Research Article GU J Sci 37(4): 2085-2094 (2024) DOI: 10.35378/gujs.1393264

Gazi University

Journal of Science

http://dergipark.gov.tr/gujs

# The Parameter Estimation of the COVID-19 Death Based on the Gumbel Distribution Through the Multi-objective Programming: Turkey Case

Ecem DEMIR YURTSEVEN[1], Emre KOCAK[2,*], H. Hasan ORKCU[2]

[1]*Rector's Office, Sivas Cumhuriyet University, Sivas, Türkiye*
[2]*Department of Statistics, Faculty of Science, Gazi University, Ankara, Türkiye*

## Highlights

- The study aims to comprehend the impacts of the virus and how it spreads.
- To estimate the daily number of COVID-19 related deaths, Gumbel distribution is used.
- The findings can be viewed as a valuable resource for analyzing the effects of any future pandemics.

**Article Info**

*Received: 20 Nov 2023*
*Accepted: 20 Mar 2024*

**Keywords**

*Gumbel distribution*
*Multi-objective optimization*
*NSGA-II*
*COVID-19-related deaths*

**Abstract**

Nearly all nations, including Turkey, were impacted by the 2019 new coronavirus (COVID-19) infections reported by Wuhan, China, as the disease's first official case. Turkey is one of the most impacted nations in the globe due to the high number of infected patients. To comprehend the pattern of the virus's propagation and its impacts, it is crucial to examine the pandemic statistics in Turkey. The Gumbel distribution is utilized when describing the maximum or minimum of several samples with different distributions. Therefore, we used the Gumbel distribution to estimate the daily number of COVID-19-related deaths. This study proposes a multi-objective programming methodology for Gumbel distribution parameter estimation based on the RMSE, $R^2$, and Theil coefficient methods. A comprehensive Monte-Carlo simulation research is performed to examine the effectiveness of single-objective RMSE, $R^2$, Theil's coefficient and multi-objective RMSE-$R^2$, RMSE-Theil, $R^2$-Theil, RMSE-$R^2$-Theil programming estimation methods. When the simulation results were analyzed, the case formed by the RMSE-$R^2$-Theil estimator has the best Def value across all cases. The application of the real dataset containing COVID-19 death data is examined, and it can be seen that Theil, RMSE-Theil, and $R^2$-Theil were better estimators for winter data. At the same time, RMSE was a better estimator for autumn and autumn-winter data.

## 1. INTRODUCTION

A disaster faced by humanity is undoubtedly epidemic disease, and throughout history, it has deeply affected states, societies, and people, struck commercial activities, and paralyzed social life [1]. People's perceptions of diseases, reactions to them, and ability to adapt to them are all related. [2]. For this reason, people's health behaviors in the face of an epidemic threat are essential in minimizing the epidemic's geographical prevalence and the speed of its spread and reducing possible loss of life. In this context, the lives of civilizations and the global economy have been profoundly impacted by the new coronavirus epidemic (COVID-19 or 2019-nCOV), which began in the Chinese city of Wuhan and has spread quickly throughout the world. Since the World Health Organization declared the COVID-19 pandemic, 514 million cases have been detected worldwide, and 6.25 million deaths have occurred. In Turkey, the COVID-19 virus has been detected in 15 million people and 99 thousand deaths have been reported [3].

Upon reviewing the publications that provide data on the quantity of COVID-19 cases and fatalities, Chen analyzed the global death rate and case-fatality rates of COVID-19 using nonlinear regression [4]. An attempt has been made to confirm the MSL-COVID-19 score generated to predict COVID-19 mortality in Mexicans, and a modified version of the equation is suggested to assess the severity of COVID-19 in a triage situation [5]. During the first wave of the pandemic in 23 countries, the Gompertz model was used to characterize the growth dynamics of COVID-19 cases, and it was compared with the straightforward

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

logistic regression model [6]. The number of hospitalized cases was utilized to estimate the number of discharged and deceased cases using the MLP-MC hybrid model [7]. Time models, conditional variance, and asymmetric effects were used to model and forecast new COVID-19 cases [8]. The epidemic model was created to comprehend the disease's spreading tendency, incorporating effective TTT (tracking, testing, and treatment) control techniques [9]. Modeling pandemics is important in understanding the spreading behavior of a virus and its effects on humanity.

