The impact of COVID-19 on the electricity demand: a case study for Turkey

Zeynep Ceylan

Samsun University, Faculty of Engineering, Industrial Engineering Department, Samsun, Turkey

Correspondence
Zeynep Ceylan, Samsun University, Faculty of Engineering, Industrial Engineering Department, Samsun, 55420, Turkey.
Email: zeynep.ceylan@samsun.edu.tr

Summary
Due to the extraordinary impact of the Coronavirus Disease 2019 (COVID-19) and the resulting lockdown measures, the demand for energy in business and industry has dropped significantly. This change in demand makes it difficult to manage energy generation, especially electricity production and delivery. Thus, reliable models are needed to continue safe, secure, and reliable power. An accurate forecast of electricity demand is essential for making a reliable decision in strategic planning and investments in the future. This study presents the extensive effects of COVID-19 on the electricity sector and aims to predict electricity demand accurately during the lockdown period in Turkey. For this purpose, well-known machine learning algorithms such as Gaussian process regression (GPR), sequential minimal optimization regression (SMOReg), correlated Nyström views (XNV), linear regression (LR), reduced error pruning tree (REPTree), and M5P model tree (M5P) were used. The SMOReg algorithm performed best with the lowest mean absolute percentage error (3.6851%), mean absolute error (21.9590), root mean square error (29.7358), and root relative squared error (36.5556%) values in the test dataset. This study can help policy-makers develop appropriate policies to control the harms of not only the current pandemic crisis but also an unforeseeable crisis.

KEYWORDS
COVID-19, economic impacts, electricity demand, lockdown period, machine learning, time series prediction

1 | INTRODUCTION

Coronavirus Disease 2019 (COVID-19) outbreak has dealt a significant blow to the ordinary course of life and caused a massive economic slowdown.¹,² Global economic recession and widespread negative growth figures are observed all over the world. According to the experts, the cost of COVID-19 to the world economy is estimated to be about $1 trillion in 2020.³ To understand the potential negative economic impact of COVID-19, Carlsson-Szlezak et al. (2020) examined three potential economic transmission channels.⁴ The first is the indirect hit to the wealth effect. Significant financial difficulties have arisen in people’s lives due to the damage of the COVID-19 to the economy. With the decline in household wealth, saving rates increased, and consumption expenditures fell. The second is the direct effect associated with reduced consumption habits of goods and services. The third is the supply-side effect, unlike the channel demand, shocks mentioned above. The virus has negatively affected supply chains, labor demand, and employment by disrupting production, causing longer layoffs and increased unemployment.⁵ The slowdown in production progress has led to business closures and operations to...
cease, as well as deeper financial problems, driving many firms to bankruptcy.\textsuperscript{6,7}

The way organizations and people use energy has also changed significantly due to the socio-economic effects of the epidemic.\textsuperscript{8} The reduced social and economic activity caused by the COVID-19 pandemic has affected all aspects of life, including the electricity sector. With the transition of educational institutions to distance learning mode, many public events and non-critical services started to be performed remotely. Many commercial and professional services started to be carried out from homes all over the world. As expected, this trend resulted in lower electricity consumption in the commercial and public sectors and an increase in residential electricity use. In addition to commercial and public sectors, the transportation sector has come under difficult conditions. For example, the cessation of the aviation industry and the disruption of related services at all airports resulted in a drastic drop in electricity demand. The reduction in traffic has also affected electricity demand, as many types of public transport run on electricity.

COVID-19 outbreak has also had a significant impact on energy production systems and the use of energy resources worldwide. Conventional power generation systems are generally affected only in terms of generation capacity. However, another essential type of energy source that has been significantly affected by the epidemic is the renewable energy generation sector, which has become widespread in recent years. Solar or wind energy systems, which require equipment, materials, and labor, are among the most affected renewable energy sectors. The economic effects of the epidemic slow down the expansion of existing basic clean energy technologies while limiting or preventing new investments. With the epidemic, except for a few types (eg, electric vehicles, lighting, and so on), renewable energy investments have been stopped. Also, serious decreases in fuel prices due to transportation restrictions and the closure of countries reduce the demand for biomass or biomass-derived biofuels such as pellets,\textsuperscript{9} biogas,\textsuperscript{10,11} biochar,\textsuperscript{12} vegetable oil-biodiesel,\textsuperscript{13} which are other clean, sustainable and renewable resources. The impact of the economic contraction in the sector necessitates the support of the governments for investments in renewable energy resources. For this reason, it is a critical issue to examine energy production during the pandemic process and to reshape energy production policies after the pandemic.

In the literature, many studies are conducted that examine the impact of the pandemic in various energy sectors. However, a few of these studies focus on the effects on the electrical power industry. Table 1 shows some studies on the impact of the COVID-19 epidemic and associated lockdowns on electricity consumption patterns in different countries.

For example, Carvalho et al. (2020) applied the joinpoint analysis to analyze the effect of mobility

## Table 1
The change in electricity consumption of different countries due to COVID-19

| Reference | Province/Country     | Impact of COVID-19 on electricity consumption                                                                 |
|-----------|----------------------|--------------------------------------------------------------------------------------------------------------|
| 14        | Brazil               | In the first quarter of 2020, the electricity consumption decreased by 0.9% compared to 2019, and residential, industrial, and commercial sectors decreased by 0.3%, 0.4%, and 2.2%. |
| 15        | Ontario, Canada      | The electricity demand declined by 14%, totaling 1267 GW for the month of April-2020.                          |
| 16        | European countries   | From 6 to April 13, 2020, Spain (25%) experienced the highest reduction in electricity demand, followed by Italy (17.7%), Belgium (15.6%), the United Kingdom (14.2), and Netherlands (11.6%). |
| 17        | United States        | The overall electricity demand has declined by less than 10%.                                                |
| 18        | India                | The daily electricity supply for the selected five weekdays in March and April 2020 decreased by 14-24% compared to the same days of 2019. |
| 19        | Australia            | The overall electricity demand fell by 6.7% in March 2020, while residential demand increased by 14% in Victoria, Australia. |
| 20        | Spain                | Power demand has decreased by 13.49% (from 14 March to 30 April 2020) compared to the average value of five previous years. |
| 21        | Lagos, Nigeria       | The industrial electricity consumption decreased by 24% and 18% under partial and total lockdown scenarios, respectively. |
| 22        | Italy                | The epidemic caused a decrease of up to 37% in electricity consumption in the first week of March 2020 compared to the same period of 2019. |
| 23        | Germany              | The average share of net electricity production using Renewable Energy Resources increased above 55% in the first half of 2020, compared to 47% for the same period in 2019. |
| 24        | China                | The demand for electricity decreases by 0.65% when the population of infected people increases by 1%.       |
restrictions caused by the COVID-19 pandemic on the Brazilian electricity demand patterns. Rayash and Dincer (2020) conducted a similar analysis for the province of Ontario, Canada. The results of the study showed a 14% decrease in monthly electricity demand, with the highest daily 25% decrease. Bahmanyar et al. (2020) compared the impact of different COVID-19 restrictions on electricity use in different European countries. Gillingham et al. (2020) used the polynomial regression and the two-step augmented regression model to provide insight into the effects of the pandemic on energy consumption in the United States. Kanitkar (2020) examined how lockdown measures affect electricity consumption in India. They used the linear input-output analysis to forecast the economic losses from the virus. Snow et al. (2020) focused on household electricity use in Australia during the COVID-19 pandemic. The study results revealed a significant increase in household electricity consumption during the COVID-19 lockdown period, such as cooking and the use of digital devices.

