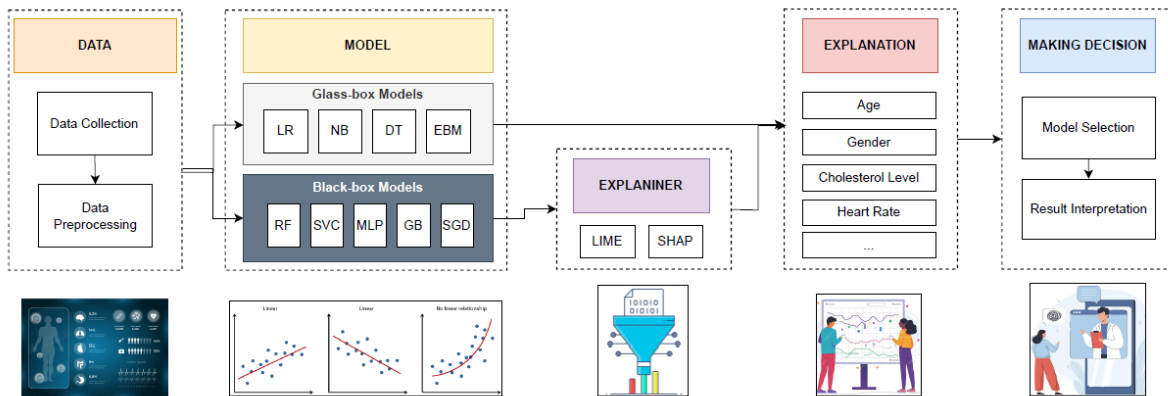


Figure 2. Flow diagram to be followed for obtaining and interpreting the results

A fine-tuned process is carried out to prepare the models for use and select the best among them. Through this process, the hyper-parameters, which are parameter values that are given externally to the model, of the models are determined. For this process, more than one value is tried for the relevant parameter in each model. The fine-tuned parameters for each model are summarized in Table 3. The established artificial learning model is included in the "Models" column in the table. Under the heading "Hyper-Parameters," hyper-parameters are differentiated in creating the established models. "Best Model Parameters" illustrates the hyper-parameters used to prepare the best model obtained due to the different hyper-parameter values run for the relevant model. Moreover, while giving the information in this column, hyper-parameter definitions are presented precisely the same as the name

in the library included in the package used. Thus, it is aimed to create a clear table for users who want to use the same models. Besides, the grid search algorithm is used to fine-tune the models. After completing this process, the parameters that yield the best results are reported for each model created using the specified method. Thus, models to be compared are obtained to select the best model. The last column, "Time (sec)," presents the time to determine the best parameters among the existing ones during the fine-tuned process. No duration has been defined since existing models are used for LR, NB, and explainable boosting methods. In other models, it is seen that the method that requires the most time in the parameter definition phase is the gradient boosting method, and the method that requires the least time is the MLP method.

Table 3. Artificial learning models with hyper-parameters tuned by grid search in the inner loop

Models	Hyper-Parameters	Best Model Parameters	Time (sec)
LR	-		
NB	-		
DT Classifier	<ul style="list-style-type: none"> • <i>Criterion</i>: Measure the split quality and is a measure of impurity. • <i>Splitter</i>: Define the strategy which is used to choose split at each node. • <i>Maximum Depth</i>: Give the tree's maximum depth. • <i>Maximum Features</i>: The number of features to consider when looking for the best split. • <i>Complexity Parameter</i>: Used for minimal cost-complexity pruning. 	<i>criterion: gini</i> <i>splitter: best</i> <i>max_depth: 1</i> <i>max_features: None</i> <i>ccp_alpha: 0</i>	20.0325
RF Classifier	<ul style="list-style-type: none"> • <i>Criterion</i>: Measure the split quality and is a measure of impurity. • <i>Number of Estimator</i>: The number of trees in the forest. • <i>Maximum Depth</i>: Give the tree's maximum depth. • <i>Weights associated with classes</i> 	<i>criterion: gini</i> <i>n_estimators: 200</i> <i>max_depth: 4</i> <i>class_weight: balanced</i> <i>subsample</i>	605.7901

Table 3. Artificial learning models with hyper-parameters tuned by grid search in the inner loop (cont.)

Models	Hyper-Parameters	Best Model Parameters	Time (sec)
SVM Classifier	<ul style="list-style-type: none"> <i>Kernel</i>: The kernel type to be used in the algorithm <i>Gamma</i>: Kernel coefficient: <i>Regularization parameter</i> 	<i>kernel: rbf</i> <i>gamma: auto</i> <i>C: 3.0</i>	1.6822
MLP Classifier	<ul style="list-style-type: none"> <i>Activation</i>: The hidden layer activation function <i>Solver</i>: The solver for weight optimization <i>Alpha</i>: Strength of the L2 regularization term <i>Learning rate</i> <i>Maximum number of iterations</i> 	<i>activation: logistic</i> <i>solver: adam</i> <i>alpha: 0.0</i> <i>learning_rate: constant</i> <i>max_iter: 300</i>	2703.402
Gradient Boosting Classifier	<ul style="list-style-type: none"> <i>Loss</i>: Type of the loss function <i>Learning Rate</i>: Learning rate shrinks the contribution of each tree <i>Criterion</i>: Measure the split quality <i>Maximum Depth</i>: Maximum depth of the individual regression estimators <i>Max Features</i>: The number of features to consider when looking for the best split. 	<i>loss: log_loss</i> <i>learning_rate: 0.9</i> <i>criterion: squared_error</i> <i>max_depth: 8</i> <i>max_features: sqrt</i>	820.8958
Stochastic Gradient Descent Classifier	<ul style="list-style-type: none"> <i>Loss</i>: Loss function type <i>Penalty</i>: Regularization term <i>Alpha</i>: Coefficient of the regularization term <i>Learning Rate</i> 	<i>loss: log_loss</i> <i>penalty: l2</i> <i>alpha: 0.003</i> <i>learning_rate: optimal</i>	23.3565
Explainable Boosting Classifier	-		

After the fine-tuning process with grid search, the best models are compared. A run is taken using the hyper-parameters determined to compare different models. As a result of these runs, the models' performance metrics calculated using the training data are shown in Table 4. Different performance metrics are listed in this resulting table. The headings in the table provide information regarding the classification of the models operated in the first column as glass-box or black-box. The model number and the name of the model are shared in the following two columns. Furthermore, there are four subheadings under the heading "Performance Metrics": "Accuracy," "Recall," "Precision," and "F1-Score". The results for each model are presented for these performance metrics specified under subheadings.

When the results in Table 4 are examined, it is seen that the best results among the models called glass-box are obtained with the explainable boosting method. Black-box models are examined; on the other hand, it is observed that the best results are obtained 100% with the gradient-boosting classifier algorithm. Moreover, these results should be evaluated in terms of over-fitting. Thus, the performance metrics obtained for the

test data set are considered to determine whether the models are usable. In this context, the performance metrics of the test data are listed in Table 5. Table 5 is created to resemble Table 4, where the values obtained with the training data set are reported. In addition to existing performance metrics, the AUC value is also reported. The graph of the AUC values given with the performance metrics of the test values is shown in Figure 3.

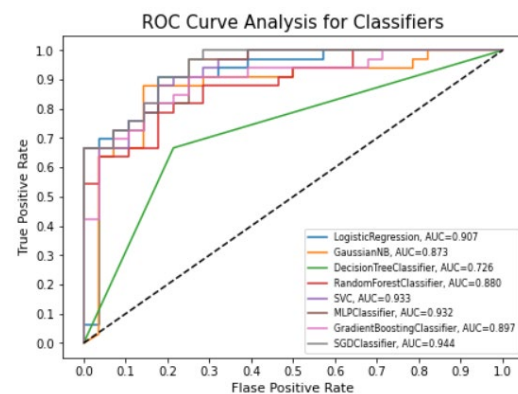


Figure 3. ROC curve graphs obtained in models for test data

Table 4. Performance metrics for train data set

	#	Model	Performance Metrics (%)			
			Accuracy	Recall	Precision	F1-Score
Glass Box Models	1	Logistic Regression	88.8430	88.8430	88.9863	88.9146
	2	Naïve Bayes	67.7686	67.7686	77.0972	72.1325
	3	Decision Tree	77.6860	77.6860	77.6528	77.6694
	4	Explainable Boosting	95.0413	95.0413	95.0489	95.0451
Black Box Models	1	Random Forest	88.0165	88.0165	88.0128	88.0147
	2	Support Vector	85.9504	85.9504	86.0250	85.9877
	3	Multi-Layer Perceptron	86.7769	86.7769	86.8602	86.8185
	4	Gradient Boosting	100.0000	100.0000	100.0000	100.0000
	5	Stochastic Gradient Descent	86.3636	86.3636	86.5579	86.4607

Table 5. Performance metrics for test data set

	#	Model	Performance Metrics (%)				
			Accuracy	Recall	Precision	F1-Score	AUC
Glass Box Models	1	Logistic Regression	81.9672	81.9672	83.0761	82.5179	0.907
	2	Naïve Bayes	75.4098	75.4098	81.6214	78.3927	0.873
	3	Decision Tree	72.1311	72.1311	73.1069	72.6158	0.726
	4	Explainable Boosting	77.0492	77.0492	76.8275	76.9382	0.872
Black Box Models	1	Random Forest	75.4098	75.4098	77.2509	76.3193	0.880
	2	Support Vector	80.3279	80.3279	81.8386	81.0762	0.933
	3	Multi-Layer Perceptron	78.6885	78.6885	79.7530	79.2172	0.930
	4	Gradient Boosting	80.3279	80.3279	82.9217	81.6042	0.897
	5	Stochastic Gradient Descent	81.9672	81.9672	82.4220	82.1940	0.944

Suppose the performance metrics obtained for the test and train data sets are examined simultaneously. In that case, it is seen that in most of the models, better results are obtained with the train data set, and there are decreases in these metrics for the test data. Moreover, this decrease is greater in black-box models than in glass-box models. In addition, the test performance metrics for the gradient boosting algorithm, which achieved 100% success, are lower than other black-box models. This shows that over-fitting is explicitly

observed for this model. For this reason, if black-box models are established and a model is selected, it would be appropriate to choose the RF method, where the change between train and test is less and good results are obtained in the training data. When the glass-box models are examined, it becomes clear that the explainable boosting algorithm is a usable model in terms of giving good results in training and good results in terms of testing.

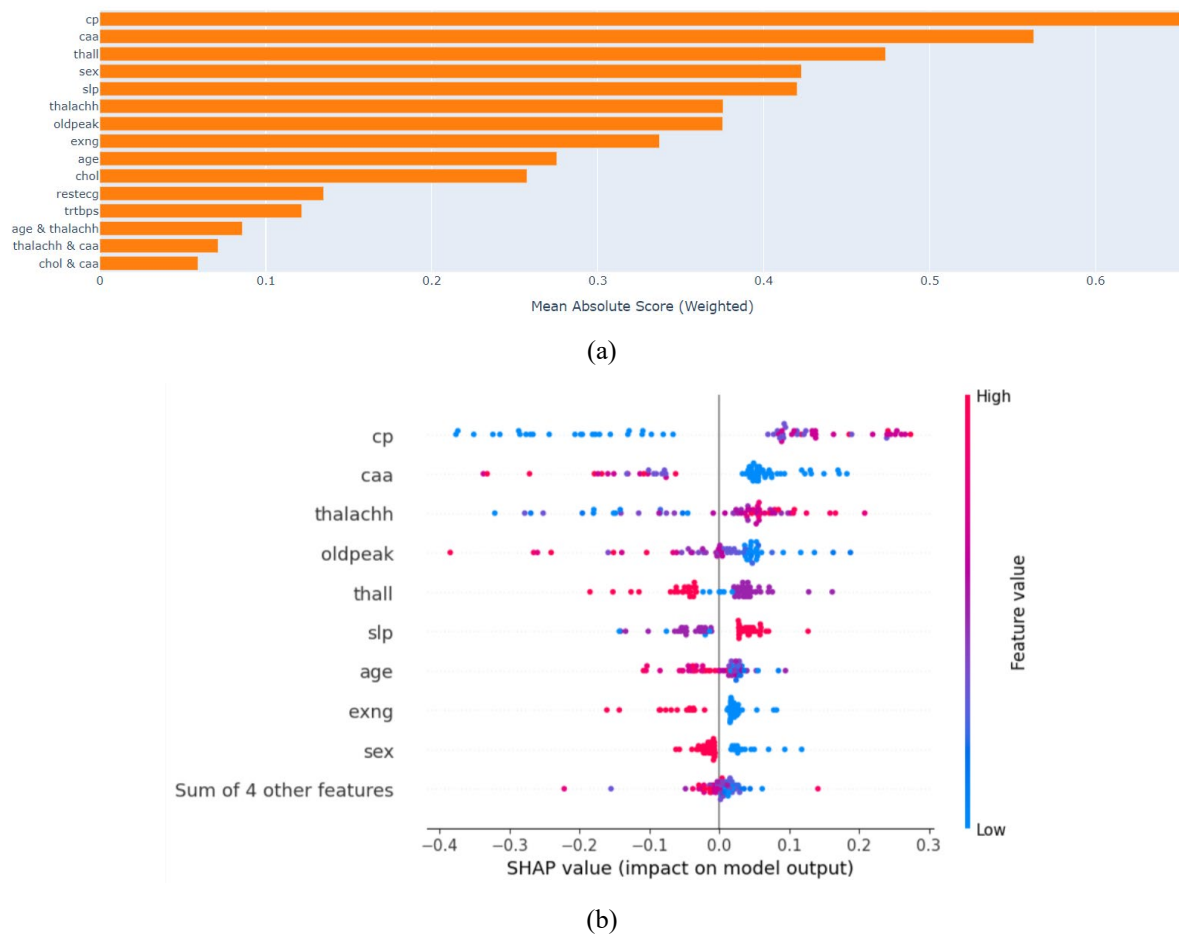


Figure 4. Global feature importance for the selected models: (a) Explainable boosting classifier and (b) RF with SHAP

After glass-box (explainable boosting classifier) and black-box (RF) models are selected, they need to be evaluated in terms of explainability. As mentioned before, due to their structure, glass-box models are easier to interpret the results individually than black-box models. In this context, a summary of the variables and information within the scope of the variables can be obtained with the explainable boosting model. In addition, comments can be obtained for individuals. However, this process cannot be performed directly in black-box models. For this reason, this information can be obtained using methods such as LIME and SHAP. Before explaining the examples locally, we can present the weights of the variations obtained by explainable boosting, as in Figure 4(a). The graph in Figure 4(a) shows importance weights on the x-axis and features on the y-axis. According to the information obtained from this graph, the “cp” variable corresponding to the chest pain type is the most crucial feature in the explainable boosting classifier model. In contrast, it seems that the least important feature is the parameter formed by the “chol&caa” combination. Similarly, importance levels of variables can be obtained for RF using SHAP. Since

different functions are used, the visualization of the results also varies. As can be seen from Figure 4(b), the most crucial variable is “cp.” In the results obtained by evaluating RF in terms of explainability, it is seen that the variables have similarities in terms of importance.