First introduced by E. J. Gumbel in 1941, the Gumbel distribution is applied to explain the maximum or minimum of several samples of different distributions [10]. In real life, the Gumbel distribution is used for numerous purposes, including predicting annual sea level rise and the flow of rivers, floods, strong winds, and earthquake magnitudes [11-14]. This distribution has different applications in modeling extreme values, so it is essential to estimate its parameters correctly. Various estimation methods, such as the method of moments, the method of maximum likelihood, the method of least squares, and the method of weighted least squares, have been used to estimate the Gumbel distribution parameters [14-18]. The Gumbel distribution function was utilized to estimate the daily and overall number of COVID-19-related deaths [19,20]. Since the root means square error (RMSE) and Theil's inequality coefficient includes the least-squares estimator function, coefficient determination ($R^2$) contains the likelihood function for simple linear regression, and the parameter estimation functions are RMSE, Theil's inequality, and $R^2$ are used.

In the second and third section of the study, Gumbel distribution and estimation methods and the NSGA-II method are respectively given. In the next section, a detailed Monte Carlo (MC) simulation research is performed to examine the effectiveness of single-objective RMSE, $R^2$, Theil's coefficient and multi-objective RMSE-$R^2$, RMSE-Theil, $R^2$-Theil, RMSE-$R^2$-Theil programming estimation method, and the findings are presented. The implementation of the suggested method has been evaluated using COVID-19 death data in Turkey in the next section. In the last section, the study is summarized, and the results are analyzed.

## 2. ESTIMATING PARAMETERS FOR GUMBEL DISTRIBUTION USING THE NSGA-II

The cumulative distribution function (CDF) and the probability density function (PDF) of the two-parameter Gumbel distribution with the location parameter μ and the scale parameter σ are given in Equations (1) and (2), respectively

$$F(x) = exp\left(-exp\left(-\frac{(x-\mu)}{\sigma}\right)\right) \tag{1}$$

$$f(x) = \frac{1}{\sigma} exp\left(-\frac{(x-\mu)}{\sigma}\right) exp\left(-exp\left(-\frac{(x-\mu)}{\sigma}\right)\right), -\infty < x < \infty, \mu \in R, \sigma \in R^{+} . \tag{2}$$

In this section, where the parameters of the two-parameter Gumbel distribution are given, the concepts of RMSE, $R^2$, and Theil's inequality coefficient estimators and the proposed multi-objective approach are presented, let be a random sample of size $n$ taken from the PDF in $x_1, x_2, \dots, x_n$.

### 2.1. Root means square error (RMSE)

The standard deviation of the estimated errors is measured by the RMSE, a quadratic metric that assesses the size of the error in a model between the expected and true values. Considering $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ as the order statistics of $X_1, X_2, \dots, X_n$ and $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ as observed ordered observations, the values of CDF in Equation (1) are estimated using the mean rank provided in Equation (3)

$$\hat{F}\left(x_{(i)}\right) = \hat{y}_{\mathrm{i}} = \frac{i}{n+1}, \qquad i = 1,2, \dots, n . \tag{3}$$

Here, *i* denotes the *i*th smallest value of $x_{(1)}, x_{(2)}, \dots, x_{(n)}$. The RMSE minimizing the function in Equation (4)

$$\psi(\mu,\sigma) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}\,. \tag{4}$$

## 2.2. Coefficient of Determination ($R^2$)

The coefficient of determination is the best measure of the linear model's goodness of fit. The coefficient in question indicates how much of the change in the dependent variable is explained by the independent variable(s). Equation (5) provides the most comprehensive definition of the coefficient of determination

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - F(x_i)\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y}))^2}\,. \tag{5}$$

## 2.3. Theil's Inequality Coefficient (TIC)

The accuracy of a set of predictions produced by certain models was assessed using Theil's inequality coefficient. One of the first applications was by Hee in 1966 [21]. Thiel's inequality coefficient, sometimes called as Thiel's U, indicates how closely a model's predicted values match the corresponding observed values. Theil's summary measure of prediction accuracy is proposed in Equation (6)

$$TIC = \frac{\sqrt{\frac{1}{N}\left(\sum_{i=1}^{N}\hat{y}_i - y_i\right)^2}}{\sqrt{\frac{1}{N}\sum_{i=}^{N}\hat{y}_i^2} + \sqrt{\frac{1}{N}\sum_{i=}^{N}y_i^2}}\,. \tag{6}$$

## 2.4. Non-dominated Sorting Genetic Algorithm II (NSGA-II)

In addition to being a multi-objective metaheuristic approach, Genetic Algorithm-based NSGA-II is among the best techniques for identifying the Pareto solution set. The fast non-dominated sorting and the crowding distance are the two main advantages of the NSGA-II method over other multi-objective genetic algorithms [22]. The search begins with a set of solutions using the population-based search method NSGA-II, each representing a possible solution for the issue. Better solutions are sought from the current set of solutions. A new population is produced through crossing and mutation operators using the best members of the existing population. For a set number of iterations, it keeps creating a population.