Similar to the study by Kanitkar (2020), Santiago et al. (2021) conducted a detailed analysis to determine how the containment measures affect the daily electricity consumption in Spain. In the study, the changes in the CO2 emission and electricity prices caused by the electricity demand change were also evaluated. Unlike the studies in the literature, Edomah and Ndulue (2020) separately examined the impact of COVID-19 electricity consumption on residential, commercial, and industrial sectors in Lagos, Nigeria. In another study, Ghiani et al. (2020) discussed the effects of the COVID-19 on the entire power system in Italy. They also conducted a statistical analysis to compare the different day-ahead electricity market price distributions pre- and post-the lockdown measures. The study results showed that the decrease in electricity consumption had a rapid effect on the day-ahead electricity market. Halbrügge et al. (2021) applied data visualization and descriptive statistics to reveal how German electricity systems were affected during the COVID-19 pandemic. The analysis results of the study revealed that COVID-19 caused not only noticeable changes in electricity consumption in Germany but also in pricing, import, and export. Norouzi et al. (2020) developed regressive and neural network models to determine the change in electricity and petroleum consumption in China, the first country to suffer from COVID-19. The study results indicated that the severity of the epidemic significantly affects the demand for electricity and petroleum, both directly and indirectly.

The studies mentioned above generally discuss the impact of the COVID-19 epidemic on electricity consumption. However, studies to predict sudden and dramatic changes in electricity demand during the pandemic are quite limited. Table 2 lists studies on electricity demand forecasting during the COVID-19.

For example, Huang et al. (2021) proposed a new optimized gray prediction model that predicts the electricity demand gap due to pandemic. They emphasized...
that a significant decrease in electricity consumption is strongly associated with the number of COVID-19 cases.\textsuperscript{25} Lu et al. (2021) developed a hybrid multi-objective optimizer-based model to predict the electricity consumption of the US during the epidemic. According to the study results, the highest prediction accuracy was obtained by including the daily number of infections in the developed model.\textsuperscript{26} Chen et al. (2020) developed a prediction model that considers mobility data as a measure of economic activities to estimate the electrical load of regions in the United States and Europe.\textsuperscript{27} Alhajeri et al. (2020) used regression analysis and the genetic algorithm to predict the electricity peak load and power generation of Kuwait during the pandemic. They also made a quantitative evaluation of the economic and environmental impacts of partial and full lockdowns.\textsuperscript{28} Alasali et al. (2021) developed a rolling stochastic Auto-Regressive Integrated Moving Average with Exogenous (ARIMAX) model to analyze the electricity demand data of three regions in Jordan.\textsuperscript{29}

The previous studies emphasize the importance and necessity of energy forecasting to help electricity companies manage their resources effectively and ensure a safe electricity supply to the end-user during the pandemic. Although previous studies have examined various effects of epidemic-like emergencies, research on the impact of the epidemic on the electricity sector is still needed due to the different severity of the epidemic in different countries.

Presenting a case study on Turkey, this paper highlights the importance and vulnerability of the energy sector due to the COVID-19 pandemic with a specific focus on electricity demand. This study aims to develop a robust and useful method that can accurately predict the electricity consumption gap during the COVID-19 lockdown period when there is an obvious and significant change in global trends. Considering the capability of the existing models, this uncontrollable and unpredictable contingency increases complexity and disrupts the predictive performance of classical models. Therefore, machine learning (ML) methods are a better option for capturing the change in electricity consumption with their nonlinear modeling capabilities. As far as the author aware, this is the first comprehensive study to use and compare different ML approaches to predict electricity consumption that deviates from its usual trend due to the COVID-19. Nowadays, as the effects of the virus continue, such studies can guide policy-makers to manage the effects of the economic recession during and after the pandemic. Thus, they can better anticipate the supply and demand reactions of the markets and ensure that these processes are overcome with the least damage.

The rest of the paper is organized as follows: Section 2 describes the case study, the data, and the methodology. Section 3 contains results and discussions, and Section 4 presents the conclusions and suggestions for future studies.

2 | METHODOLOGY

The methodology section of the study, which consists of five subsections, is shown in detail in Figure 1. Section 2.1 presents the case study for Turkey. In this subsection, the daily electricity consumption data of Turkey (starting from March 16, 2020 to June 1, 2020) is examined to better understand the changing dynamics of electricity consumption habits due to COVID-19. The daily electricity consumption data of Turkey was obtained from the Turkish Electricity Transmission Corporation (TEIAS) website.\textsuperscript{25} Section 2.2 presents some essential aspects of the data set used in the study. Details about the data preprocessing stage and the predictive variables used are given in this subsection. Section 2.3 presents the description and basic theoretical aspects of different ML techniques used within the scope of the study. Section 2.4

F I G U R E 1  Flowchart of research methodology
provides the performance evaluation criteria and statistical significance tests used in the study. Finally, Section 2.5 details the method used to rank the performance of the models.

2.1 | Case study

The COVID-19 disease was declared as a pandemic by the World Health Organization (WHO) on 11 March 2020. On the same day, the Turkish Ministry of Health reported the first case in Turkey. Immediately after the detection of the first COVID-19 case in Turkey, it was decided to be shut down the schools to curb the pandemic. Then, the government of Turkey declared a state of emergency, bars, nightclubs, theaters, cinemas, gyms, and cafes were closed. These measures were in full force during part of March 2020 and the entire month of April and May. Travel restrictions, curfews, unpaid leave practices, and working hours regulations have led to changes in sectoral demand, thus affecting both production and production demand. Many industrial centers have reduced production levels due to the lack of demand and interruptions in their supply chain. As a result, the public was subject to lockdown, which resulted in a sharp drop in economic output. The reduced social and economic activity caused by the COVID-19 pandemic had a dramatic impact on the country’s electricity demand. As seen in Figure 2, the electricity demand of Turkey fell sharply in May compared to March and April.

Total electricity consumption in April and May 2020 is 10.2% and 10.3% lower than the lowest April and May values recorded between 2016 and 2019, respectively. Compared to the same month of 2019, demand decreased by 15.4% and 16.7% in April and May.

In Figure 3, the weeks during lockdown measures in Turkey are compared with the corresponding weeks in 2019. Week 1 is 16-22 March 2020, which corresponds to 18-24 March 2019, and so forth (see Appendix). COVID-19 outbreak in Turkey has exceeded 100 cases as of 18 March 2020. In the same week, it was entered into a new era by the end of the working days in Turkey. Therefore, electricity consumption started to decrease as of 23rd March, and the demand decreased by 5.7% compared with the same week of the previous year. In the second week of April, when the full-curfew was imposed, electricity consumption decreased by 16.9%. Electricity demand fell to its lowest level in the last week of May, and electricity demand was 26.3% lower than the reference week of 2019.

As seen in Figure 4, the electricity demand fell to Sunday levels rather than a typical high-demand weekday pattern. A 48-hour curfew was imposed on the weekends in major cities where about 65 million people live. The electricity consumption hit the lowest level on Sunday at around 457 272 MW-hours, marking a nearly 34.7% decrease compared to the corresponding day last year.

