Forecasting Hourly Electricity Demand Under COVID-19 Restrictions

Ali Kök, Ergün Yükseltan, Mustafa Hekimoğlu*, Esra Agea Aktunc, Ahmet Yücekaya, Ayşe Bilge

Industrial Engineering Department, Kadir Has University, Istanbul, Turkey. *Email: mustafa.hekimoglu@khas.edu.tr

Received: 06 August 2021 Accepted: 10 November 2021 DOI: https://doi.org/10.32479/ijeep.11890

ABSTRACT

The rapid spread of the COVID-19 pandemic has severely impacted many sectors including the electricity sector. The restrictions such as lockdowns, remote-working, and -schooling significantly altered the consumers’ behaviors and demand structure especially due to a large number of people working at home. Accurate demand forecasts and detailed production plans are crucial for cost-efficient generation and transmission of electricity. In this research, the restrictions and their corresponding timing are classified and mapped with the Turkish electricity demand data to analyze the impact of the restrictions on total demand using a multiple linear regression model. In addition, the model is utilized to forecast the electricity demand in pandemic conditions and to analyze how different types of restrictions impact the total electricity demand. It is found that among three levels of COVID-19 restrictions, age-specific restrictions and the complete lockdown have different effects on the electricity demand on weekends and weekdays. In general, new scheduling approaches for daily and weekly loads are required to avoid supply-demand mismatches as COVID-19 significantly changed the consumer behavior, which appears as altered daily and weekly load profiles of the country. Long-term policy implications for the energy transition and lessons learned from the COVID-19 experience are also discussed.

Keywords: COVID-19, Pandemic, Electricity Demand, Daily Demand Curve, Restrictions, Regression

JEL Classifications: Q47, E17, Q40.

1. INTRODUCTION

The impact of COVID-19 on the economy, industry, health, education, and other critical sectors are obvious and the damage will need to be recalculated when the pandemic period is over. The demand for electric energy is highly impacted as a result of restrictions and the changing habits of the users. Electricity is a commodity that is delivered to end-users passing through planning, generation, and transmission steps. The fuel and resources are planned based on the demand projections, they are ordered and transported if necessary, and the power plants are scheduled to run to generate the electricity that will be injected into the transmission system. However, the planning is triggered using the forecasted demand. The electricity demand forecasting thus has been a subject for quite a number of studies in the literature. However, the proposed forecasting methods use past data that has cyclic behaviors and trends, assuming that the demand will follow a similar pattern. The demand still needs to be forecasted under the presence of COVID-19 restrictions as the electricity is still generated to meet the demand. It has been quite some time since the pandemic period started and the post-pandemic consumption pattern is still unknown. Hence, electricity forecasting models, considering altered consumption behaviors of both household and industrial consumers, are needed to estimate demand accurately.

Electricity demand depends on industrial facilities and household consumption behaviors. There is a sharp decline in electricity demand due to the pandemic measures. The sudden stop of the production of large companies in the industrial sector, such as iron-steel, glass, ceramics, and cement factories, has a significant effect on this decline. The limitations on the operation of industrial facilities, schools, shopping centers, and other non-critical
Electricity demand needs to be forecasted considering the changing structure of the demand drivers and the restrictive limitations of the pandemic. Although the impact of the pandemic is obvious, the demand still needs to be forecasted for market operations and system planning purposes. There are studies in the literature for electricity demand forecasting, however, they are tailored for normal circumstances with no restrictions. Hence, the electricity demand forecasting addressing the long- and short-term effects of COVID-19 on electricity demand rate is a research gap in the literature.

Linear models and time series methods are commonly used in the literature for demand forecasting. Suganthi and Samuel, 2012 present the literature on forecasting methods including Artificial Neural Networks (ANN), Genetic Algorithms (GA), Support Vector Machines (SVM), Particle Swarm Optimization (PSO), and other numerical methods. ARMA and ARIMA models are also used to include the stochastic effects in demand forecasting. A detailed description of the literature on forecasting methods is also given in (Yukseltan et al., 2017).

The impact of temperature on electricity demand depends on the infrastructure and heating resources, and the temperature is used to increase the forecast accuracy. Different aspects of the influence of the temperature on the electricity demand have been analyzed in (Baştə and Helman, 2013; De Felice et al., 2015; Hor et al., 2005; Islam et al., 1995; Luis et al., 2017; Momani, 2013; Taylor, 2012). The seasonal cycles determine the impact of temperature on electricity demand especially if the electricity is used for heating and cooling needs. On the other hand, it is shown that relative humidity, solar radiation, cloudiness, and other climatic parameters can have a role in demand forecasting. Different studies, classified according to their forecasting methods, are given in Table 1. The methodologies can be classified as time series analysis, statistical methods, surveys, ANN and simulation, heuristic approaches, and temperature-based methods. The main assumptions in all of these studies are to forecast the demand based on normal conditions without any pandemic restrictions.