In addition to the importance of the weights of the variables for the model, another essential feature of XAI is that it provides the opportunity for *local interpretation*. In other words, it means explaining each observation value (in this study, patients whose heart attack risk is measured are expressed). In this context, an observation value in the data set is chosen randomly, and local explanations are given based on this observation value. Since the explainable boosting classifier and LIME use the same library, namely *interpret*, the resulting graphics have a similar visual structure. SHAP, on the other hand, shows a different visuality because it comes from a different library. As mentioned before, an observation value is chosen randomly. In this context, the results of the 11th observation are shown in Figure 5, respectively.

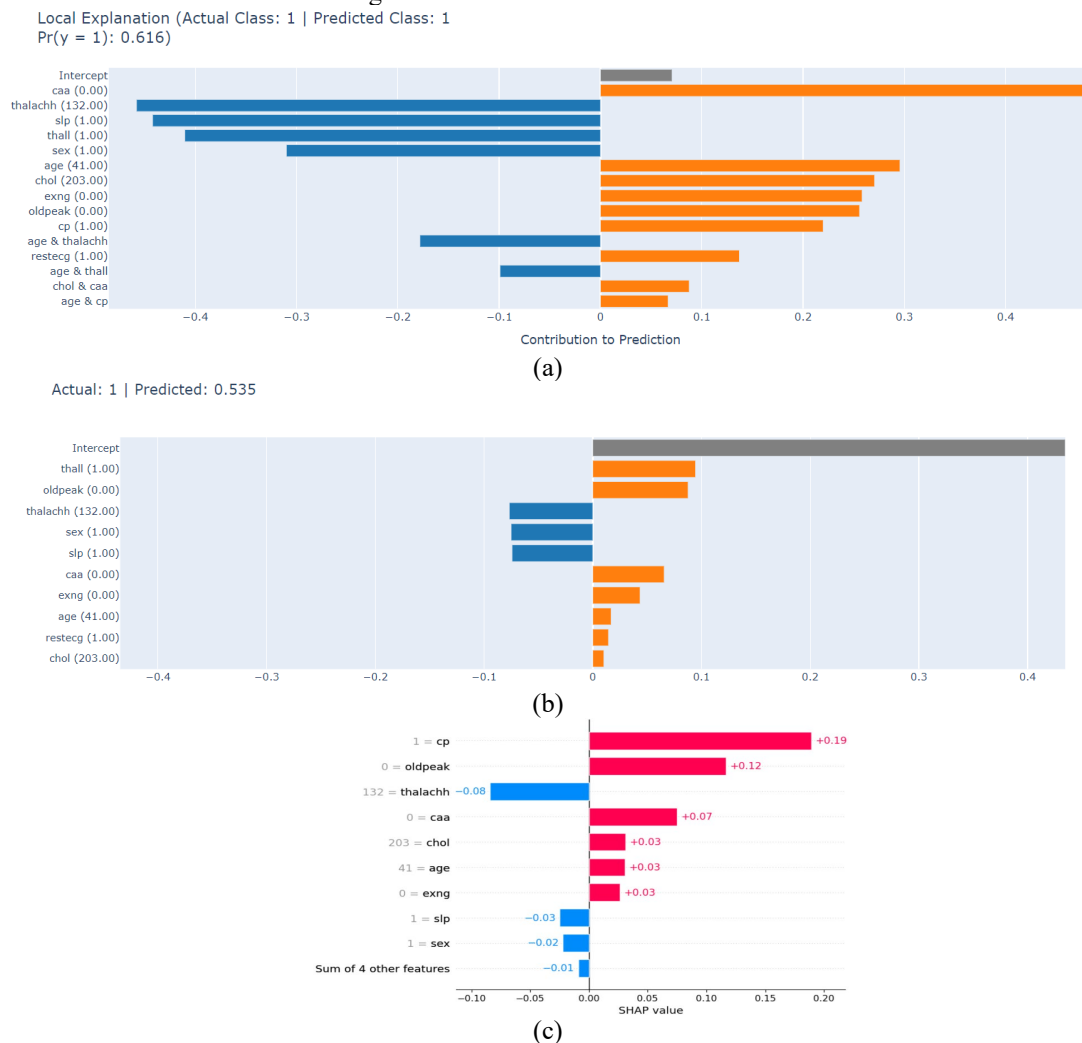


Figure 5. Local explanation of observation 11: (a) Explainable boosting classifier, (b) RF with LIME, and (c) RF with SHAP

In all three graphs in Figure 5, the x-axis shows the importance values of the variables for that patient. The y-axis shows the variables and the patient's values of those variables. For example, for this patient, the gender is shown as 1, and the age variable is expressed as 41. It can be seen from this information that it contains data from a 41-year-old male patient. In addition, the chest pain type value of this patient, expressed as " $cp = 1$ ", shows that the patient has typical angina. Moreover, the patient information and the observation value are estimated to have a high risk of heart attack in both models (explainable boosting classifier and RF). The importance of the variables that affect this estimate can also be obtained using the graphs in Figure 5.

The graphs in Figure 5 should be evaluated to show the effect of the variables in calculating the risk of having a heart attack after being used to obtain the patient's data. The most effective variable in classifying the risk in the explainable boosting algorithm for the patient whose information is given in Figure 5 is "caa," whereas the most effective variable for RF visualized with LIME is "Thall," and the variable visualized with SHAP is "cp." Similarly, it is seen that the influencing variables change as the effects of the variables change. Considering only the RF algorithm, using different visualization techniques for the same model also causes the variables and their effects to differ. According to the information obtained here, selecting and individually evaluating the method for problems that significantly impact human life, such as the models used to determine the risk of heart attack, is vital. In addition, presenting the results to decision-makers and physicians in this problem, with more than one explanatory model, will be effective in determining the treatments to be applied.

IV. APPLICATION AND RESULTS: PREDICTION OF HEART ATTACK RISK

In healthcare, AI rapidly transforms how we diagnose, predict, and classify diseases. ML techniques have proven to be powerful tools, delivering impressive accuracy. However, a major hurdle lies in the complexity of some AI models. These models, often called "black-box" models, can generate highly accurate predictions but lack transparency in their reasoning process. This lack of clarity can be concerning, particularly in critical areas like healthcare decision-making, affecting human life. XAI fills this gap by providing explanations for the complex calculations performed by these models. Ideally, researchers prefer models with complete interpretability, also known as "glass-box" models. However, these models may compromise accuracy for transparency. Finding the right balance between these two aspects is crucial for XAI implementation in healthcare.

This study addresses this challenge by investigating a model specifically designed to classify heart attack risk based on patient data. LR, NB, DT, and explainable gradient boosting algorithms, called glass-box models, are developed in this context. In addition to these models, RF, SVM, MLP, gradient boosting, and stochastic gradient boosting algorithms, classified as black-box models, are established. To select the best model among the developed models, fine-tuning is done by running the models with different parameters using grid search. The results of the fine-tuning processes, whose results provide the best models of the methods within themselves, are compared. Considering all the proposed models, it is seen that the best results are obtained with the explainable gradient boosting algorithm. In addition, the best performance in black-box models is obtained as RF when considering glass-box and black-box models. The ability to explain predictions on an individual patient basis is also essential. The RF algorithm is visualized using LIME and SHAP explanation methods to compare the patient-based description of these two methods. These methods are used to unveil the inner workings of the RF model, making its predictions more interpretable. When explainable gradient boosting and RF are compared, it is seen that the importance of the features changes, and different features are considered in determining the risk of heart attack. Moreover, this study highlights another crucial point: even within the same model (RF in this case), the choice of interpretability technique (LIME vs. SHAP) can influence the perceived importance of variables. This underscores the importance of careful method selection and individual evaluation, especially when dealing with high-impact domains like healthcare, where decisions can influence life-and-death situations. Furthermore, presenting physicians with multiple interpretable models can be highly beneficial. By considering diverse perspectives on the data, physicians gain a richer understanding of the factors contributing to a patient's heart attack risk. This comprehensive view can empower them to make more informed decisions about each individual's most effective treatment course.

That is, findings offer valuable insights for healthcare professionals. By understanding the strengths and limitations of different AI models, they can make informed decisions about which tool is best suited for their specific needs, striking a crucial balance between accuracy and interpretability in the healthcare field. Future studies can expand the study regarding data set size, application area, and model. In this context, by expanding the data set, the validity of the application can be ensured, and its applicability in large data can be addressed. Apart from heart attacks, applications can be created for different cardiovascular diseases and diseases in different areas. In addition, it can be applied not only in the health field but also in areas that affect human and living life and require individual evaluation. On more extensive data sets, deep learning algorithms

can also be used, and the performance of the explanation methods can be examined. Finally, by developing hybrid models, the model's accuracy rate and interpretability rate can be improved.

REFERENCES

- [1] Arrieta, A.B. et al. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. fusion*, vol. 58, pp. 82–115.
- [2] Longo, L. et al. (2024). Explainable artificial intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions. *Inf. Fusion*, p. 102301.
- [3] Langer, M. et al. (2021). What do we want from Explainable Artificial Intelligence (XAI)?--A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. *Artif. Intell.*, vol. 296, p. 103473.
- [4] Retzlaff, C.O. et al. (2024). Post-hoc vs ante-hoc explanations: xAI design guidelines for data scientists. *Cogn. Syst. Res.*, vol. 86, p. 101243.
- [5] Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.*, vol. 1, no. 5, pp. 206–215.
- [6] Cinà, G., Röber, T., Goedhart, R., and Birbil, I. (2022). Why we do need explainable ai for healthcare, *arXiv Prepr. arXiv2206.15363*.
- [7] Wysocki, O. et al. (2023). Assessing the communication gap between AI models and healthcare professionals: Explainability, utility and trust in AI-driven clinical decision-making. *Artif. Intell.*, vol. 316, p. 103839.
- [8] Nasarian, E., Alizadehsani, R., Acharya, U.R., and Tsui, K.-L. (2024). Designing interpretable ML system to enhance trust in healthcare: A systematic review to proposed responsible clinician-AI-collaboration framework. *Inf. Fusion*, p. 102412.
- [9] Riyaz, L., Butt, M.A., Zaman, M., and Ayob, O. (2022). Heart disease prediction using machine learning techniques: a quantitative review, in *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2021*, Volume 3, pp. 81–94.
- [10] Habebh, H. and Gohel, S. (2021). Machine learning in healthcare. *Curr. Genomics*, vol. 22, no. 4, p. 291.
- [11] Liang, Z., Zhang, G., Huang, J.X., and Hu, Q. V. (2014). Deep learning for healthcare decision making with EMRs, in *2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 556–559.
- [12] Patel, M.J., Andreescu, C., Price, J.C., Edelman, K.L., Reynolds III, C.F. and Aizenstein, H.J. (2015). Machine learning approaches for integrating clinical and imaging features in late-life depression classification and response prediction. *Int. J. Geriatr. Psychiatry*, vol. 30, no. 10, pp. 1056–1067.
- [13] Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, vol. 542, no. 7639, pp. 115–118, doi: 10.1038/nature21056.
- [14] o'Brien, A. R., Wilson, L.O.W., Burgio, G. and Bauer, D.C. (2019). Unlocking HDR-mediated nucleotide editing by identifying high-efficiency target sites using machine learning. *Sci. Rep.*, vol. 9, no. 1, p. 2788.
- [15] Pan, X., et al. (2020). ToxDL: deep learning using primary structure and domain embeddings for assessing protein toxicity. *Bioinformatics*, vol. 36, no. 21, pp. 5159–5168.
- [16] Ahsan, M.M. and Siddique, Z. (2022). Machine learning-based heart disease diagnosis: A systematic literature review. *Artif. Intell. Med.*, vol. 128, p. 102289.
- [17] Sahu, R., Mohanty, K., Dash, S.R., Brahmam, S., and Barra, P. (2023). Prediction of Heart Attack and Death: Comparison Between 1 DCNN and Conventional ML Approaches, in *2023 1st International Conference on Circuits, Power and Intelligent Systems (CCPIS)*, pp. 1–6.
- [18] Rao, K.D., Kumar, M.S.D., Akshitha, D. and Rao, K.N. (2022). Machine Learning Based Cardiovascular Disease Prediction, in *2022 International Conference on Computer, Power and Communications (ICCCPC)*, pp. 118–122.
- [19] Mahmud, I., Kabir, M.M., Mridha, M.F., Alfarhood, S., Safran, M. and Che, D. (2023). Cardiac Failure Forecasting Based on Clinical Data Using a Lightweight Machine Learning Metamodel. *Diagnostics*, vol. 13, no. 15, p. 2540.
- [20] Khan Mamun, M.M.R. and Elfouly, T. (2023). Detection of Cardiovascular Disease from Clinical Parameters Using a One-Dimensional Convolutional Neural Network. *Bioengineering*, vol. 10, no. 7, p. 796.
- [21] Ozcan, M. and Peker, S. (2023). A classification and regression tree algorithm for heart disease modeling and prediction. *Healthc. Anal.*, vol. 3, p. 100130.
- [22] Yu, H. (2023). Analysis and Prediction of Heart Disease Based on Machine Learning Algorithms, in *In 2023 8th International Conference on Intelligent Computing and Signal Processing (ICSP)*, pp. 1418–1423.
- [23] Saeed, W. and Omlin, C. (2023). Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities. *Knowledge-Based Syst.*, vol. 263, p. 110273.
- [24] Lundberg, S.M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.*, vol. 30.