Figure 5 shows daily electricity consumption from the beginning of March until the start of new normalization on 1 June 2020, and the corresponding days in 2019. This trend clearly shows that national electricity demand declined since mid-March when the first COVID-19 case was confirmed.
2.2 | Data preparation

In this study, Waikato Environment for Knowledge Analysis-version 3.9.4 (WEKA 3.9.4) software was used for time series forecasting. WEKA is developed by the University of Waikato, New Zealand, equipped with various ML algorithms written in Java programming language.32 WEKA includes a time series forecasting package that enables the selection of the basic learning model and parameters, generation of lagged variables, creation of variables derived from a date-time stamp, the addition of “overlay” data, evaluation options, and output control.

Lagged variables are effective tools that capture the relationship between past and current values of a series.33 Since the data used in this study is daily, the minimum and maximum lag lengths were set as 1 and 7, respectively. A value of “1” means a lagged variable that holds target values at time \( t - 1 \), while a value of 7 is a lagged variable that holds target values at time \( t - 7 \). All-time periods between minimum and maximum are turned into lagged variables. A fine-tuning lag selection was also performed to achieve high predictive performance by entering a random interval within the minimum and maximum variables.

Additional data that may be related to the prediction can improve the performance of learning algorithms. WEKA software uses the intervention attributes (also called overlay data) as inputs to the model. Thus, within the scope of this study, time series forecasting has been performed with the overlay data. Daily electricity consumption data in 2019 has been taken into consideration in the time series forecasting. The average temperature data for Turkey has also been taken into consideration from the Turkish State Meteorological Service Official website.34 Moreover, consumer behavior differs significantly from the regular pattern on certain days, such as national and religious holidays and special days. Thus, national holidays and curfews during the lockdown period were included in the analysis (1—yes; 0—no). WEKA provides many parameters and options to derive time-dependent attributes in time series analysis. Thus, in this study, new attributes such as “day of the week” and “weekends” that will affect the time series forecasting performance were derived.

Consumption data were collected daily in units of GWh. To check the predictive ability of the learning algorithms, the electricity consumption data was divided into two sets, namely training and testing sets. In this regard, all learning algorithms have been trained on the first 70% (54 samples) of the instances and tested on the last 30% (23 samples) of the instances. Thus, daily electricity consumption data between 16 March 2020 and 8 May 2020 were used as the training set. On the other hand, test data includes the data from 9 May 2020 to 31 May 2020.

2.3 | Regression models

In this study, various function- and tree-based ML algorithms in the WEKA software are used as a base learner
for time series forecasting. Function-based algorithms make predictions based on mathematical models. On the other hand, tree-based learning models are used for making predictions via a tree structure. Gaussian Process Regression (GPR), Sequential Minimal Optimization Regression (SMOReg), Correlated Nyström Views (XNV), and Linear Regression (LR) from function-based models and M5P model tree (M5P) and Reduced Error Pruning Tree (REPTree) from tree-based models were used. The mentioned ML algorithms are briefly explained in the subsections. Table 3 shows the default values of the hyperparameters used in regression models in this study.

### 2.3.1 Gaussian process regression

Gaussian process regression (GPR) is a powerful non-parametric kernel-based probabilistic model. It was introduced by Rasmussen and Williams.\(^{35}\) GPR is well adopted and widely used in various applications in engineering problems.\(^ {36}\) GPR implements Gaussian processes (GP) for regression without hyperparameter-tuning. GP is a set of random variables such that any finite number of them have a joint Gaussian distribution. The GP is specified by a mean \( \mu(x) \) and a covariance (kernel) function \( k(x, x') \). Given a function space \( f(x) = \phi(x)^T \omega \), and defining \( \left[ f(x^{(1)}), f(x^{(2)}), ..., f(x^{(n)}) \right] \) as the variable set, which follows a Gaussian distribution, the GPR model can be given as:

\[
f(x) \sim GP(\mu(x), k(x, x'))
\]

where

\[
\mu(x) = E[f(x)]
\]

\[
k(x, x') = E \left[ (f(x) - \mu(x))(f(x') - \mu(x'))^T \right]
\]

where \( E[.] \) denotes expectation. The function values \( f(x) \) are not achievable in most applications. In practice, only the noisy observations are available, which can be expressed by\(^ {35}\):

\[
\begin{align*}
y(x) &= f(x) + \epsilon \\
\epsilon &= \text{additive white noise}
\end{align*}
\]

where \( \epsilon \) is the additive white noise. It is assumed to be independent and identically distributed normal random noises with mean 0 (zero) and variance \( \sigma^2 \), \( \epsilon \sim \mathcal{N}(0, \sigma^2) \). Any finite number of the observed values can form an individual GP as stated by\(^ {35}\):

\[
f \sim GP(\mu, k), y \sim GP(\mu, k + \sigma^2 \delta)
\]

where \( \delta_{ij} \) is the Kronecker delta, which is set as 1 when \( i = j \) or 0 otherwise. The objective of the GPR model is to infer the function value \( f^* \) and its variance \( \text{cov}(f^*) \) based on the new test point \( x^* \). In this sense, \( X^* \) denotes the input matrix of the test dataset and \( N \) is the size of test dataset. According to the GPR assumption, the observed values and the function values at new test points follow the joint Gaussian process distribution:

\[
\begin{pmatrix}
y \\
f^*
\end{pmatrix}
\sim \mathcal{N}
\begin{pmatrix}
\mu(X) \\
\mu(X^*)
\end{pmatrix},
\begin{pmatrix}
K(X, X) + \sigma_n^2 I_n & K(X, X^*) \\
K(X^*, X) & K(X^*, X^*)
\end{pmatrix}
\]

where \( I \) denotes the \( n \times n \) unit matrix, \( K(X, X) \) is the covariance matrix of the training dataset, \( K(X^*, X^*) \) is the covariance matrix of test dataset, and \( K(X, X^*) \) is the covariance matrix obtained from the training and test
dataset, such that $K'(X',X) = K(X, X')^T$. Thus, we can obtain the predictive distribution:

$$f^* | X, y, X' \sim N \left( \left( \bar{f}^* \right), \text{cov}(f^*) \right)$$  \hspace{1cm} (7)

where

$$\bar{f}^* = \mu(X') + K(X', X) \left[ K(X, X) + \sigma_n^2 I_n \right]^{-1} (y - \mu(X))$$  \hspace{1cm} (8)

$$\text{cov}(f^*) = K(X', X') - K(X', X) \left[ K(X, X) + \sigma_n^2 I_n \right]^{-1} K(X, X')$$  \hspace{1cm} (9)

The main advantage of this method is that it has good generalization ability and is suitable for dealing with high dimensional spaces and small samples, and complex nonlinear problems.

### 2.3.2 Sequential minimal optimization regression

Support vector machine (SVM) is a supervised ML algorithm developed by Cortes and Vapnik in 1995. It has been widely used in recent years as a successful machine learning approach for solving real-life engineering problems. It is called support vector classification (SVC) and support vector regression (SVR) for classification and regression analysis, respectively. The aim of this algorithm is to find the optimal separating hyperplane that provides the maximum margin between the hyperplane and the support vectors. The schematic diagram of the SVM algorithm is given in Figure 6. The kernel function is used to construct a mapping from input vectors to a high dimensional feature space. These functions can be of various types, such as linear, polynomial, sigmoid, and radial basis function.