Methods for forecasting electricity demand may be roughly classified as the ones that are “autoregressive”, i.e., the ones using electricity data only, and the ones that use exogenous regressors. For long-term forecasts (over a few years’ horizon) these exogenous regressors should include economic and demographic growth parameters. In cases where electricity consumption for heating and cooling is negligible, linear regression, including sinusoidal functions, captures cyclic variations in the electricity consumption arising from illumination demand and economic activities; and it provides reliable medium-term forecasts (over a year’s horizon) (Yukseltan et al., 2017). In cases where electricity is heavily used for heating and cooling, deviations from comfortable temperatures should be added as an exogenous regressor. ARMA type time series models are useful for following short-term variations (over a horizon of a few days) faithfully. ARMA models can be incorporated with linear models to obtain sharp forecasts taking into account temperature effects or other irregular variations (Yukseltan et al., 2020). These methods that have been tested under “normal” conditions, may fail to explain sudden variations due to unexpected major events, such as global changes in lifestyles and in economic activities that have been in force starting from March 2020, due to the COVID-19 pandemic. The effects of COVID-19 on the economy, energy sector, and environment have been analyzed in various papers such as (Abu-Rayash and Dincer, 2020; Bahmanyar et al., 2020; Cakmakli et al., 2021; Chen et al., 2020; Corpus-Mendoza et al., 2021; Ghiani et al., 2020; López Prol and O, 2020).

Previously, authors studied the changes in the electricity consumption in Turkey due to COVID-19, during the period January-June 2020 (Yucekaya et al., 2020). In that time frame, complete lockdowns and various age and sector-dependent restrictions were applied during the period March-May 2020. In the framework of the project, a population-based index is formed to model the effects of restrictions. In a previous paper, a modulated Fourier series expansion is used for modeling the decrease in the

| Table 1: Overview of the forecasting methods and related resources |
|-----------------------------|---------------------------|
| **Methods**                  | **Sources**               |
| Time Series Analysis         | (Conejo et al., 2005), (Suganthi and Samuel, 2012), (Clements et al., 2016), (Niu et al., 2010), (Andersen et al., 2013), (Lo and Wu, 2003) |
| Statistical Methods          | (Vilar et al., 2012), (Taylor, 2010), (Fan and Hyndman, 2012), (Wang et al., 2012), (Hyndman and Fan, 2010), (McSharry et al., 2005), (Taylor, 2003), (Chakhchoukh et al., 2011), (Apadula et al., 2012), (Ren et al., 2016), (Fili̇k et al., 2011) |
| Surveys                      | (Dyner and Larsen, 2001), (Suganthi and Samuel, 2012), (Hahn et al., 2009), (Conejo et al., 2005) |
| ANN and Simulation           | (Zhang and Dong, 2001), (Wang and Ramsay, 1998), (Srinivasan et al., 1995) |
| Heuristic Approaches         | (AlRashidi and EL-Naggar, 2010), (Wang et al., 2012), (Zhu et al., 2011), (Azadeh et al., 2007), (Pai and Hong, 2005) |
| Temperature-based Methods    | (Taylor and Buizza, 2003), (De Felice et al., 2013), (De Felice et al., 2015), (Crowley and Joutz, 2003), (Luisis et al., 2017), (Islam et al., 1995), (Hor et al., 2005), (Momani, 2013), (Baştə and Helman, 2013) |
In this paper, a multiple linear regression model is proposed to forecast the total electricity demand for the Turkish Electricity Market under the COVID-19 restrictions. The restrictions are classified and mapped with the total electricity demand to analyze the change in the demand. Then the regression model is utilized to estimate the impact of each restriction on the demand. Specifically, the electricity consumption of pre-pandemic and pandemic periods are compared using average hourly demand rates for weekends and weekdays. The proposed regression model satisfactorily explains the marginal contribution of each restriction type to demand loss for weekends and weekdays. In addition, the proposed model is utilized to obtain an importance ranking among the regressors in the model. It is found that religious holidays and the effect of Monday create the most significant effect on electricity demand whereas binary variables indicating day-time hours are the least important ones in the model. The proposed methodologies present novel approaches to literature and it is conjectured that these findings can be generalized to other countries and can be helpful to explain changes in electricity demand due to the COVID-19 pandemic.