- [25] Ribeiro, M.T., Singh, S. and Guestrin, C. (2016). ‘Why should i trust you?’ Explaining the predictions of any classifier, in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1135–1144.
- [26] Schwalbe, G. and Finzel, B. (2023). A comprehensive taxonomy for explainable artificial intelligence: a systematic survey of surveys on methods and concepts. *Data Min. Knowl. Discov.*, pp. 1–59.
- [27] James, G., Witten, D., Hastie, T., Tibshirani, R. and others (2013). An introduction to statistical learning, vol. 112. *Springer*.
- [28] Shah, K., Patel, H., Sanghvi, D., and Shah, M. (2020). A comparative analysis of logistic regression, random forest and KNN models for the text classification. *Augment. Hum. Res.*, vol. 5, no. 1, p. 12.
- [29] Aborisade, O. and Anwar, M. (2018). Classification for authorship of tweets by comparing logistic regression and naive bayes classifiers, in *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, pp. 269–276.
- [30] Stephens, C.R., Huerta, H.F. and Linares, A.R. (2018). When is the Naive Bayes approximation not so naive?. *Mach. Learn.*, vol. 107, pp. 397–441.
- [31] Jadhav, S.D. and Channe, H.P. (2016). Comparative study of K-NN, naive Bayes and decision tree classification techniques. *Int. J. Sci. Res.*, vol. 5, no. 1, pp. 1842–1845.
- [32] Dong, S. (2022). Virtual currency price prediction based on segmented integrated learning, in *2022 IEEE 2nd International Conference on Power, Electronics and Computer Applications (ICPECA)*, pp. 549–552.
- [33] Pattanayak, S., Loha, C., Hauchhum, L., and Sailo, L. (2021). Application of MLP-ANN models for estimating the higher heating value of bamboo biomass. *Biomass Convers. Biorefinery*, vol. 11, pp. 2499–2508.
- [34] Visani, G., Bagli, E., Chesani, F., Poluzzi, A. and Capuzzo, D. (2022). Statistical stability indices for LIME: Obtaining reliable explanations for machine learning models. *J. Oper. Res. Soc.*, vol. 73, no. 1, pp. 91–101.
- [35] Wang, D., Thunéll, S., Lindberg, U., Jiang, L., Trygg, J. and Tysklind, M. (2022). Towards better process management in wastewater treatment plants: Process analytics based on SHAP values for tree-based machine learning methods. *J. Environ. Manage.*, vol. 301, p. 113941.
- [36] Heart Disease Prediction, dataset by informatics-edu, 2020. [Online]. Available: <https://data.world/informatics-edu/heart-disease-prediction>. [Accessed: 11-May-2024].

Examining the Relationship Between Climate Change, Energy Consumption, and the Financial System: The Case of E7 Countries

Naime İrem DURAN ¹, Özge DEMİRKALE ²

¹ *İstanbul Beykent University, Faculty of Economics and Administrative Sciences, Department of Capital Markets, İstanbul, Türkiye*

² *İstanbul Aydın University, Faculty of Economics and Administrative Sciences, Department of Business Administration, İstanbul, Türkiye*

Abstract

This study aims to investigate the correlation among climate change, energy consumption, and the financial system, using the E7 countries as a case study. The E7 comprises emerging economies including Brazil, China, India, Russia, Turkey, Indonesia, and Mexico, which are the primary focus of this research. The research delves into the factors impacting CO2 emissions over the long term, spanning from 1992 to 2020. Results reveal a positive correlation between economic growth and fossil fuel usage with CO2 emissions, while a negative correlation is identified between CO2 emissions and variables such as renewable energy consumption, temperature changes, and capital investments. The study underscores the significance of sustainability and environmental policies for the E7 nations. Recommendations include increasing investments in renewable energy sources, encouraging the adoption of carbon-neutral transportation technologies, and supporting initiatives for forest conservation and afforestation. In conclusion, this study provides valuable insights into the relationship between climate change, energy consumption, and the financial system within E7 countries, offering policy recommendations for achieving sustainability.

Keywords: Climate Change, Energy Consumption, Financial System, E7 Countries

I. INTRODUCTION

With the onset of globalization, economies have become integrated with each other in many areas, primarily trade, finance, and technology. Developments in information technology have accelerated this change dramatically. One of the most important elements of economic progress and sustainability worldwide, given advancements in technology, is energy. However, today it is understood that a significant portion of global energy consumption (EC) is not sustainable, considering current technology and general energy sources. At this point, the importance of renewable energy (RE) sources comes to the forefront [1-2].

The use of RE sources plays a central role today in key issues such as economic growth (EG) and global temperature change. RE sources are generally derived from various forms and sources such as solar energy, wind energy, water energy (including river currents, sea and ocean waves), and biomass, biogas, or biochemical energy. These sources, while supporting a sustainable energy supply, also have the potential to promote EG. In particular, increasing investments in RE can bring economic benefits such as creating new job opportunities, transitioning to a green economy, and reducing energy dependency. However, in the face of global issues such as climate change and rising temperatures, the importance of using RE sources is becoming increasingly evident [3]. Investments in RE can also play a critical role in controlling global temperature change by reducing carbon emissions. In particular, reducing greenhouse gas emissions and decreasing the use of fossil fuels are fundamental steps in mitigating the effects of climate change worldwide and ensuring a sustainable future. The Paris Agreement aims to establish a system based on the principles of common responsibilities and differentiated contributions for all developed and developing countries. The agreement aims to establish and strengthen a sustainable system encompassing socio-economic activities globally in the period after 2020 to combat climate change.

The European Green Deal (EGD) was announced in 2019 with the aim of implementing the global climate change initiatives set forth by the Paris Climate Agreement. The Green Deal aims to reduce greenhouse gas emissions by 55% by 2030 and achieve carbon neutrality by 2050 as part of the EU's growth strategy. Within this framework, the goal is to ensure a transparent, fair, and inclusive transition by reducing pollution, preserving the lives of all living beings, and assisting companies in becoming world leaders in clean products and Technologies [4].

The long-term changes in average surface temperatures worldwide and atmospheric weather conditions are attributed to the increases in greenhouse gas emissions resulting from human activities such as industrialization, fossil fuel use, and deforestation. In particular, increases in the production and consumption of fossil fuels such as coal, oil, and natural gas have led to a rise in global temperatures by approximately 1.1°C compared to pre-industrial levels. The use of fossil fuels in energy production contributes to an increase in carbon dioxide concentration in the atmosphere. The reduction of this emissions has been identified as a priority and major concern globally. According to the Intergovernmental Panel on Climate Change, it has been stated that the primary cause of the increase in surface temperatures over the past century is the rise in human-induced greenhouse gas emissions. Reducing greenhouse gas emissions is recognized as the most important objective in addressing this issue by the international community [5-6]. Doval and Negulescu [7], demonstrated in their study that green finance and production practices are depicted as a formula for Europe's recovery from crises. Emerging economies, especially those like the E7, exhibit particular sensitivity to climate change threats due to rapidly increasing EC and resulting CO₂ emissions (CO₂E). Figure 1 shows that among the E7 countries, China has the highest CO₂E, while Turkey has the lowest emissions.

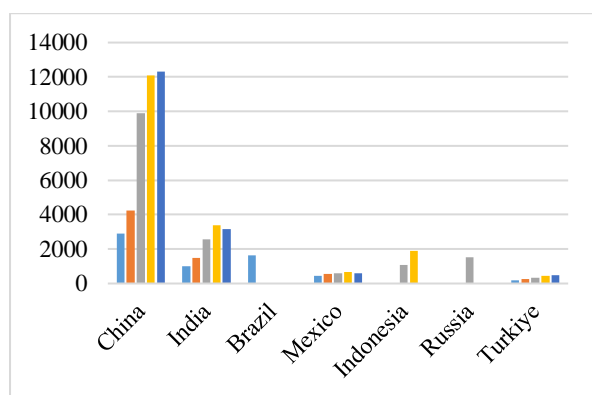


Figure 1. Greenhouse Gas Emissions

The relationship between temperature change and RE production is extremely important. Research indicates that global temperature increase has a positive impact

on RE production. For example, high temperatures and longer periods of sunshine can increase the efficiency of solar energy systems, thereby promoting RE production. Similarly, the efficiency of wind energy systems can also vary depending on temperature changes.

The efficient use of resources, especially in production, has become essential in all sectors as part of the fight against climate change. The effects of global environmental changes, particularly evident during the Covid-19 pandemic, have highlighted the need for resource efficiency. The aim of this study is to determine the relationship between climate change, EC, and the financial system in E-7 countries in the long term, and to provide recommendations on sustainability and environmental policies. In this context, the study will begin with a comprehensive literature review, followed by a detailed explanation of the methodology and data employed. Finally, the findings will be analyzed and interpreted.

Studies in the literature provide significant evidence regarding the relationship between climate change and EG. Research in the United States has established the foundational evidence for the relationship between climate change and EG. Strobl [9], conducted a study investigating the impact of hurricanes on EG in 409 coastal regions of the United States between 1970 and 2005. The research revealed that, on average, hurricanes reduced regional EG by 0.45 percentage points. Deryugina and Hsiang [10], employed the EKK method to examine the impact of daily temperature on EG in 48 states of the United States from 1969 to 2011. The study indicates that each 1°C increase in daily temperature above 15°C reduces daily economic productivity by approximately 1.7 percent. Colacito et al. [11], examined the impact of average seasonal temperatures on EG in the United States from 1957 to 2012 using the Panel EKK method. The study found that temperatures, particularly during summer months, have significant effects on EG. It was determined that a 1°F increase in average summer temperature is associated with a decrease in annual growth rate by 0.15-0.25 percentage points.

Research encompassing diverse countries provides strong evidence for understanding the relationship between climate change and EG. Bansal and Ochoa [12], examined the effects of temperature on EG for 147 countries between 1950 and 2007. The study found that a 1°C temperature shock reduced EG by approximately 0.9 percent, with stronger effects observed in countries closer to the equator. Dell et al. [13], investigated the impact of climate change on global economic activity between 1950 and 2006. They found that a 1°C increase in temperature would decrease EG by approximately 1.3 percent, with poorer countries being more adversely affected. According to Burke et al. [14], economic productivity peaks when

the annual average temperature is 13°C, with efficiency rapidly declining at higher temperature levels. Azam et al. [15], investigated the impact of environmental degradation on EG using annual data from 1971 to 2013 for China, the United States, India, and Japan. The results of the study indicate that carbon emissions have a significant positive relationship with EG for China, Japan, and the United States, while for India, it exhibits a significantly negative relationship. Sequeira et al. [16], found in their study examining the impact of climate change on economic and industrial outputs in countries that temperature increases do not lead to decreases in per capita income, except in poor countries. Henseler and Schumacher [17], state in their study, utilizing annual data for 103 countries from 1961 to 2010, that temperature is associated with per capita gross domestic product. Additionally, it was found that high temperature levels have strong effects in countries with low EG. Kahn et al. [18], investigated the impact of climate change on economic activities across countries using an ARDL model with a panel dataset covering 174 countries from 1960 to 2014. The study found that a permanent increase of 0.04°C in average global temperature would lead to a reduction of more than 7% in real per capita gross domestic product by the year 2100. Islam et al. [19], found in their study for the period 1990-2019 that in Saudi Arabia, carbon emissions and precipitation have a negative impact on EG, while temperature has a positive effect. Duan et al. [20], investigated the economic impact of climate change in China using the Panel Error Correction Model method. The study found that EG decreases by 0.78% for every 1°C increase in temperature. Additionally, the research indicates that EG will be influenced by a 0.86% increase for every 100 mm increase in precipitation and a 1.34% decrease for every 1% increase in humidity.

There are numerous studies in the literature addressing the complex relationship between RE consumption, temperature increase, and EG. These studies contribute to understanding the effects of climate change on EG and to the formulation of sustainable development policies. Zhang et al. [21], stated that renewable policies in Brazil and China have long-term positive effects on RE production and consumption. However, they found that Russia's RE policies are inadequate and decrease the growth in RE consumption in the long term. Keleş and Bilgen [2], concluded that Turkey's geographical location offers various advantages for widespread utilization of RE sources, particularly noting sufficient RE potential in terms of both fuel and electricity.

Cetin [22], found in his study, which examined the relationship between RE consumption and EG for E-7 countries during the period 1992-2012, that RE consumption has a positive impact on real GDP in E-7 countries. Marinaş et al. [23], investigated the relationship between EG and RE consumption for ten European Union (EU) member countries from Central and Eastern Europe during the period 1990-2014. The results indicate that in the short term, the dynamics of Gross Domestic Product (GDP) and RE Consumption (REC) are independent in Romania and Bulgaria, while increasing RE consumption improves EG in Hungary, Lithuania, and Slovenia. Klimenko et al. [24], state that in Russia, electricity production efficiency in thermal and nuclear power plants decreases as air temperature rises. According to climate model results, due to temperature increase, electricity production in thermal power plants and nuclear power plants will decrease by 6 billion kWh by 2050. The increase in air temperature during summer months will require higher EC for air conditioning, and by 2050, this figure will increase by approximately 6 billion kWh. Kasperowicz et al. [25], investigated the relationship between EG and RE consumption for 29 countries in Europe during the period 1995-2016. The results indicate that the use of RE as a global commodity is highly significant in the process of EG. Mele et al. [26], investigated the impact of increased RE production on the Brazilian economy, taking into account the periods of the SARS and Covid-19 pandemics. The results of the study indicate that the increasing use of RE sources may sustain economic recovery and create a Gross Domestic Product (GDP) momentum that outperforms other energy variables. Botzen et al. [27], found that under a high warming scenario, their study's results suggest predictions regarding electricity and gas consumption in Mexico. By the end of the century, electricity consumption is projected to increase by 12%, while gas consumption is expected to decrease by 10%, leading to a significant net economic cost of approximately 43 billion pesos per year. Rokicki et al. [3], stated in their study, which examined data from 23 countries in Central and Eastern Europe, that large-scale energy production from renewable sources could lead to a 60% decrease in temperature rise. Additionally, the study shows that such activities could result in a 90% increase in energy efficiency. Saqib et al. [28], revealed that technological modernization helped reduce pollution levels in E-7 countries from 1995 to 2019. Therefore, the study demonstrates that human development, technological innovation, and RE use were the most important variables for reducing carbon emissions during the examined period.

Yu et al. [1], results of the study found that emissions in E7 countries increased with the onset of development but later decreased due to potentially strong environmental regulatory policies implemented. The study also found that renewable energy, green innovations, environmental taxes, and technological innovations in all models had a significant and negative impact on carbon emissions in both the short and long term in E7 countries. Jia et al. [29], investigated the direct and indirect effects of RE consumption on EG using panel data from 90 countries participating in the Belt and Road Initiative between 2000 and 2019. The results of the study demonstrate that RE consumption directly contributes to EG. Germán-Soto et al. [30], investigated the impact of increasing temperatures on electricity consumption and economic development in Mexico between 2003 and 2019. The results of the study indicate that extreme weather conditions increase electricity demand. During extreme weather conditions, electricity consumption further increases due to fluctuations in electricity supply. The study concludes that temperatures have significant effects on economic development and electricity supply.