## 3. PROPOSED MULTI-OBJECTIVE OPTIMIZATION APPROACH

The functions of RMSE, $R^2$, and Theil were used to build the suggested multi-objective optimization methodology. It is desirable to make the function as small as possible in the RMSE and Theil but as large as possible in the $R^2$. Therefore, the additive inverse of the $R^2$ function is taken. The generated models for RMSE-$R^2$, RMSE-Theil, Theil-$R^2$, and RMSE-$R^2$-Theil are specified in Equations (7) - (10), respectively.

Rather than a single optimal solution, sometimes referred to as the ideal solution, Pareto optimum solutions are alternate solution sets that answer the multi-objective optimization problem. In order to obtain Pareto optimal solutions, multi-objective metaheuristic approaches are practical because they generate many solutions, do not require derivative computations, offer an excellent approach to Pareto optimal solutions, and are easily applicable to optimization problems

$$Minimize \begin{cases} \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(exp - \left(exp - \left(\frac{x_i - \mu}{\sigma}\right)\right) - F\left(x_{(i)}\right)\right)^2} \\ \\ \frac{\sum_{i=1}^{n}\left(exp - \left(exp - \left(\frac{x_i - \mu}{\sigma}\right)\right) - F\left(x_{(i)}\right)\right)^2}{\left(\sum_{i=1}^{n} exp - \left(exp - \left(\frac{x_i - \mu}{\sigma}\right)\right) - \frac{1}{n}\sum_{i=1}^{n} exp - \left(exp - \left(\frac{x_i - \mu}{\sigma}\right)\right)\right)^2} \end{cases} \tag{7}$$

$$Minimize\begin{cases}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)-F\left(x_{(i)}\right)\right)^2} \\ \\ \dfrac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)-F\left(x_{(i)}\right)\right)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)^2\right)}+\sqrt{\frac{1}{n}\sum_{i=1}^{n}F\left(x_{(i)}\right)^2}}\end{cases} \tag{8}$$

$$Minimize\begin{cases}\dfrac{\sum_{i=1}^{n}\left(exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)-F\left(x_{(i)}\right)\right)^2}{\left(\sum_{i=1}^{n}exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)-\frac{1}{n}\sum_{i=1}^{n}exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)\right)^2} \\ \\ \dfrac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)-F\left(x_{(i)}\right)\right)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)^2\right)}+\sqrt{\frac{1}{n}\sum_{i=1}^{n}F\left(x_{(i)}\right)^2}}\end{cases} \tag{9}$$

$$Minimize\begin{cases}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)-F\left(x_{(i)}\right)\right)^2} \\ \\ \dfrac{\sum_{i=1}^{n}\left(exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)-F\left(x_{(i)}\right)\right)^2}{\left(\sum_{i=1}^{n}exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)-\frac{1}{n}\sum_{i=1}^{n}exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)\right)^2}\ . \\ \\ \dfrac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)-F\left(x_{(i)}\right)\right)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(exp-\left(exp-\left(\frac{x_i-\mu}{\sigma}\right)\right)^2\right)}+\sqrt{\frac{1}{n}\sum_{i=1}^{n}F\left(x_{(i)}\right)^2}}\end{cases} \tag{10}$$

## 4. SIMULATION STUDY

A detailed MC simulation research is performed to investigate the performance of the proposed single-purpose RMSE, $R^2$, Theil's coefficient and multi-purpose RMSE-$R^2$, RMSE-Theil, $R^2$-Theil, RMSE-$R^2$-Theil programming estimation methods in this section. The performance of various scenarios is compared using the deficiency criterion (Def), which is used to evaluate the efficiency of the parameter estimate techniques [23]. It is defined as given in Equation (11). Additionally, Equations (12) and (13) provide the mean squared error (MSE) values for the parameters that were used to calculate the Def criterion

$$Def(\hat{\mu},\hat{\sigma}) = MSE(\hat{\mu}) + MSE(\hat{\sigma}) \tag{11}$$

$$MSE(\hat{\mu}) = Var(\hat{\mu}) + Bias^2(\hat{\mu}) \tag{12}$$

$$MSE(\hat{\sigma}) = Var(\hat{\sigma}) + Bias^2(\hat{\sigma})\ . \tag{13}$$

Here, $\hat{\mu}$ and $\hat{\sigma}$ are estimators of parameters $\mu$ and $\sigma$ respectively.