Suppose there are $N$ training elements $\{(x_i, y_i)\}$, for $i = 1, ..., N$, $y_i \in R$, $x_i \in R^D$, the SVM function can be represented, as shown in Equation (10)

$$y_i = w \cdot x_i + b$$  \hspace{1cm} (10)

where $w$ is weight, $b$ is the deviation vector. The SVM algorithm formula can be expressed as a constrained optimization problem as following equation.

$$\min \frac{1}{2} ||w||^2 + C \sum_{i=1}^{N} (\xi_i + \xi_i')$$  \hspace{1cm} (11)

subject to:

$$w \cdot x_i + b - y_i \leq \epsilon + \xi_i \forall i$$  \hspace{1cm} (12)

$$y_i - w \cdot x_i + b \leq \epsilon + \xi_i' \forall i$$  \hspace{1cm} (13)

$$\xi_i, \xi_i' \geq 0, \forall i$$  \hspace{1cm} (14)

where $\epsilon$ is the insensitive loss function, $\xi_i, \xi_i'$ are the non-negative slack variables, $C$ is the penalty factor (also known complexity parameter), and $y_i$ is the target value.

The sequential minimum optimization (SMO) technique was introduced by Platt. SMO significantly improved the computation efficiency and accuracy of the SVM model. Smola and Scholkopf improved the original SMO algorithm to solve regression problems. Then, Shevade et al. made further modifications to the SMO regression (SMOreg) model to achieve better performance. The improved SMOreg algorithm enhances the computation speed of SVM regression, even with small sample size data sets. Besides, it can reduce the risk of overfitting problems, which helps find the generalized pattern of the real-life problems.

### 2.3.3 Correlated Nyström views

Correlated Nyström views (XNV) was first proposed by McWilliams et al. Function XNV is a semi-supervised kernel-based modeling algorithm. It is assumed that the input set $x_i$ contains $n$ feature vectors and that only a small fraction of them is labeled $\{x_i, y_i\}_{i=1}^n$. As shown in
the pseudo-code algorithm below, the working procedure of this algorithm follows two main steps:

**Step 1: Generating random views.** Two view data sets are generated using random features that are not computationally expensive. In this stage, the Nyström method is used for constructing random features.

**Step 2: Multi-view regression.** Canonical correlation analysis (CCA) is applied to the two-view data set to bias the optimization procedure toward useful features. Finally, linear regression penalized with a canonical norm is applied.

According to McWilliams et al., the function XNV outperforms a state-of-the-art algorithm for semi-supervised learning. It is a fast algorithm that can be used for regression and classification tasks. The main advantage is that it significantly improves predictive performance and reduces performance variability across various real-world datasets in a shorter processing time.

### 2.3.4 Linear regression

Linear regression (LR) is a well-known and simple regression algorithm. LR presents a mathematical model of the relationship between a dependent variable and one or more independent variables. LR model can be expressed as in Equation (15):

$$y = \omega_0 + \omega_1 a_1 + \omega_2 a_2 + \omega_3 a_3 + \ldots + \omega_k a_k$$  \hspace{1cm} (15)

where \(\omega\) is the weight, \(a\) is the attribute, and \(y\) is predicted the class.

The algorithm iteratively attributes weights to the model and calculates the value of the class. The algorithm then tries to minimize the sum of the squares of the differences between the actual class value and the predicted class value using Equation (16).

$$\sum_{i=1}^{n} \left(x_i - \sum_{j=0}^{k} \omega_j a_i\right)^2$$  \hspace{1cm} (16)

### 2.3.5 M5P model tree

M5P model tree (M5P) is one of the decision tree deduction tools in machine learning. The original M5 algorithm was first introduced by Quinlan and then rearranged into M5P (M5') by Wang and Witten. It is a widely used decision tree algorithm by many researchers for prediction purposes.

There are three main steps in the M5P algorithm: (1) constructing the tree, (2) pruning the tree, and (3) smoothing the tree. The first step involves identifying different data classes and building a tree. M5P model trees split the input progressively. In the M5P algorithm, a split process starts based on the SD of the target value, which measures error at the node. The SD reduction (SDR) is expressed as follows:

$$SDR = \sigma(T) - \sum_{i=1}^{n} \sigma(T_i) \times \frac{|T_i|}{|T|}$$  \hspace{1cm} (17)

where \(T\) denotes the case that reaches the node, \(|T_i|\) is the number of cases that reach the node, \(n\) is the number of branches, \(T_i\) is the case belonging to the \(i^{th}\) branch, \(|T_i|\) is the number of cases related to the \(i^{th}\) branch, \(\sigma\) is the SD.

The splitting process is performed recursively by selecting the assignment with the maximum expected error reduction. The exact process is applied to the subsets. The splitting process in M5P ends when a slight change is observed in the output value or some examples are preserved. After the tree is constructed, a linear regression technique is then created for each subset of samples at the leaves. In the second step, the pruning process begins after the tree is constructed to overcome the overfitting problem by eliminating the undesired subtree. The third step involves the smoothing process to compensate for the sharp discontinuities between adjacent linear models at the pruned tree leaves.
2.3.6 | Reduced error pruning tree

Reduced error pruning tree (REPTree) is a fast decision tree learner with low computational cost.\(^\text{59}\) It is a popular method for classification/regression problems because of the simple configuration.\(^\text{60}\) It builds a classification tree using the information gain or makes a regression tree from the variance.\(^\text{61}\) At first, the model uses regression tree logic to create multiple trees in various iterations. The best tree is chosen from multiple trees. Then, the model uses the reduced error pruning technique with backfitting as a pruning procedure in order to improve generalization. Finally, the algorithm sorts the values of the numerical attributes using the embedded method and then handles the missing values using a C4.5 algorithm.\(^\text{62}\)

2.4 | Performance evaluation

In this study, four performance metrics, namely mean absolute percent error (MAPE%), mean absolute error (MAE), root mean square error (RMSE), and root-relative square error (RRSE%), were used to evaluate and compare models. MAPE is the percentage ratio of the mean absolute prediction errors to the absolute actual data. MAE is applied to reveal how close estimations are to the final results. RMSE gives information about the short-term performance of the prediction models. RRSE is used to evaluate the error rate. The model with the lowest MAPE%, MAE, RMSE, and RRSE% values is considered the best candidate model. The detailed equations are defined in Table 4.

In addition to the performance criteria mentioned above, two non-parametric statistical tests, the Wilcoxon signed-rank test\(^\text{63}\) and the Friedman ranking test\(^\text{64}\) were used to see any differences among the forecasting models. The following hypotheses were formulated:

H0. There is no significant difference in the results of the ML models used in this study.

H1. There is a significant difference in the results of the ML models used in this study.

The Wilcoxon signed-rank test is based on the “one by one” rule, while the Friedman test (also referred to as analysis of variance by ranks) is used for multiple comparisons.\(^\text{40,65}\) Both statistical techniques were performed on the testing dataset. Besides, the two-tailed test was conducted with a significance level of \(\alpha = 0.05\). Statistical analysis was carried out with SPSS Version 20.0 statistic software package.

2.5 | Model performance ranking

Evaluating the algorithms according to a single criterion sometimes misleads the decision to choose the best algorithm from the available candidate algorithms. Therefore, the regression models used in this study were ranked using the Technique for Order Preference by Similarity to Ideal Solutions (TOPSIS), one of the popular Multi-Criteria Decision Making (MCDM) approaches. In other words, the problem was handled as an MCDM problem that includes six models and four criteria.

TOPSIS was introduced by Hwang and Yoon in 1981.\(^\text{66}\) The underlying principle of the TOPSIS method is to determine the best alternative with the minimum distance from the positive ideal solution and the maximum distance from the negative ideal solution.