II. METHODOLOGY AND DATA

The factors considered for the analysis comprise carbon dioxide emissions, gross domestic product, renewable and non-RE usage, capital investment, and temperature variations in the E-7 countries from 1990 to 2020. Table 1 provides the units of measurement, abbreviations used within the article, and the sources from which the data were obtained.

Table 1. Information on Variables

Variable	Unit/Source	Symbol
CO2 Emission	CO2E from fuel combustion (MtCO2)/World Bank	CO ₂
Economic Growth	GDP (constant 2015 US\$)/World Bank	EG
Fossil Energy Consumption	Fossil fuel energy cons/mtoe/World Bank	FOSENCONS
Renewable Energy Consumption	RE consumption (% of total final Energy Cons.)/World Bank	RENENCONS
Gross Capital Formation	Gross capital formation (constant 2015 US\$)/World Bank	GCAP
Temperature Change	Annual Surface Temperature Change/IMF and Turkish Meteorological Service	TEMP

This research aims to examine the relationship between renewable and non-RE consumption and the financial indicators of countries, such as growth and capital formation, along with climate change in the E-7 countries. In investigating the relationship between CO₂E, EG, financial indicators, and temperature changes, the panel ARDL method, which does not require the series to be stationary and enables the analysis of both long-term and short-term relationships, was used. The results of the analysis revealed a long-term relationship between renewable and non-RE consumption and CO₂E, which are indicators of climate change. The analysis was conducted using Stata 18 software.

2.1. Method

In the study, a unit root test was initially conducted to prevent spurious regression due to the inclusion of data over time. Researchers working with panel unit root tests are divided into two groups, and the tests they develop are known as first-generation and second-generation tests [1]. First-generation tests assume no correlation between units, and if there is correlation between units, the power of these tests is weak. The most well-known of these tests include Levin, Lin, and Chu (2002) [31], Harris and Tzavalis (1999) [32], Breitung (2000) [33], Hadri (2000) [34], Im, Pesaran, and Shin (IPS, 2003) [35], Fisher ADF (Maddala and Wu 1999) [36], Fisher Philips, and Perron (Choi 2001) [37], panel unit root tests. The main characteristic of second-generation panel unit root tests is that they assume correlation between units. The most commonly used tests among these are Pesaran (2004) [38], Bai and Ng (2004) [39], Philips and Sul (2003) [40], Moon and Perron (2004) [41], panel unit root tests.

Consider the standard panel data model where y represents the dependent variable and x represents the independent variable, with i indicating the unit and t indicating the time dimension:

$$y_{it} = \alpha_i + \beta' x_{it} + u_{it}, i:1,2,..N, t=1,2,....T \quad (1)$$

Before estimating this model (1), to determine which unit root tests are more suitable, cross-sectional dependence was examined. Breusch and Pagan (1980) [42], proposed an LM statistic, which is valid for fixed N (individual) as T (time) $\rightarrow \infty$ and is given by Equation (2):

$$LM = T \sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij}^2 \quad (2)$$

Where $\hat{\rho}_{ij}$ is the sample estimate of pairwise correlation of the residuals:

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} = \frac{\sum_{t=1}^T \hat{u}_{it} \hat{u}_{jt}}{(\sum_{t=1}^T \hat{u}_{it}^2)^{1/2} (\sum_{t=1}^T \hat{u}_{jt}^2)^{1/2}} \quad (3)$$

and \hat{u}_{it} is the estimate of u_{it} in the Equation 1. However, this test typically exhibits significant size

distortions when N is large and T is finite; therefore, Pesaran (2004) [38], has proposed the CD (Cross Section Dependent) test which can be applied when both T and N are large, was conducted to determine the correlation between units. The statistic for Pesaran's CD test by equation 4.

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij} \right) \quad (4)$$

In equation 4, $\hat{\rho}_{ij}^2$, i, j represents the residual correlation coefficient between i and j . The test statistic is distributed as χ^2 with $d = \frac{N(N-1)}{2}$ degrees of freedom. Under the null hypothesis of no correlation between units, as N approaches infinity and T is sufficiently large, CD converges to $N(0,1)$.

The first step in the study was to investigate whether the series contained unit roots. For this purpose, correlations between units were initially examined. To determine the correlation between units, the Pesaran CD test and LM tests were conducted. The test results are summarized in Table 2.

Table 2. Cross Section Dependence Test's Results

Variables	Pesaran CD	LM	LM adj*	LM CD*
CO ₂ E	15.84***	27.23	1.55*	-1.38
EG	23.51** *	40.15	6.67*	-1.12
FOSENCON S	15.72*	34.02**	4.17***	-0.36
RENENCO NS	15.11*	45.71** *	8.92***	-.13
TEMP	11.07*	59.35** *	14.02** *	3.97** *
GCAP	9.25*	29.99*	2.70***	1.54*

***, **, * 1%, 5%, and 10% significance levels.

Upon evaluating Table 2, the Pesaran CD and LM adj* tests yield statistically significant results for all variables, indicating the presence of cross-sectional dependence. The LM test shows significant results for the variables EG, FOSENCONS, RENENCONS, TEMP, and GCAP, while the LM CD* test indicates significant results for the variables TEMP and GCAP. Particularly, the Temp and Renencons variables consistently show significant results across all tests, suggesting strong cross-sectional dependence in these variables.

In summary, the results of all conducted tests indicate that the null hypothesis, H_0 : There is no cross-sectional dependence, is rejected for many variables. This implies that cross-sectional dependence is generally

present. Considering this dependence is crucial for the reliability of modeling and results.

Pesaran (2007) [43], developed a simple method to eliminate cross-sectional correlation instead of estimating factor loadings. The Pesaran CADF(PESCADF) test is designed to check for the presence of unit roots (stationarity) in time series data within a panel data framework. In this method, he utilizes an extended version of the ADF regression with lagged cross-sectional averages, and the first difference(Δ) of this regression eliminates cross-sectional correlation. This generalized Dickey Fuller regression across cross-sections is referred to as the simple Cross-sectionally Augmented Dickey-Fuller regression. (CADF) regression:

$$\Delta Y_{it} = \alpha_i + \rho_i^* Y_{it-1} + d_0 \bar{Y}_{t-1} + d_1 \Delta \bar{Y}_t + \varepsilon_{it} \quad (5)$$

In this regression model, ΔY_{it} represents the first difference of the dependent variable, Y_{it-1} represents the one-period lag of the dependent variable, and, \bar{Y}_t is the average of all N observations at time t. The presence of lagged cross-sectional averages and first differences accounts for inter-unit correlation through factor structure. If there is autocorrelation in the error term or in the factor, the regression in the univariate case with the addition of lagged first differences of Y_{it} and \bar{Y}_t can be extended as follows:

$$\Delta Y_{it} = \alpha_i + \rho_i^* Y_{it-1} + d_0 \bar{Y}_{t-1} + d_1 \Delta \bar{Y}_t + \sum_{j=0}^p d_{j+1} \Delta \bar{Y}_{t-j} + \sum_{j=0}^p c_k \Delta \bar{Y}_{i,t-k} + \varepsilon_{it} \quad (6)$$

For this equation, the degree of extension can be selected using an information criterion or consecutive tests.

Pesaran (2007) introduces the CIPS(cross-sectionally augmented IPS) test, which remains robust in the presence of cross-sectional dependence among individual series in the panel. After estimating the CADF regression, the averages of the t-statistics of lagged variables are taken to obtain the CIPS statistic:

$$CIPS = \frac{1}{N} \sum_{i=1}^N CADF_i \quad (7)$$

For the CIPS statistic expressed as above, the combined asymptotic limit is not standard, and critical values have been calculated for various T and N values. In the test, the null hypothesis is that the variables are not stationary.

Table 3. Panel Unit Root Test Results

Variables	The Pesaran Cross-sectionally Augmented Dickey-Fuller regression (PESCADF) Results	
	Level	
	Constant ¹	Constant and Trend ²
CO2E	-2.246*	-2.737*
EG	-2.101	-1.835
FOSENCONS	-2.255	-2.461
RENENCONS	-1.472	-1.600
TEMP	-4.024**	-3.695***
GCAP	-2.160*	-2.622
	First Difference	
	Constant	Constant and Trend
CO2E	-2.942***	-2.898**
EG	-2.579*	-2.588
FOSENCONS	-2.687***	-2.732*
RENENCONS	-3.055***	-3.134***
TEMP	-5.651**	-3.965***
GCAP	3.628***	-3.796***
	Results	
CO2E	I(1)	
EG	I(1)	
FOSENCONS	I(1)	
RENENCONS	I(1)	
TEMP	I(0)	
GCAP	I(1)	

¹shows the result of the model estimated with the constant parameter. ² shows the the result of the model estimated with the constant parameter and trend. ***, **, * 1%, 5%, and 10% significance levels.

According to Table 3, the PESCADF results test the null hypothesis that the series have a unit root, indicating non-stationarity. For CO2E, EG, FOSENCONS, RENENCONS and GCAP, the null

hypothesis cannot be rejected at levels, implying these series are non-stationary. However, after taking the first difference, the null hypothesis is rejected, indicating these series become stationary (I(1)). In contrast, for the TEMP serie, the null hypothesis is rejected at levels, indicating it is stationary (I(0)) without needing differencing. This means TEMP does not have a unit root and is stable over time, unlike the other variables. After determining the stationarity levels of the series, cointegration tests are conducted to identify the presence of a long-term relationship between them.

The Mean Group (MG) estimation method, proposed by Pesaran and Smith (1995) [44], obtains the long-run parameter by averaging the long-run parameters of autoregressive distributed lag models created for each unit. Thus, it allows for the valuation of long-run parameters according to units. The Pooled Mean Group estimation method, proposed by Pesaran, Shin, and Smith (1999) [45], consists of a mixture of the MG estimator, which allows both slope and intercept parameters to vary across units, and the fixed effects estimator, which imposes that the slope parameter is constant while allowing the intercept parameter to vary. PMG keeps the long-run parameters fixed while allowing short-run parameters and error variances to vary across units [45], :

$$\Delta Y_{it} = \phi_i (Y_{it-1} - \beta_i' X_{it}) + \sum_{j=1}^{p-1} \lambda_{ij} \Delta Y_{it-j} + \sum_{j=0}^{q-1} \delta_{ij} \Delta X_{it-j} + \varepsilon_{it} \tag{8}$$

where: Y_{it} is the dependent variable for unit i at time t X_{it} is a vector of independent variables for unit i at time t , β_i are the long term coefficients. λ_{ij} and δ_{ij} are short run coefficients and ε_{it} is error term. ϕ_i is the error correction parameter, and the calculation of the ϕ_i , and β_i' parameters is as follows :

$$\phi_i = -(1 - \sum_{j=1}^p \lambda_{ij}) \tag{9}$$

$$\beta_i = (\sum_{j=0}^p \delta_{ij}) \tag{10}$$

Here, if ϕ_i is significant and negative, there is a long-run relationship between the dependent and independent variables. [46-47].

III. RESULTS AND DISCUSSION

Table 4 shows the PMG and MG estimates for the variables CO2, GDP, RE consumption, non- RE consumption, temperature increase and gross capital formation for E-7 countries.

Table 4: Panel ARDL Long Run Estimation

Dependent Variable: CO2E	Coefficients	
	PMG	MG
Long Run		
EG	0.055*	0.103*
FOSENCONS	0.997***	0.843***
RENENCONS	-0.0011*	-0.003
TEMP	-0.055**	-0.062
GCAP	-0.003***	-0.0008
EC Coefficient	-0.259***	-9.92***
Short Run		
EG	0.1089*	0.047
FOSENCONS	0.537***	0.323***
RENENCONS	-0.005***	-0.002
TEMP	0.007	0.026
GCAP	0.0004	0.0006
Hausman	3.78 (0.5817)	

t statistics in paranthesis: * p < 0.10, ** p < 0.05, *** p < 0.01

Before interpreting the results in Table 4, we first examined the error correction coefficient and found it to be statistically significant at the 1%, 5%, and 10% significance levels. In terms of indicating how quickly the series reach equilibrium, this parameter is important. Accordingly, approximately 25% of the imbalances occurring in one period will be corrected in the next period, leading towards approaching the long-term equilibrium. After conducting PMG and MG estimations, a Hausman test was performed to determine which estimator is consistent to use, and based on the test results, it was decided that using the PMG estimator is appropriate.

The estimated PMG model finds that all parameters are statistically significant. According to the results, a 1% increase in EG leads to a 0.055% increase in CO2E while a 1% increase in FOSENCONS results in a 0.99% increase in CO2E. Looking at other long-term parameters, a 1% increase in RENENCONS leads to a 0.0011% decrease in CO2E, while a 1% increase in TEMP results in a 0.05% decrease in CO2E emissions. A 1% increase in GCAP leads to a 0.003% decrease in CO2E emissions. In this context, Table 6 summarizes the long-term effects of all variables on the CO2E variable.

Table 5: Summary of the long-term effects of variables on CO₂E emission

Variable (%1 increase)	Effect on CO ₂ E Emissions in the long term
Economic Growth (EG) ↑	0.055% increase ↑
Fossil Energy Consumption (FOSENCOS) ↑	0.99% increase ↑
Renewable Energy Consumption (RENECONS) ↓	0.0011% decrease ↓
Temperature Change (TEMP) ↓	0.05% decrease ↓
Domestic Investment in Fixed Capital Formation ↓	0.003% decrease ↓

Upon a detailed evaluation of the results, it is known that there is a relationship between growth, one of the most important economic indicators of a country, and CO₂E, and this relationship is also complex depending on the country's EC and production structure. While previous literature has shown a positive relationship between EG and CO₂E, recent studies suggest that this relationship may vary. Given that the E-7 countries are developing economies, it is expected that there would be a positive relationship between EG and CO₂E. The relationship between fossil EC and CO₂E is also complex, but existing literature indicates that as fossil EC increases, CO₂E also increase, which is consistent with the findings in our model.

When examining studies on RE consumption, it is observed that there is a negative relationship between CO₂ and RE consumption. RE technologies such as solar panels and wind turbines do not rely on the combustion of fossil fuels, hence they do not contribute to CO₂E, or they minimize these emissions significantly. Our model also exhibits a similar effect, as the relationship between RE consumption and CO₂E is found to be negative as the relationship between RE consumption and CO₂E is found to be negative.