Selecting the point that corresponds to the best $E^2$ value of the prediction points provides the optimal solution among the Pareto points in the parameter space for RMSE-$R^2$, RMSE-Theil, $R^2$-Theil, and RMSE-

$R^2$-Theil. The sum of the differences between the estimated and real parameter values gives the $E^2$ value, which is determined in Equation (14)

$$E^2 = \sum_{i=1}^{n}(\gamma - \hat{\gamma})^2 \ . \quad (14)$$

The real parameter values for the two-parameter Gumbel distribution are to be $(\mu, \sigma) = (3{,}4)$. The mean, MSE, and Def values of the parameter estimators are calculated using 100000/n Monte Carlo runs for sample sizes of 10, 50, 100, and 1000. RMSE, $R^2$, Theil, RMSE-$R^2$, RMSE-Theil, $R^2$-Theil, and RMSE-$R^2$-Theil estimates of the parameters for the two-parameter Gumbel distribution are calculated using the NSGA-II algorithm. Crossover Fraction = 0.8 and Pareto Fraction = 0.35 are the parameters considered for the NSGA-II algorithm. Population Size (Pop) is also considered 50, 100, and 250, respectively, while the Search Space for $\mu$ and $\sigma$ is chosen as (0, inf). The simulated mean and Def values for $(\mu, \sigma) = (3{,}4)$ are given in Table 1.

The simulation results demonstrate that the case created by the RMSE-$R^2$-Theil estimator has the best Def value across all cases. Compared to its parameter estimation, this case does not yield the best results in any scenario. When the simulation results were analyzed in terms of parameter estimations, it was seen that the $R^2$ state made the best parameter estimation for the location and scale parameters in all parameter values. This situation is because achieving objectives in cases with more than one objective function is difficult to compare with cases with an objective function, and the spread of predicted values is less in multi-objective function cases when estimating parameters. However, the parameter estimate values are related under various circumstances.

It is seen in Table 1 that as the number of samples increases, there is a decrease in Def values in all cases. On the other hand, it shows that the growth of the population number does not affect the Def values as much as the number of samples. Although RMSE-$R^2$-Theil is the estimator that gives the best def values for all cases, this is not the same for the population size parameter. When the sample size is 10, 50, 100, and 1000, the population sizes that give the best Def results are 100, 250, 50, and 50, respectively. Using these results, the population size can be small when working with large sample sizes. Thus, an advantage can be achieved in terms of algorithm time.

## 5. APPLICATION OF COVID-19 DEATH DATA

This section examines, using the NSGA-II algorithm, the application of parameter estimates approaches to a real dataset that contains COVID-19 death data. Consisting of 181 observations, the number of deaths from COVID-19 in Turkey between September 1, 2021, and February 28, 2022, was used as a dataset, and the data were analyzed in three datasets. In the first dataset, the dataset for the autumn-winter period consists of 181 observations covering all the data, and the second reflects the autumn period of 91 observations between September 1 and November 30, 2021, and the winter period of 90 observations between December 1, 2021-February 28, 2022.

A goodness-of-fit test using Kolmogorov-Smirnov (KS) was used to demonstrate how well the data fit the Gumbel distribution. The KS test statistics and p values are KS=0.0559 and p=0.9233 for the autumn period, KS=0.0908 and p=0.4238 for the winter period and KS=0.0647 and p=0.4168 for the autumn-winter period, respectively.

When the results were examined, it was seen that the dataset was suitable for the Gumbel distribution ($p<0.05$). When the parameter estimation values of logL and AIC of the COVID-19 death dataset are examined in Table 2, it can be seen from logL and AIC values that RMSE estimates are better than in the other cases for autumn data. Theil, RMSE-Theil, and $R^2$-Theil estimates are better for the winter datasets, and RMSE cases ensure the best estimates for the autumn-winter data. Additionally, the Pareto optimal solution sets are given in Figure 1.