### Table 4  A summary of the performance evaluation metrics used in this study

| Metric | Equation | Value range | Description |
|--------|----------|-------------|-------------|
| MAPE   | \[ \frac{100}{n} \times \sum_{t=1}^{n} \left| \frac{x_t - y_t}{x_t} \right| \] | MAPE \(\leq 10\%\) | High prediction accuracy |
|        |          | 10\% \(<\) MAPE \(\leq 20\%\) | Good prediction |
|        |          | 20\% \(<\) MAPE \(\leq 50\%\) | Reasonable prediction |
|        |          | MAPE \(> 50\%\) | Inaccurate prediction |
| MAE    | \[ \frac{1}{n} \sum_{t=1}^{n} |x_t - y_t| \] | — | The lower the MAE, RMSE, and RRSE values, the better the model performance |
| RMSE   | \[ \sqrt{\frac{1}{n} \sum_{t=1}^{n} (y_t - x_t)^2} \] | — | — |
| RRSE   | \[ 100 \times \sqrt{\frac{\sum_{t=1}^{n} (y_t - x_t)^2}{\sum_{t=1}^{n} (x_t - \bar{x})^2}} \] | — | — |

*Note: \(y_t\) and \(x_t\) are the predicted and measured values at time point \(t\), respectively. Also, \(\bar{x}\) is the mean of measured values, and \(n\) is the number of time points.*
Assuming an MCDM problem with \( m \) alternatives (regression models) and \( n \) criteria (model performance measures) the TOPSIS procedure is as follows.\(^6\)

Develop a normalized decision matrix \( (r_{ij}) \) using Equation (18):

\[
  r_{ij} = \frac{X_{ij}}{\sqrt{\sum_{i=1}^{m} X_{ij}^2}} \quad i = 1, 2, ..., m \text{ and } j = 1, 2, ..., n \quad (18)
\]

where \( X \) is the \( m \times n \) decision matrix of performance ratings of each alternative at each criterion.

Construct a weighted normalized decision matrix \( (v_{ij}) \) by assigning weights to each criterion as follows:

\[
  v_{ij} = w_j \times r_{ij} \quad i = 1, 2, ..., m \text{ and } j = 1, 2, ..., n \quad (19)
\]

where \( w_i \) is the weight of \( j \)-th criteria and \( \sum_{j=1}^{n} w_j = 1 \).

Compute the positive ideal \( (A^+) \) and negative ideal solution \( (A^-) \) using Equations (20) and (21), respectively.

\[
  A^+ = \{ (\max_i v_{ij} | j \in h), (\min_i v_{ij} | j \in l) \} \quad i = 1, 2, ..., m \quad (20)
\]

\[
  A^- = \{ (\min_i v_{ij} | j \in h), (\max_i v_{ij} | j \in l) \} \quad i = 1, 2, ..., m \quad (21)
\]

where \( h \) and \( l \) indicate the sets of indices \( j \) corresponding to the benefit criteria and negative criteria.

Compute the separation measures \( S_i^+ \) from positive ideal and \( S_i^- \) from negative ideal solution by using Euclidean distance as follows:

\[
  S_i^+ = \sqrt{\sum_{j=1}^{n} (v_{ij} - A^+)^2} \quad i = 1, 2, ..., m \quad (22)
\]

\[
  S_i^- = \sqrt{\sum_{j=1}^{n} (v_{ij} - A^-)^2} \quad i = 1, 2, ..., m \quad (23)
\]

Determine the relative closeness of each alternative by calculating the performance score \( (p_i) \):

\[
  p_i = \frac{S_i^-}{S_i^- + S_i^+} \quad i = 1, 2, ..., m \quad (24)
\]

Rank the alternatives according to the scores \( p_i \).

### 3 | RESULTS AND DISCUSSION

In this section, the results of the time series analysis are given. Section 3.1 shows the performance comparison of models, statistical test results, and residual analysis. Finally, a detailed discussion is provided in Section 3.2.

#### 3.1 | Comparison of machine learning algorithms

As discussed previously, the objective of the study is to examine the predictability of the electricity demand of Turkey via six different ML algorithms during the COVID-19 lockdown period. To evaluate the performance of the models, 70% of the dataset was used for training and the remaining 30% for testing. The reliability of each model was measured by four accuracy criteria, namely, MAPE (%), MAE, RMSE, and RRSE (%). Table 5

| Dataset | Algorithm | MAPE (%) | MAE | RMSE | RRSE (%) |
|---------|-----------|----------|-----|------|----------|
| Training set | GPR | 2.6115 | 16.5504 | 20.5693 | 33.7073 |
| | SMOReg | 1.3415 | 8.6483 | 13.9753 | 22.9016 |
| | XNV | 1.7246 | 11.0944 | 14.5786 | 23.8903 |
| | LR | 1.7200 | 11.0724 | 14.3515 | 23.8579 |
| | M5P | 2.5504 | 16.0555 | 19.8384 | 39.3506 |
| | REPTree | 2.9718 | 19.1722 | 25.9112 | 42.4612 |
| Test set | GPR | 5.2732 | 33.9158 | 41.5626 | 51.0949 |
| | SMOReg | 3.6851 | 21.9590 | 29.7358 | 36.5556 |
| | XNV | 3.9717 | 24.8220 | 31.8105 | 39.1062 |
| | LR | 3.9237 | 24.4739 | 31.4022 | 38.6043 |
| | M5P | 4.7526 | 28.1889 | 40.4870 | 49.7727 |
| | REPTree | 7.8421 | 45.3113 | 58.2825 | 71.6496 |

### TABLE 5 Prediction performance of learning algorithms
depicts the performance metric values of the learning algorithms. All regression models performed well in the range of $8.6483 \leq \text{MAE} \leq 19.1722$, $13.9753 \leq \text{RMSE} \leq 25.9112$, and $22.9016\% \leq \text{RRSE} \leq 42.4612\%$ in the training stage. The MAPE values vary between $1.3415\%$ and $2.9718\%$ in the training stage and $3.6851\%$ and $7.8421\%$ in the testing stage. Based on these findings, it can be concluded that the performance of all ML algorithms used to predict electricity consumption in this study is significantly good and reliable (MAPE $\leq 10$).

To select the best performing model according to the evaluation criteria mentioned above, the TOPSIS method was used. This provides a more systematic ranking system that makes the algorithm selection process more applicable and robust. Equal weight was assigned to each performance metric such that the sum of all weights is equal to 1. Table 6 shows the ranking of the algorithms according to their relative closeness values. Rank 1 is the higher preferred algorithm, while rank 6 is less preferred. Accordingly, the SMOReg algorithm is the best model for predicting electricity consumption change due to the COVID-19 pandemic (MAPE: 3.6851\%, MAE: 21.9590, RMSE: 29.7358, RRSE: 36.5556\%). It is followed by the LR model (MAPE: 3.9237\%, MAE: 24.4739, RMSE: 31.4022, RRSE: 38.6043\%), the XNV algorithm (MAPE: 3.9717\%, MAE: 24.8220, RMSE: 31.8105, RRSE: 39.1062\%), respectively.