Gross fixed capital formation represents the total value of new fixed capital investments made in an economy during a specific period, typically involving the construction and expansion of production facilities, infrastructure projects, and other long-term asset investments. Such investments often occur in energy-intensive sectors, which can contribute to CO₂E. Existing literature suggests that the relationship between fixed capital and CO₂E tends to be positive in countries experiencing rapid EG and industrialization processes. However, in the PMG estimation conducted for these E-7 countries, this relationship was found to

be negative. Nevertheless, it is well known that several E-7 countries, particularly China, Brazil, India, and Mexico, also prioritize sustainable development and environmental protection. China, for instance, makes significant investments in RE to reduce greenhouse gas emissions and implements various environmental protection policies. India takes measures to combat air and water pollution and invests in renewable energy. Brazil makes efforts to protect the Amazon rainforest, reduce deforestation, and conserve biodiversity. Mexico implements various policies to increase energy efficiency, improve waste management, and protect natural resources. Additionally, Russia also conducts studies on the conservation of natural resources, environmental laws and regulations, and industrial waste control. Similarly, Turkey engages in activities related to waste management, clean energy production, forest conservation, and afforestation. Therefore, it is not surprising that the relationship between gross fixed capital formation and CO₂E appears negative in the model, as it can be influenced by the environmental policies implemented by these countries. In conclusion, the relationship between gross fixed capital and CO₂E in E-7 countries is complex, with policies varying from country to country. However, the small magnitude of this effect compared to other coefficients indicates that the impact of this variable on E-7 countries is minimal. The model estimations revealed a negative correlation between temperature change and CO₂E. While literature studies provide sample evidence that temperature change influences CO₂E, the general expectation is for this relationship to be positive. However, the negative relationship observed in the predictions for E-7 countries can be explained by their active engagement in environmental protection and sustainability efforts, particularly highlighted in countries like Brazil, China, India, and Mexico, as mentioned above. Additionally, each country has its own unique policies. Understanding this relationship necessitates taking into account the specific circumstances and policies of each country.

IV. CONCLUSION

Climate change and energy consumption are pivotal issues dominating global agendas, profoundly shaping the economic policies of nations. The objective of this study is to explore the long-term interrelationships among climate change, EC, and the financial system within the E-7 countries, and to offer recommendations for sustainability and environmental policies. The E-7 countries, which include Brazil, China, India, Russia, Turkey, Indonesia, and Mexico, are recognized as the seven major emerging economies. Given their substantial populations and growth potential, the development strategies implemented by these nations play a pivotal role in addressing climate change. To underscore this significance, the study analyzed factors influencing CO₂ emissions over the period from 1992 to 2020. The analysis identified that variables such as EG, Renewable EC, Fossil EC, fixed capital, and

temperature changes have a significant impact on CO₂ emissions in the long run. This research highlights the critical importance of integrating sustainable energy practices and robust environmental policies to mitigate the adverse effects of climate change. Additionally, the findings underscore the necessity for tailored policy measures that address the unique developmental and environmental challenges faced by emerging economies.

The study's findings highlight a clear connection between CO₂ emissions (CO₂E) and economic growth (EG), as well as fossil fuel consumption. Conversely, CO₂E shows an inverse relationship with RE consumption, temperature changes, and capital investment variables. A broad evaluation of the results confirms that they meet the expected outcomes. The E-7 nations present substantial potential for growth, and targeted investments in renewable energy sources—such as solar, wind, hydroelectric, and biomass—tailored to the specific needs of each country, could effectively reduce CO₂E. Additionally, adopting carbon-neutral transportation solutions, including public transit and electric vehicles, is promising for further CO₂E reduction. Strategies for forest conservation, reforestation, and the advancement of clean energy and environmental technologies would also be advantageous. E-7 countries can also cut CO₂E by investing in carbon-free transportation and enhancing public transit systems. Shifting to clean transportation technologies like electric vehicles will improve air quality and lower greenhouse gas emissions. Moreover, the growth and widespread adoption of clean energy and environmental Technologies will not only boost economic growth but also advance environmental sustainability. Additionally, the positive correlation between EG and fossil fuel usage with CO₂E suggests that financial support in these sectors mirrors their environmental impact. Considering the role of financial systems in addressing climate change, it is crucial to assess how financial institutions in E-7 countries handle environmental risks and opportunities, and how they can channel resources into sustainable projects. Sustainable finance, known for its pivotal role in combating climate change through green finance practices, necessitates that policymakers and financial institutions collaboratively develop and implement strategies to manage environmental risks and prioritize sustainability in investments.

Considering the study's limitations, it is important to acknowledge that the analysis is based on data available up to the end of 2020. This data constraint regarding renewable and fossil energy sources may affect the scope and depth of the findings. While this timeframe allows for a comprehensive examination of trends and relationships up to 2020, developments in energy consumption and policy changes occurring after this period are not included. Future research could gain

valuable insights by integrating more recent data, providing an updated perspective on the dynamics between renewable and fossil energy sources and their impact on CO₂ emissions. Additionally, detailed analyses for each country within the E-7 group could further enhance these findings. Expanding the study to include neighboring countries could also help identify which variables are influenced by climatic factors.

REFERENCES

- [1] Yu Y, Radulescu M, Ifelunini AI, Ogwu SO, Onwe JC, Jahanger A. Achieving Carbon Neutrality Pledge through Clean Energy Transition: Linking the Role of Green Innovation and Environmental Policy in E7 Countries. *Energies*. 2022; 15(17):6456. <https://doi.org/10.3390/en15176456>
- [2] Keleş, S.& Bilgen, S. (2012). Renewable energy sources in Turkey for climate change mitigation and energy sustainability. *Renewable and Sustainable Energy Reviews*, 16(7), p. 5199-5206, <https://doi.org/10.1016/j.rser.2012.05.026>
- [3] Rokicki, T., Koszela, G., Ochnio, L., Perkowski, A., Bórawski, P., Beldycka-Bórawska, A., Gradziuk, B., Gradziuk, P., Siedlecka, A., Szeberényi, A., & Dzikuć, M. (2022). Changes in the production of energy from renewable sources in the countries of Central and Eastern Europe. *Frontiers in Energy Research*.
- [4] European Commission. (2019). What is the European Green Deal. European Commission: https://ec.europa.eu/commission/presscorner/api/files/attachment/859152/What_is_the_European_Green_Deal_en.pdf.pdf (Erişim Tarihi: 04.02.2024).
- [5] Bankacılık Düzenleme ve Denetleme Kurumu (BDDK). (2021). 2022-2025 Sürdürülebilir Bankacılık Stratejik Planı.
- [6] European Commission. (2019). What is the European Green Deal. European Commission: https://ec.europa.eu/commission/presscorner/api/files/attachment/859152/What_is_the_European_Green_Deal_en.pdf.pdf (Erişim Tarihi: 04.02.2024).
- [7] Doval, E. ve Negulescu, O. (2014). A Model of Green Investments Approach. *Procedia Economics and Finance*, 15, 847-852.
- [8] https://www.climatewatchdata.org/ghg-emissions?end_year=2020&start_year=1990. Erişim Tarihi: 05.04.2024
- [9] Strobl, E. (2011). The economic growth impact of hurricanes: Evidence from US coastal counties. *Review of Economics and Statistics*, 93(2), 575-589.
- [10] Deryugina, T. & Hsiang, S. M. (2014). Does the environment still matter? Daily temperature and income in the United States (No. w20750). *National Bureau of Economic Research*.

- [11] Colacito, R., Hoffmann, B. & Phan, T. (2019). Temperature and growth: A panel analysis of the United States. *Journal of Money, Credit and Banking*, 51(2-3), 313-368.
- [12] Bansal, R. & Ochoa, M. (2011). Temperature, aggregate risk, and expected returns (No. w17575). National Bureau of Economic Research.
- [13] Dell, M., Jones, B. F. & Olken, B. A. (2012). Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics*, 4(3), 66-95.
- [14] Burke, M., Hsiang, S. M. & Miguel, E. (2015). Global non-linear effect of temperature on economic production. *Nature*, 527(7577), 235-239.
- [15] Azam, M., Khan, A. Q., Abdullah, H. B. & Qureshi, M. E. (2016). The impact of CO2 emissions on economic growth: evidence from selected higher CO2 emissions economies. *Environmental Science and Pollution Research*, 23(7), 6376-6389.
- [16] Sequeira, T. N., Santos, M. S. & Magalhães, M. (2018). Climate change and economic growth: a heterogeneous panel data approach. *Environmental Science and Pollution Research*, 25(23), 22725-22735.
- [17] Schumacher, I & Henseler, M. (2019). The impact of weather on economic growth and its production factors. *Climatic change*, 154(3), 417-433.
- [18] Kahn, M. E., Mohaddes, K., Ng, R. N., Pesaran, M. H., Raissi, M. & Yang, J. C. (2021). Long-term macroeconomic effects of climate change: A cross-country analysis. *Energy Economics*, 104, 105624.
- [19] Islam, M.M., Alharthi, M. & Murad, M.W. (2021). The effects of carbon emissions, rainfall, temperature, inflation, population, and unemployment on economic growth in Saudi Arabia: An ARDL investigation, *PLoS ONE* 16(4): e0248743.
- [20] Duan, H., Yuan, D., Cai, Z. & Wang, S. (2022). Valuing the impact of climate change on China's economic growth. *Economic Analysis and Policy*, 74, 155-174.
- [21] Zhang, H., Li, L., Cao, J., Zhao, M. & Wu, O. (2011). Comparison of renewable energy policy evolution among the BRICs. *Renewable and Sustainable Energy Reviews*, 15(9), p. 4904-4909, <https://doi.org/10.1016/j.rser.2011.07.063>.
- [22] Cetin, M. A. (2016) Renewable energy consumption-economic growth nexus in E-7 countries, *Energy Sources, Part B: Economics, Planning, and Policy*, 11:12, 1180-1185, DOI: 10.1080/15567249.2016.1156195
- [23] Marinaş M-C, Dinu M, Socol A-G, Socol C (2018) Renewable energy consumption and economic growth. Causality relationship in Central and Eastern European countries. *PLoS ONE* 13 (10): e0202951. <https://doi.org/10.1371/journal.pone.0202951>
- [24] Klimenko, V.V., Klimenko, A.V., Tereshin, A.G. *et al.* Impact of Climate Change on Energy Production, Distribution, and Consumption in Russia. *Therm. Eng.* 65, 247–257 (2018). <https://doi.org/10.1134/S0040601518050051>
- [25] Kasperowicz, R., Bilan, Y. & Štreimikienė, D. (2020). The renewable energy and economic growth nexus in European Countries. *Sustainable Development*, 28(5), p. 1086,10093, <https://doi.org/10.1002/sd.2060>
- [26] Mele, M., Gurrieri, A.R., Morelli, G. *et al.* Nature and climate change effects on economic growth: an LSTM experiment on renewable energy resources. *Environ Sci Pollut Res* 28, 41127–41134 (2021). <https://doi.org/10.1007/s11356-021-13337-3>
- [27] Botzen WJW, Nees T, Estrada F. Temperature Effects on Electricity and Gas Consumption: Empirical Evidence from Mexico and Projections under Future Climate Conditions. *Sustainability*. 2021; 13(1):305. <https://doi.org/10.3390/su13010305>
- [28] Saqib N, Usman M, Radulescu M, Sinisi CI, Secara CG and Tolea C (2022) Revisiting EKC hypothesis in context of renewable energy, human development and moderating role of technological innovations in E-7 countries?. *Front. Environ. Sci.* 10:1077658. doi: 10.3389/fenvs.2022.1077658
- [29] Jia H, Fan S, Xia M. The Impact of Renewable Energy Consumption on Economic Growth: Evidence from Countries along the Belt and Road. *Sustainability*. 2023; 15(11):8644. <https://doi.org/10.3390/su15118644>
- [30] Germán-Soto, V., Bordallo Favela, R.A. How rising temperatures affect electricity consumption and economic development in Mexico. *Environ Dev Sustain* (2024). <https://doi.org/10.1007/s10668-024-04527-3>
- [31] Levin, A., Lin, C.F. and Chu, C.S.J. (2002) Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. *Journal of econometrics*, 108, 1-24. [https://doi.org/10.1016/S0304-4076\(01\)00098-7](https://doi.org/10.1016/S0304-4076(01)00098-7)
- [32] Harris, R. D. F., & Tzavalis, E. (1999). Inference for Unit Roots in Dynamic Panels Where the Time Dimension Is Fixed. *Journal of Econometrics*, 91, 201-226. [https://doi.org/10.1016/S0304-4076\(98\)00076-1](https://doi.org/10.1016/S0304-4076(98)00076-1)
- [33] Breitung, J. (2000) The Local Power of Some Unit Root Tests for Panel Data. *Advances in Econometrics*, 15, 161-177. [http://dx.doi.org/10.1016/S0731-9053\(00\)15006-6](http://dx.doi.org/10.1016/S0731-9053(00)15006-6)
- [34] Hadri, K. (2000). Testing for Stationarity in Heterogeneous Panel Data. *The Econometrics Journal*, 3, 148-161. <https://doi.org/10.1111/1368-423X.00043>
- [35] Im, K., Pesaran, H. and Shin, Y. (2003), “Testing for Unit Roots in Heterogeneous Panels”, *Journal of Econometrics*, 115(1): 53-74.