***Table 1.*** *Simulation results for estimating of the parameters* $(\mu, \sigma) = (3,4)$

| Pop. Size | Method | n=10 | | | n=50 | | |
|---|---|---|---|---|---|---|---|
| | | Location | Scale | Def | Location | Scale | Def |
| 50 | RMSE | 2.98155 | 4.40919 | 4.12578 | 2.95997 | 4.06743 | 0.72047 |
| | $R^2$ | 3.07926 | 4.08480 | 3.91429 | 2.97788 | 4.00810 | 0.71184 |
| | Theil | 2.95816 | 4.36057 | 4.06099 | 2.95435 | 4.05842 | 0.71832 |
| | RMSE- $R^2$ | 3.02888 | 4.23969 | 3.57662 | 2.96806 | 4.04329 | 0.68822 |
| | RMSE-Theil. | 2.90480 | 4.56541 | 4.07379 | 2.91971 | 4.18334 | 0.85841 |
| | $R^2$-Theil. | 3.01913 | 4.21960 | 3.59038 | 2.96553 | 4.03837 | 0.69270 |
| | RMSE- $R^2$-Theil. | 3.02021 | 4.24371 | 3.55770 | 2.96592 | 4.04249 | 0.68636 |
| 100 | RMSE | 2.95718 | 4.41868 | 4.05257 | 2.99992 | 4.10002 | 0.74665 |
| | $R^2$ | 3.05659 | 4.09421 | 3.81098 | 3.01869 | 4.03757 | 0.73533 |
| | Theil | 2.93323 | 4.37055 | 3.98688 | 2.99416 | 4.09023 | 0.74364 |
| | RMSE- $R^2$ | 3.01198 | 4.22724 | 3.52775 | 3.01003 | 4.06571 | 0.70997 |
| | RMSE-Theil. | 2.92943 | 4.42663 | 3.89131 | 2.99170 | 4.11012 | 0.74139 |
| | $R^2$-Theil. | 3.00256 | 4.21017 | 3.54615 | 3.00749 | 4.06135 | 0.71248 |
| | RMSE- $R^2$-Theil. | 3.00449 | 4.23100 | 3.50397 | 3.00830 | 4.06608 | 0.70823 |
| 250 | RMSE | 2.95953 | 4.43113 | 4.04294 | 3.02035 | 4.06750 | 0.68841 |
| | $R^2$ | 3.05742 | 4.10781 | 3.81600 | 3.03923 | 4.00526 | 0.68135 |
| | Theil | 2.93606 | 4.38293 | 3.98026 | 3.01459 | 4.05782 | 0.68583 |
| | RMSE- $R^2$ | 3.01369 | 4.23945 | 3.53338 | 3.03018 | 4.03432 | 0.65595 |
| | RMSE-Theil. | 2.94333 | 4.40265 | 3.93539 | 3.01701 | 4.06242 | 0.68098 |
| | $R^2$-Theil. | 3.00479 | 4.22291 | 3.55338 | 3.02735 | 4.03038 | 0.65803 |
| | RMSE- $R^2$-Theil. | 3.00715 | 4.24166 | 3.51801 | 3.02827 | 4.03496 | 0.65412 |
| Pop. Size | Method | n=100 | | | n=1000 | | |
| | | Location | Scale | Def | Location | Scale | Def |
| 50 | RMSE | 2.99275 | 4.03382 | 0.32519 | 3.00490 | 4.00598 | 0.03111 |
| | $R^2$ | 3.00199 | 4.00300 | 0.32345 | 3.00582 | 4.00295 | 0.03112 |
| | Theil | 2.98988 | 4.02888 | 0.32477 | 3.00461 | 4.00550 | 0.03110 |
| | RMSE- $R^2$ | 2.99534 | 4.02606 | 0.32009 | 3.00250 | 4.01068 | 0.03208 |
| | RMSE-Theil. | 2.95417 | 4.15141 | 0.46849 | 2.99545 | 4.02163 | 0.03580 |
| | $R^2$-Theil. | 2.99474 | 4.02100 | 0.31753 | 3.00223 | 4.01979 | 0.04604 |
| | RMSE- $R^2$-Theil. | 2.99501 | 4.02387 | 0.31704 | 3.00506 | 4.00542 | 0.03035 |
| 100 | RMSE | 2.98978 | 4.04095 | 0.35403 | 2.98431 | 4.00201 | 0.03688 |
| | $R^2$ | 2.99920 | 4.00972 | 0.35153 | 2.98515 | 3.99920 | 0.03688 |
| | Theil | 2.98682 | 4.03604 | 0.35342 | 2.98403 | 4.00156 | 0.03689 |
| | RMSE- $R^2$ | 2.99464 | 4.02457 | 0.34228 | 2.98474 | 4.00063 | 0.03655 |
| | RMSE-Theil. | 2.98519 | 4.04966 | 0.35384 | 2.98203 | 4.01003 | 0.03943 |
| | $R^2$-Theil. | 2.99328 | 4.02226 | 0.34310 | 2.98462 | 4.00035 | 0.03659 |
| | RMSE- $R^2$-Theil. | 2.99370 | 4.02478 | 0.34162 | 2.98465 | 4.00070 | 0.03653 |
| 250 | RMSE | 2.99363 | 4.03571 | 0.35691 | 2.98992 | 4.00088 | 0.03088 |
| | $R^2$ | 3.00278 | 4.00521 | 0.35496 | 2.99087 | 3.99772 | 0.03083 |
| | Theil | 2.99076 | 4.03089 | 0.35638 | 2.98961 | 4.00037 | 0.03089 |
| | RMSE- $R^2$ | 2.99822 | 4.02026 | 0.34584 | 2.99044 | 3.99917 | 0.03056 |
| | RMSE-Theil. | 2.99213 | 4.03325 | 0.35451 | 2.98975 | 4.00061 | 0.03082 |
| | $R^2$-Theil. | 2.99687 | 4.01783 | 0.34688 | 2.99031 | 3.99891 | 0.03060 |
| | RMSE- $R^2$-Theil. | 2.99737 | 4.02046 | 0.34525 | 2.99033 | 3.99923 | 0.03054 |