Table 6  Ranking of regression models

| Algorithm | $S_+^i$ | $S_-^i$ | Performance score | Rank |
|-----------|---------|---------|-------------------|------|
| GPR       | 0.0876  | 0.0662  | 0.4305            | 5    |
| SMOReg    | 0.1531  | 0.0000  | 0.0000            | 1    |
| XNV        | 0.1402  | 0.0134  | 0.0870            | 3    |
| LR         | 0.1423  | 0.0113  | 0.0737            | 2    |
| M5P        | 0.1055  | 0.0488  | 0.3161            | 4    |
| REPTree    | 0.0000  | 0.1531  | 1.0000            | 6    |

As shown in Table 7, the considerable differences among the models were also analyzed using the Friedman test and Wilcoxon signed-rank tests at the level of 95\% confidence. Using the Friedman test, it was observed that the $P$-value is less than the significant level of .05. Thus, it can be stated that there is a statistically significant difference in the performance among all regression models ($\chi^2 = 40.216$, $df = 5$, $p = 0.000$). On the other hand, the Wilcoxon signed-rank test was also performed to determine whether there was a significant one-to-one difference between SMOReg and other models. The $P$-values of the model pairs are less than .05, and the $z$-values are below $-1.960$ (Critical values for 5\% two-tailed: $\pm 1.96$). Accordingly, the null-hypothesis between the SMOReg model and any of the compared ML algorithms can be rejected. The results of both statistical tests confirm the existence of a difference between the models used in this study.

To better understand the predictive success of the SMOReg model, the actual values were plotted against the predicted values in Figure 8A,B. Using the data recorded from TEIAS, the electricity demand forecast is formulated with a 95\% confidence level. It can be said that the SMOReg technique successfully captures actual electricity consumption data, especially when the data suddenly decreased. When the curve of the electricity load data reaches a peak or trough, the SMOReg model can adopt these changes well.

3.2 | Discussion

The nation and/or provincial lockdown implemented to slow down the spread of infection has affected the operational activities of businesses and industries, which resulted in a significant decrease in electricity demand. This unprecedented situation makes it difficult to forecast electricity demand. This unexpected change disrupts the
predictive performance of existing models, making the electrical system difficult to manage. In this uncertain and unclear period, policy-makers need reliable models to monitor the impact of COVID-19 in real-time and develop policies that can respond quickly to this change. Thus, accurate forecasting of electricity demand is essential for the management, operation, and control of the power system. Therefore, forecasting the future trend of the power demand with reliable and robust models is crucial.

As in other countries, restrictive measures were taken to slow the spread of COVID-19 in Turkey since mid-March 2020. Taking the daily electricity demand in Turkey as a case study, this study discusses the impact of the COVID-19 outbreak on electricity consumption. It provides a comprehensive comparison of ML techniques for predicting electricity demand during the lockdown period. Based on the analysis results, it can be seen clearly that the SMOReg model yielded better results than the benchmark models to predict the daily electricity consumption during the pandemic. This may be due to the advantage of SMOReg to cope with higher dimensional feature spaces compared to other regression algorithms. Another strength of the SMOReg model is that it can be successfully applied to different data sets.

Given that the impacts of COVID-19 vary from country to country and governments take different degrees of restrictions at different periods, it should be noted that electricity consumption patterns are likely to vary. Therefore, this model can be used in other countries and regions to monitor the extent of the disruption caused by the pandemic in real-time. This only requires data on electrical load, temperature, and COVID-19 restrictions, and such information is available to the public in almost all developed countries in the world.

However, there are some limitations in this study. First, it can be said that electricity demand patterns will gradually return to their normal trajectories as countries begin to ease lockdown and shutdowns. Although the reduction in electricity demand will not be persistent, the trajectory of the epidemic is still unclear. With the second wave of the epidemic, if the lockdowns are renewed, the
measures may be reactivated, which directly affects the consumption behavior of electricity demand. Thus, the study results will need to be updated periodically to ensure high prediction accuracy. Second, this study addresses only the short-term impacts of COVID-19 on the electricity sector during the lockdown period.

**FIGURE 7** Comparison of models using (A) violin plot (B) histogram

**TABLE 7** Results of Wilcoxon signed-rank test and Friedman test

| Compared models       | Wilcoxon signed-rank test | Friedman test |
|-----------------------|---------------------------|---------------|
|                       | z-value       | P value | Chi-square | P value |
| SMOReg vs GPR         | −4.076        | .000'   | 40.267     | .000'   |
| SMOReg vs XNV         | −3.619        | .000'   |            |         |
| SMOReg vs LR          | −3.133        | .002'   |            |         |
| SMOReg vs M5P         | −2.007        | .045'   |            |         |
| SMOReg vs REPTree     | −2.220        | .026'   |            |         |

*Note: 'Statistically different with α = 0.05.*
Therefore, this study does not consider the potential for long-term and recurrent lockdowns with cumulative and increasingly adverse economic impacts. Finally, while satisfactory short-term forecast results are provided with national electricity load change data in the study, individual effects of the epidemic in various sectors such as the transport, industrial sector, residential, and commercial and public services are not evaluated. Analyzing changes in electricity consumption patterns on a sector basis can be helpful to evaluate the impact of COVID-19 measures on particular sectors.

4 | CONCLUSIONS AND FUTURE RESEARCH

The lockdown measures implemented by many countries to control the spread of COVID-19 have direct and indirect effects on human life and the world economy. The electricity sector is an essential part of economic growth in every country. Thus, it is essential to determine a suitable model that can estimate fluctuations in electricity consumption. In this context, this study provides a comprehensive comparison of ML techniques to estimate the short-term impact of COVID-19 on the economy using daily electricity consumption data of Turkey. The findings are summarized as follows:

- There has been a significant reduction in electricity demand in Turkey in May 2020. In this regard, demand decreased approximately 17% compared to 2019 levels.
- The most significant decrease in electricity usage was on the weekends, with a decline of 34.7% compared to the same period of 2019.
- The SMOReg model outperformed the GPR, XNV, LR, M5P, and REPTree models in predicting electricity consumption during the COVID-19 lockdown period.

Consequently, this study discusses the severity and impact of COVID-19 on the electrical power industry as a case study. The results of this study showed that the
SMOReg method could be used as a stable and robust method for short-term estimation of electricity demand in any crisis (ie, pandemics, economic collapses, etc.). Therefore, this study can be useful for governments and policy-makers to develop appropriate policies.

While still experiencing the pandemic, it is essential to comprehensively examine the crisis and analyze its comprehensive and multidimensional effects. Many more issues remain to be investigated regarding the impact of the COVID-19 pandemic and the resulting lockdown measures on the energy sector. For example, it can be useful to examine the effects of measures during the COVID-19 lockdown period on fossil fuel demand and renewable energy. From an environmental perspective, the estimation of CO2 emission levels associated with energy demand during the epidemic could be another critical subject for future research.

ACKNOWLEDGMENTS

No funding to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Zeynep Ceylan https://orcid.org/0000-0002-3006-9768

REFERENCES

1. Norouzi N, Zarazua de Rubens GZ, Enevoldsen P, et al. The impact of COVID-19 on the electricity sector in Spain: an econometric approach based on prices. Int J Energy Res. 2020;1-13. https://doi.org/10.1002/er.6259.

2. Ceylan Z. Estimation of COVID-19 prevalence in Italy, Spain, and France. Sci Total Environ. 2020;729:138817. https://doi.org/10.1016/j.scitotenv.2020.138817.

3. Elavarasan RM, Shafuillah GM, Raju K, et al. COVID-19: impact analysis and recommendations for power sector operation. Appl Energy. 2020;279:115739. https://doi.org/10.1016/j.apenergy.2020.115739.