- [36] Maddala, G.S. and Wu, S. (1999) A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. *Oxford Bulletin of Economics and Statistics*, 61, 631-652.
<http://dx.doi.org/10.1111/1468-0084.61.s1.13>
- [37] Choi, I. (2001) Unit Root Tests for Panel Data. *Journal of International Money and Finance*, 20, 249-272.
[http://dx.doi.org/10.1016/S0261-5606\(00\)00048-6](http://dx.doi.org/10.1016/S0261-5606(00)00048-6)
- [38] Peseran, M. (2004). General Diagnostic Tests for Corss Section Dependence in Panels. IZA Discussion Paper, 1240.
- [39] Bai, J. & Ng, Serena (2004). A PANIC Attack on Unit Roots and Cointegration, 72, 4, 1127-1177.
- [40] Phillips, P.C. and Sul, D. (2003). Dynamic Panel Estimation and Homogeneity Testing under Cross Section Dependence. *Econometrics Journal*, 6, 217-259.
<https://doi.org/10.1111/1368-423X.00108>
- [41] Moon, H. & Perron, B. (2004). Testing for a unit root in panels with dynamic factors. *Journal of Econometrics*, 122, 1, 81-126.
- [42] T. S. Breusch, A. R. Pagan (1980), The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics, *The Review of Economic Studies*, Volume 47, Issue 1, 1980, Pages 239–253. <https://doi.org/10.2307/2297111>
- [43] Pesaran, M.H. (2007), A Simple Panel Unit Root Test in the Presence of Cross Section Dependence. *Journal of Applied Econometrics*, 22, 265-312.
<https://doi.org/10.1002/jae.951>
- [44] Pesaran, M.H. and Smith, R. (1995), Estimating Lon-Run Relationship from Dynamic Heterogeneous Panels. *Journal of Econometrics*, 68, 79-113. [http://dx.doi.org/10.1016/0304-4076\(94\)01644-F](http://dx.doi.org/10.1016/0304-4076(94)01644-F)
- [45] Pesaran, M.H., Shin, Y. and Smith, R.P. (1999) Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. *Journal of the American Statistical Association*, 94, 621-634.
<https://doi.org/10.1080/01621459.1999.10474156>
- [46] Tatoğlu Y. F. (2013). Panel Veri Ekonometrisi Stata Uygulamalı,2. Baskı Beta Yayınları, İstanbul, s:214-220.
- [47] Tatoğlu Y. F. (2013).İleri Panel Veri Analizi ,2. Baskı Beta Yayınları, İstanbul, s:223-256.

Elektrokimyasal Ayırma Yöntemi ile Sulu Çözeltilerden Bor Giderim Prosesinin Optimizasyonu

Optimization of Boron Removal Process from Aqueous Solutions by Electrochemical Separation Method

Sevgi POLAT¹ 

¹Marmara Üniversitesi, Mühendislik Fakültesi, Kimya Mühendisliği Bölümü, İstanbul, Türkiye

Öz

Canlı yaşamının en temel ihtiyacı su, günümüzde teknolojinin ve endüstri alanlarının gelişmesiyle beraber ciddi oranda kirlenmektedir. Su kaynaklarının kısıtlı olması ve yakın gelecekte dünya nüfusunun su sıkıntısı yaşama olasılığının yüksek olduğu göz önüne alındığında, su kirliliğine sebep olabilen ve ülkemiz için önemli bir mineral olan borun sulardan uzaklaştırılması gerekmektedir. Bu kapsamda bu çalışmada, bor iyonlarının sulu çözeltilerden yenilikçi elektrokimyasal ayırma yöntemi kullanılarak giderilmesi incelenmiştir. Bu amaçla, aktif karbon ile kaplanarak hazırlanmış elektrotlar ve elektrokimyasal akış hücresi kullanılmıştır. Sürekli akış koşullarında bor adsorpsiyon verimini etkileyen akış hızı, hücreye uygulanan potansiyel ve bor konsantrasyonu parametrelerinin optimizasyonu Box-Behnken deney tasarımı yöntemi kullanılarak gerçekleştirilmiştir. 1-5 ml/dak akış hızı, 0,5-1,5 V potansiyel ve 10-90 ppm bor konsantrasyonu aralıklarında bor iyonlarının adsorpsiyon deneyleri yapılmıştır. Deney tasarımı sonuçlarına göre, çalışılan üç parametre arasından akış hızının bor giderimini etkileyen en önemli parametre olduğu belirlenmiştir. Bor adsorpsiyon veriminin en yüksek olduğu (%94,1) optimum koşullar 1 ml/dak akış hızı, 0,5 V potansiyel ve 50 ppm bor konsantrasyonu olarak tespit edilmiştir. Ayrıca, optimum koşullarda çalışılan elektrotların yüzey morfolojileri taramalı elektron mikroskobu kullanılarak karakterize edilmiştir.

Anahtar Kelimeler: Bor giderimi, Optimizasyon, Deney tasarımı, Elektrokimyasal ayırma yöntemi.

Abstract

Water, the most basic need of both humans and ecosystems, is being seriously polluted today with the development of technology and industry. Considering that water resources are limited and the probability of the world population experiencing water shortage soon is high, boron, which can cause water pollution and is an important mineral for our country, needs to be removed from the water. In this study, the removal of boron ions from aqueous solutions using an innovative electrochemical separation method was investigated. For this purpose, electrodes prepared by coating with activated carbon and electrochemical flow cell were used. Optimization of flow rate, potential applied to the cell and boron concentration parameters affecting boron adsorption efficiency under continuous flow conditions was carried out using the Box-Behnken experimental design method. Adsorption experiments of boron ions were carried out in the ranges of 1-5 ml/min flow rate, 0.5-1.5 V potential and 10-90 ppm boron concentration. According to the experimental design results, it was determined that flow rate was the most important parameter affecting boron removal among the three parameters studied. The optimum conditions with the highest boron adsorption efficiency (94.1%) were determined to be 1 ml/min flow rate, 0.5 V potential and 50 ppm boron concentration. Surface morphologies of the electrodes studied under optimum conditions were characterized using a scanning electron microscope.

Keywords: Boron removal, Optimization, Experimental design, Electrochemical separation method.

I. GİRİŞ

Türkiye açısından stratejik öneme sahip olan bor, cam, tekstil, deterjan, ilaç, deri, kozmetik gibi çok sayıda endüstride geniş kullanım alanına sahiptir [1,2]. Bu endüstrilerde ortaya çıkan atık suların taşımış olduğu bor miktarı belirli konsantrasyon değerlerinin üzerine çıktığında ve su yollarına karıştığında su kirliliğine neden olmaktadır [3]. Bu durum, bitki ve insan sağlığı açısından sorun yaratmakta ve çevresel açıdan risk oluşturmaktadır [4]. Örneğin bor bitkilerin büyümesinde önemli bir role sahipken, bitki türlerine bağlı olarak bor iyon konsantrasyonunun fazla olması bitkilerin gelişiminin durmasına neden olmaktadır [5].

Bitkileri korumak için, sulama suyu bor sınır değeri 1 ppm'den yüksek olmamalıdır [6]. Benzer şekilde içme sularında da bor iyon konsantrasyonunun fazla olması insan sağlığı açısından olumsuz durumlar yaratmakta ve kardiyovasküler, koroner, sinir ve üreme sistemlerinde hastalıkların oluşmasına sebep olmaktadır [7,8]. Dünya sağlık örgütü içme suları için bor limitini maksimum 2,4 mg/L olarak belirlemiştir [9-10]. Bu kapsamda, belirli ve düşük konsantrasyonlarda bor iyonlarının ortamda bulunması insan ve çevre sağlığı açısından olumlu etkiler yaratsa da yüksek miktarlarda bulunması ciddi sorunlara neden olmaktadır. Ayrıca, dünya nüfusunun hızla artması ve hızlı endüstrileşme suya olan ihtiyacın giderek artmasına neden olmaktadır. Buna karşın bu ihtiyacın karşılandığı su kaynakları ise sınırlıdır. Bu durum, su kaynaklarının iyi değerlendirilmesi ve tekrar kullanılabilir hale getirilmesini zorunlu hale getirmiştir.

Bor giderimi ile ilgili literatürde koagülasyon ve sedimentasyon [11], adsorpsiyon [12,13], ekstraksiyon [14], kristalizasyon [15], iyon değiştirme [16], membran prosesleri [17,18], ters osmoz [19], elektrodializ [20] gibi farklı arıtım yöntemleri kullanılarak yapılan çalışmalar mevcuttur. Bu teknolojiler arasında koagülasyon ve sedimentasyon yöntemleri ile bor bileşiklerinin sulu çözeltilerden giderimi istenen yüksek bor giderim kapasitesine sahip olmayıp düşük verimde bir ayırım gerçekleşmektedir. Ayrıca, çöktürme işlemi sırasında da fazla miktarda ilave kimyasal maddelere ihtiyaç duyulmaktadır. Bor gideriminde en çok kullanılan yöntemlerden biri olan adsorpsiyon yönteminde ise, iyon değiştirici reçineler ve selüloz gibi çok çeşitli inorganik ve organik adsorbanlar kullanılmıştır. Bor iyonlarının giderimi için geliştirilen iyon değiştirici reçinelerin maliyetleri yüksek olmasının yanı sıra rejenerasyon sonucu ortaya çıkan yıkama suyu da yeni bir atıksu oluşturmaktadır. Evaporasyon, kristalizasyon ve ekstraksiyon yöntemleri ise yüksek bor konsantrasyonuna sahip çözeltiler için daha etkili yöntemlerdir. Ters osmoz ve elektrodializ ise tek başına kullanıldıklarında bor yüksek verimde geri kazanılamamakta ve elektrokoagülasyon, kimyasal çöktürme, iyon değiştirme gibi ilave işlemlere ihtiyaç duyulmaktadır. Özellikle elektrokoagülasyonun yüksek bor konsantrasyonuna sahip çözeltiler için oldukça etkili olduğu ve borun uzaklaştırılmasında anyonik iyon değiştirici kullanımının olumlu etkisi olduğu gösterilmiştir. Adsorpsiyon, elektrokoagülasyon ve elektrodializ ile bor giderimine dair yapılan çalışmalarda başarılı sonuçlara ulaşılmakta ve günümüzde halen sıklıkla kullanılmalarına rağmen, adsorbentlerin rejenerasyon süreçlerindeki zorluk ve maliyet, iyon değiştirici reçine ihtiyaçları, membran

maliyetlerinin pahalı olması, sürekli bakım gerektirmeleri gibi nedenlerle çeşitli dezavantajlara da sahiptirler [11-20].

Bu çalışmada ise sulu çözeltilerden bor iyonlarının giderilmesi ve geri kazanılması için elektrotların rejenerasyon prosesi sırasında ek kimyasallara gerek olmaması (yani sistem yan kirletici üretmediği için ek maliyet oluşmaması), adsorpsiyon ve desorpsiyon sürecinin hızlı olması, elektrotların tekrar tekrar kullanılabilir olması ve buna bağlı olarak düşük karbon ayak izi ve ekonomik fayda sağlaması, düşük elektriksel potansiyel değerlerinde çalışılarak düşük enerji tüketimi ve prosesin modüler olması gibi su ekonomisi ve atık sürdürülebilirliği açısından diğer yöntemlerle karşılaştırıldığında avantajlara sahip olması nedeniyle elektrokimyasal ayırma yöntemi kullanılmıştır. Bu kapsamda, sürekli akış koşullarında elektrokimyasal akış hücresi kullanılarak sulu çözeltilerden bor iyonlarının giderimi incelenmiştir. Bu yöntem ile literatürde mevcut olmayan maksimum bor adsorpsiyon veriminin sağlanabileceği proses koşulları deney tasarımı yöntemi kullanılarak incelenmiş ve proses parametrelerinin optimizasyonu yapılmıştır.

Deney tasarımı uygulaması genel olarak problemin, bağımsız parametrelerin ve bunların seviyelerinin belirlenmesi, cevap değişkeninin seçilmesi ve uygun deney tasarım yönteminin belirlenerek deney yapılması ve verilerin analiz edilmesini kapsamaktadır. Box-Behnken deney tasarımı, üç seviyeli faktörlerin ve ikinci seviyeden terimlerin söz konusu olduğu sistemlerin değerlendirilmesinde kullanılan istatistiksel bir yöntem olup minimum deney sayısı ile maksimum bilgiyi elde ederek, araştırmacının prosesin davranışını belirlemesini sağlamaktadır [21-24]. Bu çalışmada da Box-Behnken deney tasarım yöntemi kullanılarak akış hızı, potansiyel ve bor konsantrasyonu gibi proses parametrelerinin bor iyonlarının adsorpsiyon verimine etkisi incelenmiştir.

II. MATERYAL VE METOD

2.1. Deneyin Yapılışı

Bu çalışmada, bor iyonlarını ayırmak için kullanılacak elektrotların hazırlanmasında Toray karbon kağıdı, aktif karbon, karbon siyahı, poliviniliden florür (PVDF) ve aseton kullanılmıştır. Şekil 1'de gösterildiği üzere, ağırlıkça aktif karbon: karbon siyahı: PVDF oranı 85:10:5 olacak şekilde tartılmıştır. Toz halde bulunan bu karışıma çözücü olarak kullanılan 20 ml aseton eklenmiş, manyetik karıştırıcı aracılığıyla süspansiyonun 2 saat karıştırılması sağlanmıştır. Daha sonra bu süspansiyon daldırılmalı ultrasonik prob kullanılarak 1 saat ultrasonikasyon işlemine tabi tutulmuştur. Damlatma (drop casting) yöntemi kullanılarak aktif yüzey alanı $5 \times 5 \text{ cm}^2$ olan elektrotlar hazırlanmıştır. Hazırlanan ve kurutulmuş elektrotlar

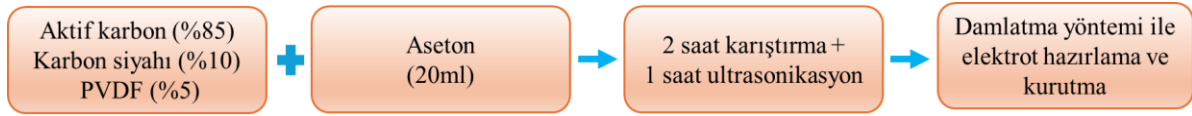
elektrokimyasal akış hücresine yerleştirilmiştir. Akış hücresi Rhinoceros 3D modelling yazılımı kullanılarak tasarlanmış ve akrilikten oluşan plaka, aktif karbon ile kaplanmış elektrot, titanyum kollektör ve polimerden yapılmış dağıtıcıdan oluşmaktadır.

Bor iyonlarının adsorpsiyon deneyleri, bu akış hücresi kullanılarak, en az 5 döngü olacak şekilde sürekli akışın olduğu koşullarda Şekil 2’de şematik olarak gösterilen deney düzeneği kullanılarak yürütülmüştür. Elektrokimyasal ayırma prosesi sırasıyla besleme tankı, peristaltik pompa, elektrokimyasal akış hücresi, güç kaynağı, iletkenlik ölçer, data kaydedici ve çıkış tankından oluşmaktadır. Akış hücresini terk eden çözeltinin iletkenlik değeri sürekli olarak ölçülmüş ve adsorplanan bor miktarını belirleyebilmek için numuneler alınarak iyon derişimi indüktif eşleşmiş plazma atomik emisyon spektroskopisi (ICP-OES) kullanılarak belirlenmiştir. Adsorpsiyon verimi aşağıda verilen eşitlik kullanılarak hesaplanmıştır. C_i ve C_t sırasıyla başlangıç ve çıkış akımındaki bor iyon konsantrasyonunu göstermektedir.