***Table 2.*** *Parameter estimate values for the dataset containing the number of deaths from COVID-19*

| Period | Method | Location | Scale | RMSE | $R^2$ | Theil | logL | AIC |
|---|---|---|---|---|---|---|---|---|
| Autumn | RMSE | 0.00174 | 0.00036 | 0.03723 | - | - | -87.92316 | 179.84631 |
| | $R^2$ | 0.00174 | 0.00035 | - | 0.98361 | - | -99.36461 | 202.72922 |
| | Theil | 0.00173 | 0.00035 | - | - | 0.03245 | -98.58736 | 201.17472 |
| | RMSE-$R^2$ | 0.00174 | 0.00035 | 0.03751 | 0.98361 | - | -99.36461 | 202.72922 |
| | RMSE-Theil | 0.00174 | 0.00036 | 0.03724 | - | 0.03213 | -89.80119 | 183.60238 |
| | $R^2$-Theil | 0.00174 | 0.00036 | - | 0.98340 | 0.03213 | -89.80119 | 183.60238 |
| | RMSE-$R^2$-Theil | 0.00174 | 0.00035 | 0.03740 | 0.98359 | 0.03224 | -97.15839 | 198.31679 |
| Winter | RMSE | 0.00209 | 0.00022 | 0.01714 | - | - | -456.90823 | 917.81647 |
| | $R^2$ | 0.00209 | 0.00021 | - | 0.99643 | - | -460.41864 | 924.83729 |
| | Theil | 0.00209 | 0.00022 | - | - | 0.01485 | -456.72442 | 917.44883 |
| | RMSE-$R^2$ | 0.00209 | 0.00021 | 0.01716 | 0.99643 | - | -460.41864 | 924.83729 |
| | RMSE-Theil | 0.00209 | 0.00022 | 0.01714 | - | 0.01485 | -456.72442 | 917.44883 |
| | $R^2$-Theil | 0.00209 | 0.00022 | - | 0.99642 | 0.01485 | -456.72442 | 917.44883 |
| | RMSE-$R^2$-Theil | 0.00209 | 0.00021 | 0.01716 | 0.99643 | 0.01487 | -460.41864 | 924.83729 |
| Autumn-Winter | RMSE | 0.00192 | 0.00035 | 0.02195 | - | - | -265.10074 | 534.20148 |
| | $R^2$ | 0.00193 | 0.00033 | - | 0.99545 | - | -325.71813 | 655.43627 |
| | Theil | 0.00193 | 0.00033 | - | - | 0.01697 | -320.19816 | 644.39632 |
| | RMSE-$R^2$ | 0.00193 | 0.00033 | 0.01955 | 0.99545 | - | -325.71813 | 655.43627 |
| | RMSE-Theil | 0.00193 | 0.00033 | 0.01952 | - | 0.01697 | -320.19816 | 644.39632 |
| | $R^2$-Theil | 0.00193 | 0.00033 | - | 0.99544 | 0.01697 | -320.19816 | 644.39632 |
| | RMSE-$R^2$-Theil | 0.00193 | 0.00033 | 0.01954 | 0.99545 | 0.01699 | -323.06819 | 650.13638 |