4. What Coronavirus Could Mean for the Global Economy [Internet]. https://hbr.org/2020/03/what-coronavirus-could-mean-for-the-global-economy. Cited January 23, 2021

5. Beyer RCM, Franco-Bedoya S, Galdo V. Examining the economic impact of COVID-19 in India through daily electricity consumption and nighttime light intensity. World Dev. 2021; 140:105287. https://doi.org/10.1016/j.worlddev.2020.105287.

6. Kliestik T, Misankova M, Valaskova K, Svabova L. Bankruptcy prevention: new effort to reflect on legal and social changes. Sci Eng Ethics. 2018;24:791-803. https://doi.org/10.1007/s11948-017-9912-4.

7. Kovacova M, Kliestik T, Valaskova K, Durana P, Juhaszova Z. Systematic review of variables applied in bankruptcy prediction models of Visegrad group countries. Oecconomia Copernicana. 2019;10:743-772. https://doi.org/10.24136/oc.2019.034.

8. Sausen ATZR, de Campos M, Sausen PS, et al. Classification of the social distance during the COVID-19 pandemic from electricity consumption using artificial intelligence. Int J Energy Res. 2021;1-11. https://doi.org/10.1002/er.6418.

9. Jandačka J, Mičieta J, Holubčík M, et al. Experimental determination of bed temperatures during wood pellet combustion. Energy Fuel. 2017;31:2919-2926. https://doi.org/10.1021/acs.energyfuels.6b03071.

10. Maroušek J, Strunecký O, Kolář L, et al. Advances in nutrient management make it possible to accelerate biogas production and thus improve the economy of food waste processing. Energy Sources, Part A Recover Util Environ Eff. 2020;00:1-10. https://doi.org/10.1080/15567036.2020.1776796.

11. Maroušek J, Bartoš P, Filip M, et al. Advances in the agrochemical utilization of fermentation residues reduce the cost of purpose-grown phytomass for biogas production. Energy Sources, Part A Recover Util Environ Eff. 2020;1-11. https://doi.org/10.1080/15567036.2020.1738597.

12. Mardoyan A, Braun P. Analysis of czech subsidies for solid biofuels. Int J Green Energy. 2015;12:405-408. https://doi.org/10.1080/15435075.2013.841163.

13. Maroušek J. Economic analysis of the pressure shockwave disintegration process. Int J Green Energy. 2015;12:1232-1235. https://doi.org/10.1080/15435075.2014.895740.

14. Carvalho M, de Delgado DB, de Lima KM, et al. Effects of the COVID-19 pandemic on the Brazilian electricity consumption patterns. Int J Energy Res. 2021;45(2):3358-3364. https://doi.org/10.1002/er.5877.

15. Abu-Rayash A, Dincer I. Analysis of the electricity demand trends amidst the COVID-19 coronavirus pandemic. Energy Res Soc Sci. 2020;68:101682 https://doi.org/10.1016/j.erss.2020.101683.

16. Bahmanyar A, Estebarsi A, Ernst D. The impact of different COVID-19 containment measures on electricity consumption in Europe. Energy Res Soc Sci. 2020;68:101683. https://doi.org/10.1016/j.erss.2020.101683.

17. Gillingham KT, Knittel CR, Li J, et al. The short-run and long-run effects of Covid-19 on energy and the environment. Joule. 2020;4:1337-1341. https://doi.org/10.1016/j.joule.2020.06.010.

18. Kanitkar T. The COVID-19 lockdown in India: impacts on the economy and the power sector. Glob Transitions. 2020;2:150-156. https://doi.org/10.1016/j.glt.2020.07.005.

19. Snow S, Bean R, Glencross M, Horrocks N. Drivers behind residential electricity demand fluctuations due to COVID-19 restrictions. Energies. 2020;13:5738. https://doi.org/10.3390/en13215738.

20. Santiago I, Moreno-Munoz A, Quintero-Jiménez P, et al. Electricity demand during pandemic times: the case of the COVID-19 pandemic in Spain. Energy Policy. 2021;148:111964. https://doi.org/10.1016/j.enpol.2020.111964.

21. Edomah N, Ndulue G. Energy transition in a lockdown: an analysis of the impact of COVID-19 on changes in electricity demand in Lagos Nigeria. Glob Transitions. 2020;2:127-137. https://doi.org/10.1016/j.glt.2020.07.002.

22. Ghiani E, Galici M, Mureddu M, Pilo F. Impact on electricity consumption and market pricing of energy and ancillary services during pandemic of COVID-19 in Italy. Energies. 2020;13:3357. https://doi.org/10.3390/en13113357.

23. Halbrügge S, Schott P, Weibelzahl M, et al. How did the German and other European electricity systems react to the
COVID-19 pandemic? Appl Energy. 2021;285:116370. https://doi.org/10.1016/j.apenergy.2020.116370.

24. Norouzi N, Zarazua de Rubens G, Choubanpishehzafar S, et al. When pandemics impact economies and climate change: exploring the impacts of COVID-19 on oil and electricity demand in China. Energy Res Soc Sci. 2020;68:101654. https://doi.org/10.1016/j.erss.2020.101654.

25. Huang L, Liao Q, Qiu R, et al. Prediction-based analysis on power consumption gap under long-term emergency: a case in China under COVID-19. Appl Energy. 2021;283:116339. https://doi.org/10.1016/j.apenergy.2020.116339.

26. Lu H, Ma X, Ma M. A hybrid multi-objective optimizer-based model for daily electricity demand prediction considering COVID-19. Energy. 2021;219:119568. https://doi.org/10.1016/j.energy.2020.119568.

27. Chen Y, Yang W, Zhang B. Using mobility for electrical load forecasting during the COVID-19 pandemic. arXiv. 2020;1-8.

28. Alhajeri HM, Almutairi A, Alenezi A, et al. Energy demand in the State of Kuwait during the Covid-19 pandemic: technical, economic, and environmental perspectives. Energies. 2020;13(17):4370. https://doi.org/10.3390/en13174370.

29. Alasali F, Nusair K, Alhmoud L, et al. Impact of the COVID-19 pandemic on electricity demand and load forecasting. Sustainability. 2021;13(3):1435. https://doi.org/10.3390/su13031435.

30. WHO | World Health Organization [Internet]. https://www.who.int/. Cited October 9, 2020.

31. Covid19 [Internet]. https://covid19.saglik.gov.tr/?_Dil=2. Cited October 9, 2020

32. Weka 3—Data mining with open source machine learning software in Java [Internet]. https://www.cs.waikato.ac.nz/ml/weka/. Cited July 9, 2020.

33. Time series analysis and forecasting with Weka—Pentaho data mining—Pentaho Wiki [Internet]. https://wiki.pentaho.com/pages/viewpage.action?pageId=23527839.

34. Williams CKI, Rasmussen CE. Gaussian Processes for Machine Learning. Cambridge, MA: MIT Press; 2006.

35. Turkish State Meteorological Service Official Web Sites [Internet]. https://mgm.gov.tr/eng/forecast-cities.aspx. Cited October 8, 2020.

36. Ceylan Z. Estimation of municipal waste generation of Turkey. Waste Manage Res. 2020;38(8):840-850. https://doi.org/10.1177/0734242X20906877.