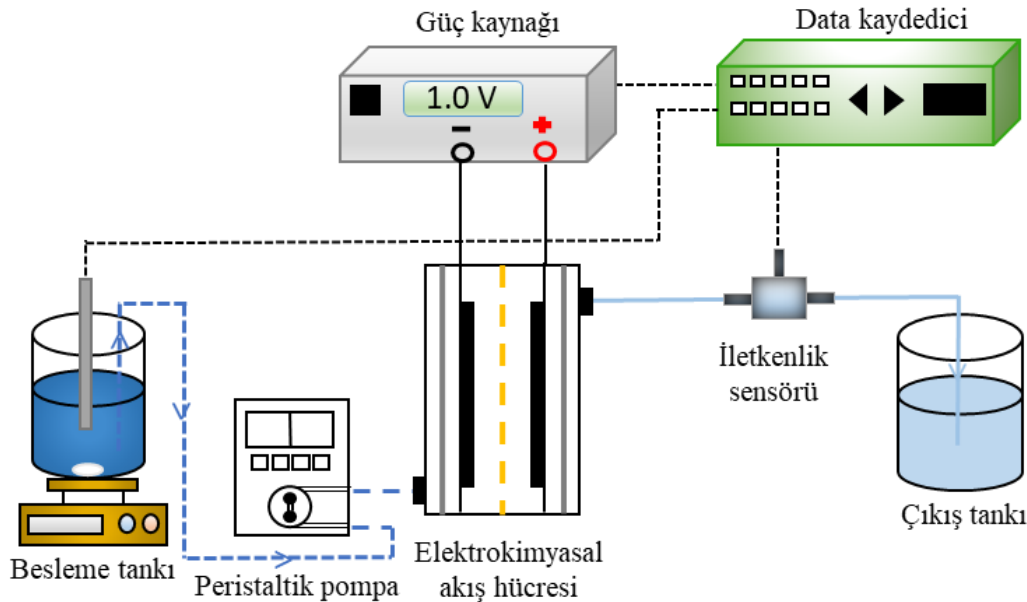
$$Verim(\%) = \frac{(C_i - C_t)}{C_i} \times 100 \quad (1)$$

2.2. Deney Tasarımı

Bu çalışmada, bor iyon giderimine etki eden parametreler yapılan ön deneyler sonucunda akış hızı, hücre potansiyeli ve iyon konsantrasyonu olarak belirlenmiştir. En yüksek adsorpsiyon verimini elde edebilmek için yanıt yüzey metodolojisi ve Box-Behnken deney tasarım modeli kullanılmıştır. Deney tasarımı ve veri analizleri için Design Expert 10.0 yazılımı kullanılmıştır. Adsorpsiyon deneyleri için merkez noktası 3 tekrar içeren toplam 15 deney yapılmıştır. Akış hızı, potansiyel ve konsantrasyon değişkenleri sırasıyla A, B ve C olarak tanımlanmış ve çalışılan aralık ve düzeyler Tablo 1’de verilmiştir. Değişken değerlerinin istenen aralıkları -1 (minimum), 0 (merkez nokta) ve +1 (maksimum) olarak kodlanmıştır. Box-Behnken deney tasarımına ait deney algoritması Tablo 2’de verilmiştir. İkinci dereceden polinomu deneysel verilere uydurmak ve ilgili model terimlerini belirlemek için doğrusal olmayan bir regresyon yöntemi kullanılmış ve model denklemin istatistiksel açıdan anlamlılığı varyans analiziyle (ANOVA) incelenmiştir.



Şekil 1. Bor iyonlarının adsorpsiyon için kullanılacak elektrotların hazırlanma diyagramı



Şekil 2. Deney düzeneği

Tablo 1. Bağımsız değişkenlerin deneysel aralığı ve düzeyleri

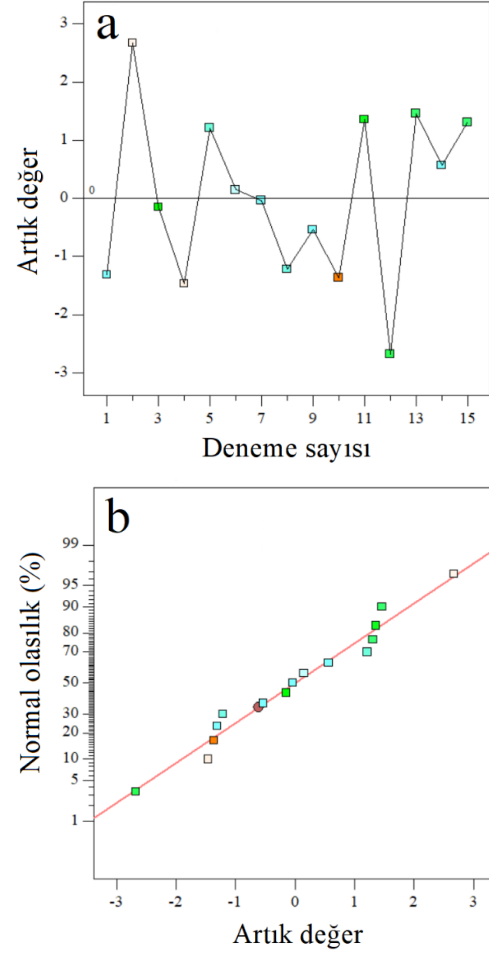
Değişkenler	Sembol	Aralık ve Düzeyler		
		-1	0	1
Akış Hızı	A	1,0	3	5
Potansiyel	B	0,5	1	1,5
Konsantrasyon	C	10	50	90

III. SONUÇLAR VE TARTIŞMA

3.1. Box-Behnken Deney Tasarım Sonuçları

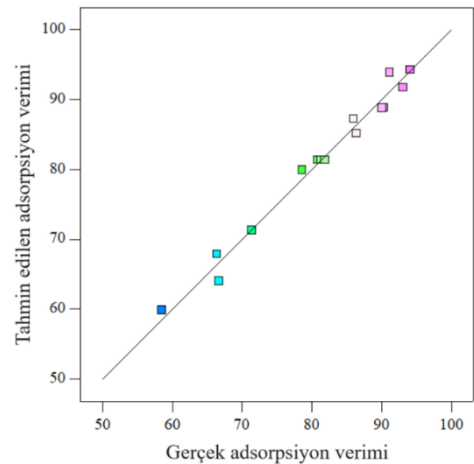
Elektrokimyasal ayırma yöntemi kullanılarak sulu çözeltilerden bor iyonu giderim sürecine etki eden akış hızı, hücre potansiyeli ve besleme çözeltisi konsantrasyonu parametrelerin etkisinin incelendiği bu çalışmada Box-Behnken deney tasarım yöntemi kullanılmış ve farklı koşullarda elde edilen adsorpsiyon verimi sonuçları Tablo 2’de verilmiştir. Bu üç bağımsız değişken ile proses cevabı arasındaki ilişki varyans analizi ile belirlenmiş, istatistiksel açıdan önemi %95 güven aralığıyla F-değeri ile analiz edilmiş ve yanıt yüzey ikinci dereceden modeli için ANOVA sonuçları Tablo 3’te verilmiştir. Kurulan matematik model 31,03 değeri ile %95 güven aralığında istatistiksel olarak anlamlıdır. Prob > F değeri yani P değerinin 0,05’ten küçük olması faktörlerin önemli olduğunu göstermesi nedeniyle A, C, A² ve B² modelde etkin ve anlamlı parametrelerdir. Bu parametreler arasında akış hızı (A) 238 F değeri ile bor iyonlarının geri kazanılmasında en etkili parametredir. Çalışılan deney koşulları aralığında hücreye uygulanan potansiyel ve bu potansiyelin akış hızı ve konsantrasyon ile etkileşimi adsorpsiyon verimini önemli ölçüde etkilememiştir. Ayrıca, Tablo 3’te verilen tanımlayıcı istatistikler incelendiğinde kurulan matematiksel modelin korelasyon katsayısı (R²=0.9824) deneysel verileri açıklayabilmiş ve adsorpsiyon verimi ile bağımsız değişkenler arasında iyi bir korelasyon olduğunu göstermiştir. Design Expert programı yeterli kesinlik (adequate precision) değerinin 4’ten büyük olmasını önermektedir. Bu çalışmada, bu değer 17,386 olması nedeniyle modelin uyumluluğu desteklenmektedir.

Şekil 3a’da her bir deneye artık değerleri veren grafik, Şekil 3b’de ise sulu çözeltilerden bor iyonlarının giderim verimi için artık değer normal olasılık grafiği verilmiştir. Bu grafikler oluşturulan matematiksel modelin uygunluğunu kontrol etmek amacıyla çizilmiştir. Artık değer normal olasılık grafiğinde noktaların bir doğru boyunca dizilmesi artık değerlerin normal dağılıma uyduğunu göstermektedir. Bor adsorpsiyon verimi için bu durum sağlanmakta, noktalar düz bir doğruyu temsil ettiğinden artık değerler normal dağılıma uygunluk göstermektedir.



Şekil 3. a) Deneme sayısına karşı gelen artık değerler grafiği ve b) bor iyonlarının adsorpsiyon verimi için artık değer normal olasılık grafiği

Şekil 4’te verilen grafikte deneysel olarak elde edilen adsorpsiyon verimi değerlerine karşılık önerilen matematiksel model ile hesaplanmış olan tahmini değerler görülmektedir. Tahmin edilen değerlerin deneysel değerlere yakın olması, geliştirilen modelin elektrokimyasal ayırma yöntemi ile bor gideriminin incelenmesinde etkili olduğunu göstermektedir.



Şekil 4. Bor iyonlarının adsorpsiyon verimi için deneysel ve tahmini değerlerin karşılaştırılması

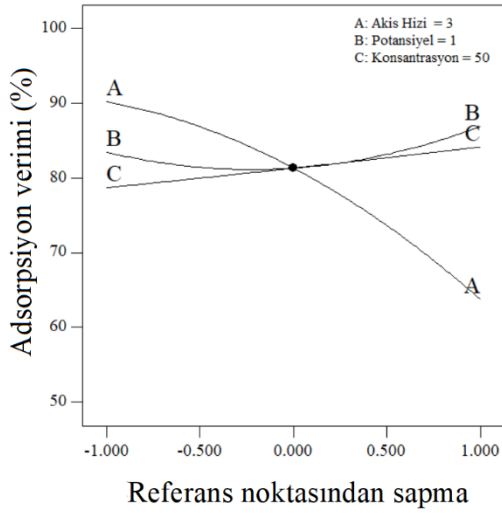
Tablo 2. Box-Behnken deney tasarımı algoritması

Deney No	Gerçek Değerler			Kodlanmış Değerler			Cevap Adsorpsiyon Verimi (%)
	Akış Hızı	Potansiyel	Konsantrasyon	A	B	C	
1	3	0,5	10	0	-1	-1	78,6
2	3	1,5	10	0	+1	-1	86,4
3	3	0,5	90	0	-1	+1	86,0
4	3	1,5	90	0	+1	+1	90,1
5	1	0,5	50	-1	-1	0	94,1
6	1	1,5	50	-1	+1	0	91,2
7	5	0,5	50	+1	-1	0	66,7
8	5	1,5	50	+1	+1	0	71,4
9	1	1,0	10	-1	0	-1	90,3
10	1	1,0	90	-1	0	+1	93,1
11	5	1,0	10	+1	0	-1	58,5
12	5	1,0	90	+1	0	+1	66,4
13	3	1,0	50	0	0	0	81,3
14	3	1,0	50	0	0	0	80,8
15	3	1,0	50	0	0	0	81,9

Tablo 3. ANOVA sonuçları

Kaynak	Kareler Toplamı	df	Kareler Ortalaması	F-Değeri	P-değeri Prob > F
Model	1638.72	9	182.08	31.03	0.0007
A-Akış Hızı	1396.56	1	1396.56	238.00	< 0.0001
B- Potansiyel	23.46	1	23.46	4.00	0.1020
C-Konsantrasyon	59.40	1	59.40	10.12	0.0245
AB	14.44	1	14.44	2.46	0.1775
AC	6.50	1	6.50	1.11	0.3407
BC	3.42	1	3.42	0.58	0.4795
A ²	69.60	1	69.60	11.86	0.0184
B ²	54.97	1	54.97	9.37	0.0281
C ²	0.026	1	0.026	4.370E-003	0.9499
Artık	29.34	5	5.87	-	-
Uyum Eksikliği	28.73	3	9.58	31.57	0.0309
R ²	0.9824				

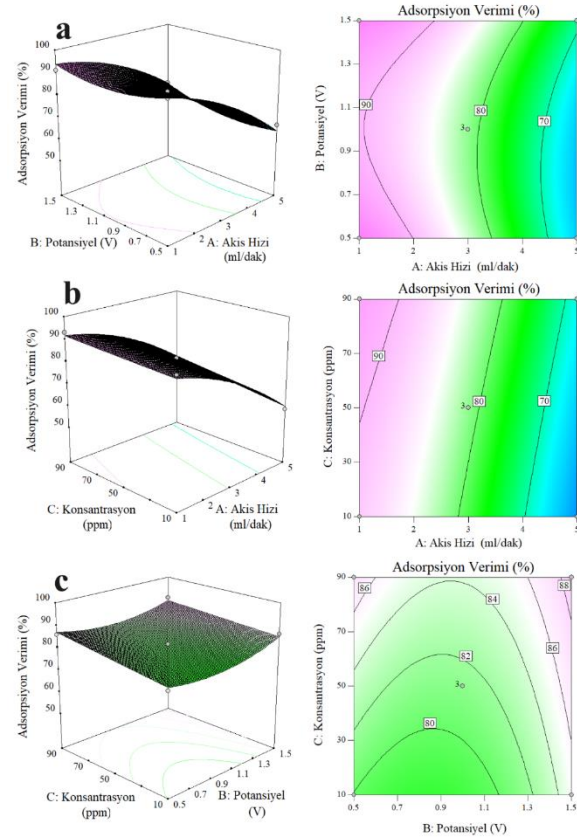
Şekil 5'te elektrokimyasal ayırma yöntemi ile bor iyonlarının adsorpsiyon verimini artırabilmek için geliştirilen modele ait değişkenlerin etkilerini açıklayan pertürbasyon eğrisi verilmiştir. Pertürbasyon eğrisinde 3, 1 ve 50 olarak verilen değerler, değişkenlere ait en düşük ve en yüksek değerlerin aritmetik ortalamasını ifade etmektedir. Bu değerler, pertürbasyon eğrisinde 0 noktasına karşılık gelen değerleri belirtmektedir. Pertürbasyon eğrileri, geliştirilen modele ait değişkenlerin, cevaplar üzerindeki sonuçlarını daha iyi açıklayabilmek için kullanılmaktadır. Şekil 5 incelendiğinde akış hızı 3-5 ml/dak aralığındayken adsorpsiyon veriminin önemli ölçüde azaldığı yani artan akış hızının verime negatif etkisi olduğu belirlenmiştir. 1,0-1,5 V potansiyel ve 10-50 ppm konsantrasyon değerlerinde ise bor giderim veriminde artış gözlemlenmiştir.