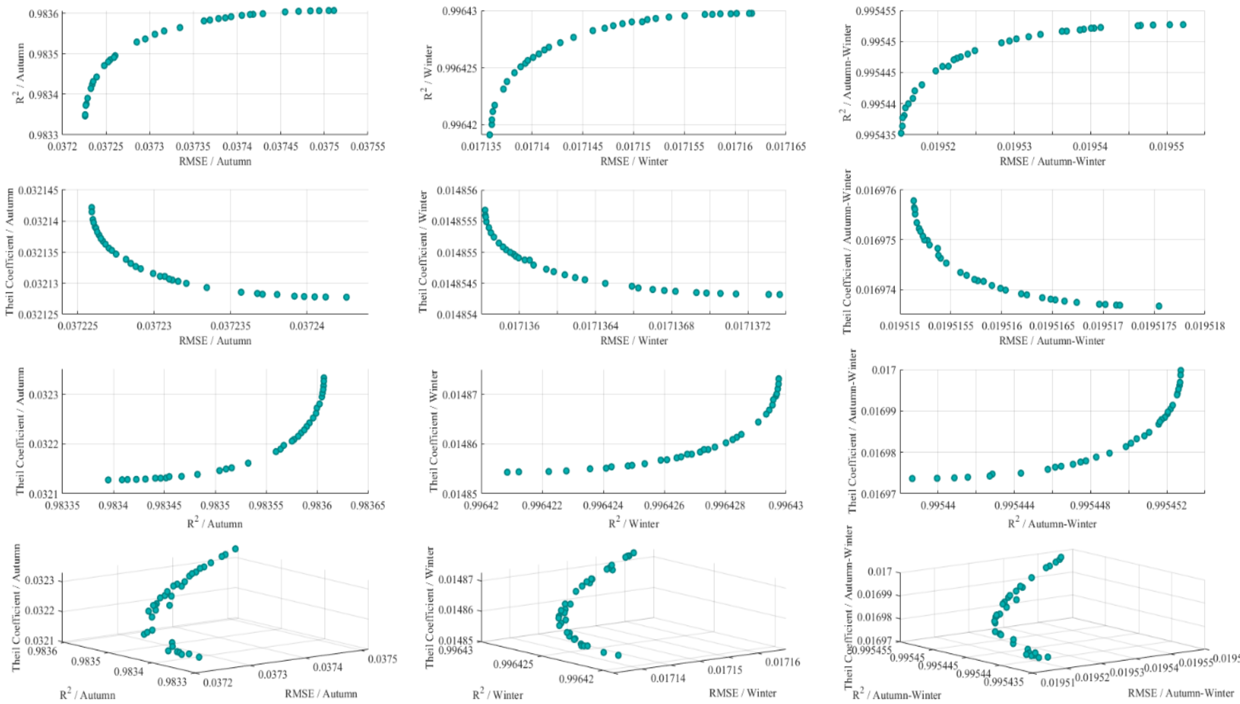


***Figure 1.*** *Pareto Points in the Parameter Space*

## 6. CONCLUSION

The Gumbel distribution is applied to explain the maximum or minimum of a set of situations of different distributions in probability theory and statistics. A multi-objective programming approach has been presented for the parameter estimation of the two-parameter Gumbel distribution, wherein methods are jointly assessed during the estimate process for the RMSE, $R^2$, and Theil performance criteria in this study.

A detailed MC simulation research is carried out to evaluate the effectiveness of this suggested strategy. The simulation results demonstrate that the case created by the RMSE-$R^2$-Theil estimator has the best Def value across all cases. The $R^2$ example produces the best parameter estimation for all parameters when the simulation results have been investigated in terms of parameter estimations. A case-involving Theil's coefficient method in parameter estimation is weaker than the others, as it does not have the best estimator case for any particular case. The findings demonstrate how well this multi-objective programming method estimates the Gumbel distribution's parameters concerning deficient criteria.

The number of deaths from COVID-19 in Turkey between September 1, 2021, and February 28, 2022, shows the Gumbel distribution. Accurate estimation of Gumbel distribution parameters gives us information about the period when more deaths will occur. In line with these possibilities, the state and institutions can take the necessary measures to reduce the number of deaths in advance. In the subsequent investigation, the parameters of other statistical distributions can be estimated using the recommended methodology.

## CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

## REFERENCES

[1] Ekiz, T., Ilıman, E., Dönmez, E., "Comparison of health anxiety level and control perception of COVID-19", International Journal of Health Management and Strategies Research, 6(1): 139-154, (2020).