37. Wang J, Hu J. A robust combination approach for short-term wind speed forecasting and analysis—combination of the ARIMA (autoregressive integrated moving average), ELM (extreme learning machine), SVM (support vector machine) and LSSVM (least square SVM) forecasts using a GPR (Gaussian Process Regression) model. Energy. 2015;93:41-56. https://doi.org/10.1016/j.energy.2015.08.045.

38. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20:273-297.

39. Chen W, Hasanipanah M, Rad HN, et al. A new design of evolutionary hybrid optimization of SVR model in predicting the blast-induced ground vibration. Eng Comput. 2019;1-17. https://doi.org/10.1007/s00366-019-00895-x.

40. Zhang Z, Ding S, Sun Y. A support vector regression model hybridized with chaotic krill herd algorithm and empirical mode decomposition for regression task. Neurocomputing. 2020;410:185-201. https://doi.org/10.1016/j.neucom.2020.05.075.

41. Ceylan Z, Bulkan S, Elevli S. Prediction of medical waste generation using SVR, GM (1,1) and ARIMA models: a case study for megacity Istanbul. J Environ Health Sci Eng. 2020;18:687-697. https://doi.org/10.1007/s40201-020-00495-8.

42. Fan GF, Qing S, Wang H, Hong WC, Li HJ. Support vector regression model based on empirical mode decomposition and auto regression for electric load forecasting. Energies. 2013;6(4):1887-1901. https://doi.org/10.3390/en6041887.

43. Smola AJ, Schölkopf B. A Tutorial on support vector regression. Statistics and Computing. 2004;14(3):199-222. https://doi.org/10.1023/B:STCO.0000035301.49549.88.

44. Platt J. Fast training of support vector machines using sequential minimal optimization. In Advances in Kernel Method: Support Vector Learning. Cambridge, MA, United States: MIT Press; 1999;185-208.

45. Chen S, Mihara K, Wen J. Time series prediction of CO2, TVOC and HCHO based on machine learning at different sampling points. Build Environ. 2018;146:238-246. http://www.science direct.com/science/article/pii/S0360132318306127.

46. Shevade SK, Keerthi SS, Bhattacharyya C, Murthy KRK. Improvements to the SMO algorithm for SVM regression. IEEE Trans Neural Netw. 2000;11:1188-1193.

47. McWilliams B, Balduzzi D, Buhmann M. Correlated random features for fast semi-supervised learning. Adv Neural Inf Proces Syst. 2013;1-15.

48. Trokicić A, Todorović B. Randomized Nyström features for fast regression: an error analysis. In: Ćirić M, Droste M, Pin J-É, eds. Algebra Informatics. Cham, Switzerland: Springer International Publishing; 2019:249-257.

49. Gao W, Alsarraf J, Moayed H, et al. Comprehensive preference learning and feature validity for designing energy-efficient residential buildings using machine learning paradigms. Appl Soft Comput J. 2019;84:105748. https://doi.org/10.1016/j.asoc.2019.105748.

50. Zhao J, Lam KP. Influential factors analysis on LEED building markets in U.S. East coast cities by using support vector regression. Sustainable Cities and Society. 2012;5:37-43. https://doi.org/10.1016/j.scs.2012.05.009.

51. Mohammed A, Rafiq S, Sihag P, et al. ANN, M5P-tree and ANNs and the M5P model tree. Appl Soft Comput. 2020;98(20):106097. https://doi.org/10.1016/j.asoc.2020.106097.

52. Quinlan JR. Learning with continuous classes. Aust Jt Conf Artif Intell. 1992;3:343-348.

53. Wang Y, Witten IH. Induction of Model Trees for Predicting Continuous Classes. Hamilton, New Zealand: University of Waikato, Department of Computer Science; 1996;

54. Balouchi B, Nikoo MR, Adamowski J. Development of expert system for the prediction of scour depth under live-bed conditions at river confluences: application of different types of ANNs and the M5P model tree. Appl Soft Comput. 2015;34:51-59. https://doi.org/10.1016/j.asoc.2015.04.040.

55. Azadi S, Amirji H, Rakshshande Shroo GR. Evaluating the ability of artificial neural network and PCA-MSP models in predicting leachate COD load in landfills. Waste Manag. 2016;55:220-230. https://doi.org/10.1016/j.wasman.2016.05.025.

56. Behnood A, Behnood V, Modiri Ghareshveran M, et al. Prediction of the compressive strength of normal and high-performance concrete in Kordestan and Behbahan using the linear correlation coefficient and artificial neural network methods, the hybrid model and the bayesian optimization. J Environ Health Sci Eng. 2019;18:687-697.
concretes using M5P model tree algorithm. Constr Build Mater. 2017;142:199-207. https://doi.org/10.1016/j.conbuildmat.2017.03.061.

57. Khosravi K, Cooper JR, Daggupati P, et al. Bedload transport rate prediction: application of novel hybrid data mining techniques. J Hydrol. 2020;585:124774 https://doi.org/10.1016/j.jhydrol.2020.124774.

58. Bhattacharya B, Solomatine DP. Neural networks and m5 model trees in modelling water level—discharge relationship. Neurocomputing. 2005;63:381-396. https://doi.org/10.1016/j.neucom.2004.04.016.

59. Witten IH, Frank E. Data mining: practical machine learning tools and techniques with Java implementations. ACM SIGMOD Rec. 2002;31(1):76-77. https://doi.org/10.1145/507338.507355.

60. Chen W, Hong H, Li S, et al. Flood susceptibility modelling using novel hybrid approach of reduced-error pruning trees with bagging and random subspace ensembles. J Hydrol. 2019;575:864-873. https://doi.org/10.1016/j.jhydrol.2019.05.089.

61. Joseph KS, Ravichandran T. A comparative evaluation of software effort estimation using REPTree and K* in handling with missing values. Aust J Basic Appl Sci. 2012;6:312.

62. Jayanthi SK, Sasikala S. Reptree classifier for identifying link spam in web search engines. IJSC. 2013;3(2):498-505. https://doi.org/10.21917/ijsc.2013.0075.

63. Wilcoxon F. Individual comparisons of grouped data by ranking methods. J Econ Entomol. 1946;39(2):269-270. https://doi.org/10.1093/jee/39.2.269.

64. Friedman M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J Am Stat Assoc. 1937;32(200):675-701. https://doi.org/10.2307/2279372.

65. Li MW, Geng J, Hong WC, et al. Periodogram estimation based on LSSVR-CCPSO compensation for forecasting ship motion. Nonlinear Dyn. 2019;97:2579-2594. https://doi.org/10.1007/s11071-019-05149-5.

66. Hwang C-L, Yoon K. Methods for multiple attribute decision making. Multiple Attribute Decision Making. Berlin, Heidelberg: Springer; 1981:58-191. https://doi.org/10.1007/978-3-642-48318-9_3.

67. Meng D, Shao C, Zhu L. A new energy efficiency evaluation solution for ethylene cracking furnace based on integrated TOPSIS-DEA model. Int J Energy Res. 2020;44:732-755. https://doi.org/10.1002/er.4826.

68. Lazaridis A, Mporas I, Ganchev T, Kokkinakis G, Fakotakis N. Improving phone duration modelling using support vector regression fusion. Speech Communication. 2011;53(1):85-97. http://dx.doi.org/10.1016/j.specom.2010.07.005.

How to cite this article: Ceylan Z. The impact of COVID-19 on the electricity demand: a case study for Turkey. Int J Energy Res. 2021;45:13022–13039. https://doi.org/10.1002/er.6631