Şekil 5. Bor iyonlarının adsorpsiyon verimi için geliştirilen modelin pertürbasyon eğrisi

Şekil 6'da akış hızı-potansiyel, akış hızı-konsantrasyon ve potansiyel-konsantrasyon parametrelerinin bor iyonlarının adsorpsiyon verimine etkisini gösteren kontur ve 3 boyutlu yüzey grafikleri gösterilmiştir. Üç faktörlü Box-Behnken tasarımına göre en yüksek adsorpsiyon verimi için optimum çalışma koşullarının 1 ml/dak akış hızı, 0,5 V potansiyel ve 50 ppm bor konsantrasyonu olduğu belirlenmiş ve en yüksek adsorpsiyon verimi %94,1 olarak elde edilmiştir. Çalışılan üç bağımsız değişkenin etkileri karşılaştırıldığında, her bir parametrenin borun geri kazanımına etkisi olsa da akış hızının en önemli etkiye sahip parametre olarak belirlenmiştir. Adsorpsiyon verimi artan akış hızıyla önemli ölçüde düşüş göstermektedir. Maksimum bor adsorpsiyon verimi de en düşük akış hızı olan 1ml/dak hızında bulunmuştur. Literatürde aktif karbon kullanılarak bor iyon giderimi üzerine yapılan çalışmalar incelendiğinde, Kluczka ve arkadaşları [25] tarafından modifiye edilmiş aktif karbon kullanıldığında maksimum bor giderimi %51, Halim ve arkadaşları [26] tarafından ise maksimum adsorpsiyon verimi pH 5,5'te ~%40, Irawan ve

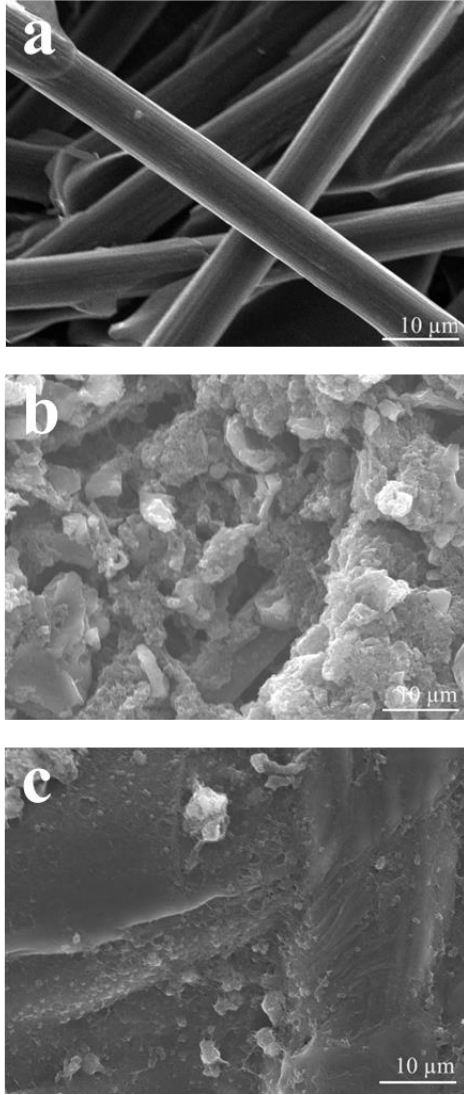
arkadaşları [27] tarafından ise kimyasal çöktürme yöntemini kullanıldığında 750 ppm bor konsantrasyonundaki çözelti için bu değer %87 olarak belirlenmiştir.



Şekil 6. a) Akış hızı-potansiyel, b) akış hızı-konsantrasyon ve c) potansiyel-konsantrasyon parametrelerinin bor iyonlarının adsorpsiyon verimine etkisini gösteren kontur ve 3 boyutlu yüzey grafikleri

3.2. Karakterizasyon Sonuçları

Box-Behnken deney tasarımı ile belirlenen ve bor adsorpsiyon veriminin en yüksek olduğu koşullarda elektrotların yüzey özellikleri taramalı elektron mikroskopu (SEM) kullanılarak belirlenmiştir. SEM analizi elektrot yüzeylerini kaplayan malzemelerin element türleri ve miktarlarının belirlenmesinin yanı sıra yüzey topografyalarının saptanabilmesine de olanak sağlamaktadır. Şekil 7'de işlem görmemiş karbon kağıdı, aktif karbon ile kaplanmış karbon kağıdı ve adsorpsiyon sonucunda akış hücresinin anot bölgesindeki karbon kağıdının SEM görüntüleri verilmiştir. Literatür ile uyumlu olarak [28,29], Toray karbon kağıdının yüzeyi düzgün yüzeyli ve homojen dağılımlı fiber yapılarından oluşmaktadır. Karbon kâğıdı, aktif karbon ile kaplandığında yüzey değişmiş ve fiber yapılar yerini homojen olmayan gözenekli bir yüzeye bırakmıştır. Bor iyonlarının adsorpsiyon işleminden sonra yüzeyde bulunan boşluklar tamamen kapanmış, homojen yüzeyli düz bir yüzey elde edilmiştir.



Şekil 7. a) Toray karbon kağıdı, b) aktif karbon ile kaplanmış çalışma elektrodu ve c) bor adsorpsiyonu sonrasında çalışma elektrodunun SEM görüntüleri

IV. SONUÇ

Bu çalışmada, sürekli akış koşullarında elektrokimyasal ayırma yöntemi ile bor gideriminin Box-Behnken deney tasarım yöntemi ile optimizasyonu incelenmiştir. Bu kapsamda akış hızı, hücre voltajı ve bor konsantrasyonunun adsorpsiyon verimine etkileri araştırılmıştır. Bu değişkenler arasında akış hızının bor adsorpsiyon veriminde etkili parametre olduğu belirlenmiştir. Maksimum adsorpsiyon verimi 1 ml/dak akış hızı, 0,5 V potansiyel ve 50 ppm bor konsantrasyonunda %94,1 olarak belirlenmiştir. Ayrıca, optimum çalışma koşullarda elektrotlar SEM analiz yöntemiyle karakterize edilmiş ve bor adsorpsiyonu sonrasında yüzeyin homojen olarak kaplandığı ve adsorpsiyon öncesi var olan gözeneklerin kapandığı görülmüştür.

Yapılan çalışma sonucunda Box-Behnken deney tasarım yönteminin sürekli akış koşullarında bor iyonlarının giderimi proses parametrelerinin optimizasyonunda uygun bir yöntem olduğu belirlenmiştir. Bu çalışmadan elde edilen sonuçların, su kaynakların korunması, sürdürülebilir şekilde bu kaynaklarının yönetiminin sağlanması, yenilikçi, ekonomik ve çevre dostu alternatif proseslerin geliştirilmesine katkı sağlayacağı düşünülmektedir.

TEŞEKKÜR

Bu çalışma, Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) tarafından 123M952 numaralı proje ile desteklenmiştir. Projeye verdiği destekten ötürü TÜBİTAK'a teşekkürlerimi sunarım.

KAYNAKLAR

- [1] Ngwabebhoh, F.A., Mammadli, N., Yıldız, U. (2019). Bioinspired modified nanocellulose adsorbent for enhanced boron recovery from aqueous media: Optimization, kinetics, thermodynamics and reusability study. *J. Environ. Chem. Eng.*, 7, 103281.
- [2] Xu, Y., Jiang, J.Q. (2008). Technologies for boron removal. *Ind. Eng. Chem. Res.*, 47, 16-2410.
- [3] Yetgin, A.G., DüNDAR, O.A., Çakmakçı, E. (2023). Removal of boron from aqueous solution by modified cellulose. *Biomass Convers. Biorefin.*, 13, 13081–13090.
- [4] Wolska, J., Bryjak, M. (2013). Methods for boron removal from aqueous solutions — A review, *Desalination.*, 310, 18-24.
- [5] Kim, H., Kim, S., Kim, C. (2024). Enhanced boron removal without pre-pH adjustment via redox-mediated electrodialysis assisted by ion-exchange resins. *J. Environ. Chem. Eng.*, 12(4), 113159.
- [6] Hilal, N., Kim, G.J., Somerfield, C. (2011). Boron removal from saline water: A comprehensive review. *Desalination.*, 273, 23–35.
- [7] Guan, Z., Lv, J., Bai, P., Guo, X. (2016). Boron removal from aqueous solutions by adsorption — A review. *Desalination.*, 383, 29–37.
- [8] Ee, L.Y., Chia, S.Y.R., Xue, K., Chin, S.Y., Cho, C.A.H., Tan, X.Y., Li, S.F.Y. (2023). Hyperbranched nanocellulose enabling rapid boron removal from aqueous environment. *Chem. Eng. J.*, 454, 140218.
- [9] Akdağ, S., Keyikoglu, R., Karagunduz, A., Keskinler, B., Khataee, A., Yoon, Y. (2023). Recent advances in boron species removal and recovery using layered double hydroxides. *Appl. Clay Sci.*, 233, 106814.
- [10] Çelebi, H., Şimşek, İ., Bahadır, T., Tulun, Ş., (2023). Use of banana peel for the removal of boron from aqueous solutions in the batch adsorption system. *Int. J. Environ. Sci. Technol.*, 20, 161-176.

- [11] Lin, J.Y., Mahasti, N.N., Huang, Y.H. (2021). Recent advances in adsorption and coagulation for boron removal from wastewater: A comprehensive review. *J. Hazard. Mater.* 407, 124401.
- [12] Xia, N.N., Zhang, H.Y., Hu, Z.H., Kong, F., He, F. (2021). A functionalized bio-based material with abundant mesopores and catechol groups for efficient removal of boron. *Chemosphere.*, 263, 128202.
- [13] Qiu, X., Sasaki, K., Hirajima, T., Ideta, K., Miyawaki, J. (2014). One-step synthesis of layered double hydroxide-intercalated gluconate for removal of borate. *Sep Purif Technol.*, 123, 114-123.
- [14] Mutlu-Salmanli, O., Koyuncu, I. (2022). Boron removal and recovery from water and wastewater. *Rev. Environ. Sci. Biotechnol.*, 21, 635-664.
- [15] Lin, J.Y., Mahasti, N.N.N., Huan, Y.H. (2021). Fluidized-bed crystallization of barium perborate for continuous boron removal from concentrated solution: Supersaturation as a master variable. *Sep Purif Technol.*, 278, 119588.
- [16] Biçak, N., Bulutçu, N., Şenkal, B.F., Gazi, M. (2001). Modification of crosslinked glycidyl methacrylate-based polymers for boron-specific column extraction. *React. Funct. Polym.*, 47, 175-184.
- [17] Wang, S., Bing, S., Zhang, H., Zhou, Y., Zhang, L., Gao, C. (2022). Surface engineering design of polyamide membranes for enhanced boron removal in seawater desalination. *J. Membr. Sci.*, 651, 120425.
- [18] Ghiasi, S., Mohammadi, T., Tofighy, M.A. (2022). Hybrid nanofiltration thin film hollow fiber membranes with adsorptive supports containing bentonite and LDH nanoclays for boron removal. *J. Membr. Sci.* 655, 120576.
- [19] Tang, Y.P., Luo, L., Thong, Z., Chung, T.S. (2017). Recent advances in membrane materials and technologies for boron removal. *J. Memb. Sci.* 541, 434-446.
- [20] Guesmi, F., Louati, I., Hannachi, C., Hamrouni, B. (2020). Optimization of boron removal from water by electrodialysis using response surface methodology. *Water Sci. Technol.* 81, 293-300.
- [21] Khajeh, M. (2011). Optimization of process variables for essential oil components from *Satureja hortensis* by supercritical fluid extraction using Box-Behnken experimental design. *J. Supercrit. Fluids*, 55(3), 944-948.
- [22] Duranoğlu, D., Yılmaz, Y. (2024). Optimization of titanium dioxide production process for photocatalytic degradation of 2,4 dichlorophenoxyacetic acid. *J. Fac. Eng. Archit. Gazi Univ.*, 39(1), 233-242.
- [23] Alrefaee, S.H., Aljohani, M., Alkhamis, K., Shaaban, F., El-Desouky, M.G., El-Bindary, A.A., El-Bindary, M.A. (2023). Adsorption and effective removal of organophosphorus pesticides from aqueous solution via novel metal-organic framework: Adsorption isotherms, kinetics, and optimization via Box-Behnken design. *J. Mol. Liq.*, 384, 122206
- [24] Al-dhawi, B.N.S., Kutty, S.R.M., Hayder, G., Elnaim, B.M.E., Mnzool, M., Noor, A., Saeed, A.A.H., Al-Mahbashi, N.M.Y., Al-Nini, A., Jagaba, A.H. (2023). Adsorptive Removal of Boron by DIAION™ CRB05: Characterization, Kinetics, Isotherm, and Optimization by Response Surface Methodology. *Processes.*, 11(2), 453.
- [25] Kluczka, J., Pudło, W., Krukiewicz, K. (2019). Boron adsorption removal by commercial and modified activated carbons, *Chem. Eng. Res. Des.*, 147, 30-42.
- [26] Halim, A.A., Roslan, N.A., Yaacob, N.S., Latif, M.T. (2013). Boron removal from aqueous solution using curcumin-impregnated activated carbon, *Sains Malaysiana*, 42(9), 1293-1300.
- [27] Irawan, C., Kuo, Y.L., Liu, J.C. (2011). Treatment of boron-containing optoelectronic wastewater by precipitation process, *Desalination*, 280, 146-151.
- [28] Polat, S. (2024). Sürekli akış koşullarında sulu çözeltilerden bor giderimi için elektrokimyasal ayırma prosesi tasarlanması ve işletilmesi, *Journal of Boron*, 9(4), 135-142.
- [29] Gurten-Inal, I.I., Aktas, Z. (2020). Enhancing the performance of activated carbon based scalable supercapacitors by heat treatment, *Appl. Surf. Sci.*, 514, 145895.