[2] Hekler, E. B., Lambert, J., Leventhal, E., Levethal, H., Jahn, E, Contrada, R. J., "Commonsense illness beliefs, adherence behaviors and hypertension control among African Americans", Journal of Behavioral Medicine, 31: 391-400, (2008).

[3] https://covid19.who.int/. Access date: 05.04.2022

[4] Chen, J. M., "Novel statistics predict the COVID-19 pandemic could terminate in 2022", Journal of Medical Virology, 94(6): 2845-2848, (2022).

[5] Bello-Chavolla, O. Y., Antonio-Villa, N. E., Ortiz-Brizuela, E., Vargas-Vázquez, A., González-Lara, M. F., de Leon, A. P., Sifuentes-Osornio, J., Aguilar-Salinas, C. A., "Validation and repurposing of the MSL-COVID-19 score for prediction of severe COVID-19 using simple clinical predictors in a triage setting: The Nutri-CoV score", PLoS One, 15(12), (2020).

[6] Pelinovsky, E., Kokoulina, M., Epifanova, A., Kurkin, A., Kurkina, O., Tang, M., Macau, E., Kirillin, M., "Gompertz model in COVID-19 spreading simulation", Chaos, Solitons and Fractals, 154: 111699, (2022).

[7] Haghighat, F., "Predicting the trend of indicators related to Covid-19 using the combined MLP-MC model", Chaos, Solitons and Fractals, 152: 111399, (2021).

[8] Ekinci, A., "Modelling and forecasting of the growth rate of new COVID-19 cases in top nine affected countries: Considering conditional variance and asymmetric effect", Chaos, Solitons and Fractals, 151: 0111227, (2021).

[9] Mishra, B. K., Keshri, A. K., Saini, D. K., Ayesha, S., Mishra, B. K., Rao, Y. S., "Mathematical model, forecast and analysis on the spread of COVID-19", Chaos, Solitons and Fractals, 147: 110995, (2021).

[10] Gumbel, E. J., "The return period of flood flows", The Annals of Mathematical Statistics, 12(2): 163-190, (1941).

[11] Kang, D., Ko, K., Huh, J., "Determination of extreme wind values using the Gumbel distribution", Energy, 86: 51-58, (2015).

[12] Niemann, H. J., Diburg, S., "Statistics of extreme climatic actions based on the Gumbel probability distributions with an upper limit", Computers and Structures, 126: 193-198, (2013).

[13] García, Bustos, S. L., Navarrete, S., Chancay, A., Mendoza, M., Pincay, M., Teran, M., "Zoning of Ecuador according to maximum magnitudes of earthquakes and their frequency of occurrence using statistical models estimated by maximum likelihood", Gazi University Journal of Science, 34(3): 916-935, (2021).

[14] Aydın, D., Şenoğlu, B., "Monte Carlo comparison of the parameter estimation methods for the two-parameter Gumbel distribution", Journal of Modern Applied Statistical Methods, 14(2): 123-140, (2015).

[15] Dietrich, D., Hüsler, H., "Minimum distance estimators in extreme value distributions", Communications in Statistics - Theory and Methods, 25(4): 695–703, (1996).

[16] Mousa, M. A., Jaheen, Z. F., Ahmad, A. A., "Bayesian estimation, prediction and characterization for the Gumbel model based on records", Statistics: A Journal of Theoretical and Applied Statistics, 36(1): 65-74, (2002).

[17] Mahdi, S., Cenac, M., “Estimating parameters of Gumbel distribution using the methods moments, probability weighted moments and maximum likelihood”, Revista de Matemática: Teoría y Aplicaciones, 12: 151-156, (2005).

[18] Yılmaz, A., Kara, M., Özdemir, O., “Comparison of different estimation methods for extreme value distribution”, Journal of Applied Statistics, 48: 2259-2284, (2021).

[19] Furutani, H., Hiroyasu, T., Okuhara, Y., “Simple method for estimating daily and total COVID-19 deaths using a Gumbel model”, Researchsquare, (2020).

[20] Furutani, H., Hiroyasu, T., Okuhara, Y., “Method for estimating time series data of COVID-19 deaths using a Gumbel model”, Archives of Clinical and Biomedical Research, 6(1): 50-64, (2022).

[21] Hee, O., “Tests for predictability of statistical models”, Journal of Farm Economics, 48(5):1479-1484, (1996).

[22] Deb, K., Pratab, A., Agarwal, S., Meyarivan, T., “A fast and elitist multi-objective genetic algorithm: NSGA-II”, IEEE Transactions on Evolutionary Computation, 6(29): 182-197, (2002).