The remainder of the paper is as follows. In Section 2, an overview of the Turkish Power market and the impact of COVID-19 on the total and sectoral electricity demand are presented. In Section 3, a multiple linear regression model, which is proposed to forecast the total electricity demand for pandemic and non-pandemic conditions, is presented. Section 4 is devoted to the presentation of importance ranking among regressors and its implications for electricity demand forecasting. Section 5 includes a discussion on our results and Section 6 presents the conclusion and suggestions for future directions.

2. DATA ANALYSIS AND PROCESSING

Aggregate statistics for the electricity consumption, generation, and pricing data for Turkey are released as hourly values by Energy Exchange Istanbul (EPIAS in Turkish), the electricity system operator of Turkey. In addition, sector-based statistics are released by Energy Market Regulatory Authority (EMRA) as monthly data. EPIAS data includes hourly demand and generation rates, the distribution of total generation by different resources, and price information. The present work is based on the data from EPIAS and EMRA for the period January-June 2020. The period July-December 2020 is also analyzed, but it is not included in this paper since the industrial sector started operating and the effects of restrictions are not significant for that period.

The COVID-19 epidemic starting in December 2019 in China, progressed rapidly as a pandemic. By March 2020, many European countries had applied various social and economic restrictions. An extensive database for the timing of the restrictions is available at the University of Oxford’s Coronavirus Government Response Tracker, together with a contingency index based on the types of restrictions (University of Oxford, 2020). Nevertheless, this index targets health-related effects and it was found to fail to explain the changes in electricity consumption. For the purpose of explaining the effects of restrictions on electricity consumption, we formed our own contingency indices for Turkey.

In Turkey, the first COVID-19 case and the first fatality were reported on March 11th, 2020, and March 18th, 2020, respectively. The first restriction, imposing a stay-at-home requirement for people of ages above 65 and below 20, was announced on March 21st. Subsequent restrictions are classified as age-based, travel-based, sector-based (closure of the noncritical facilities) restrictions, and complete lockdowns. A detailed description of these restrictions is given below. Level 2 restrictions include the shutdown of mostly service industry while Level 3 restrictions include the shutdown of the production industry.

2.1. Level 1 (Age-Specific Restrictions)
COVID-19 presents high risk for people above 65, while young people may have the infections without showing any symptoms. Age restrictions were imposed as stay-home requirements for people above 65 and below 20, to diminish health risks for the first group and to prevent the spread of the epidemic due to undetected positive cases.

2.2. Level 2 (Social Restrictions and Business Shutdown)
Social restrictions that were imposed in Turkey starting in April 2020 included travel restrictions between cities, closing of restaurants and cafés, suspension of sports events, online learning in schools and universities, and remote working in most offices.

2.3. Level 3 (Lockdowns and Industrial Shutdown)
In Turkey complete lockdowns were imposed during weekends and public holidays in April and May 2020. Complete lockdowns during the period March-June 2020 included the shutdown of all non-critical industries.

The total demand for the period January-June 2020 is presented in Figure 1, together with the timing of the restrictions. The first restriction was imposed in March, various types of restrictions were imposed in April and May, and the transition to normal life started at the beginning of June. During summer 2020, a limited form of age-based travel restrictions was applied. Although COVID-19 cases started to increase in the fall, no restrictions had been imposed until December 2020. Hence, the March-June 2020 period is selected as the analysis period. A qualitative analysis of Figure 1 reveals that the total demand decreases gradually as the level of restrictions increase and the effects of curfews are more dominant compared to others. Creating classification scenarios as low, medium, and high level is a useful method for more effective investigation and analysis of the impact of these restrictions.
Furthermore, the sudden decline in total demand observed at the end of May corresponds to the Ramadan Feast.

The sudden decline in demand is expected to affect the revenues of electricity producers significantly, and if the pandemic process is prolonged, it is worried that this will have permanent consequences in the medium term. Declining demand is expected to lead to lower prices for the market under normal conditions due to the merit-based market system. In this case, the production of more costly power plants such as those using natural gas is likely to decrease. Total electricity demand in Turkey in January-June 2019 and 2020 is shown in Figure 2.

2.4. Effect of COVID-19 Restrictions on Electricity Demand of Different Sectors

Examining sector-based electricity consumption is essential to reveal the underlying reasons for the change in aggregate demand. Because, while total demand has increased or decreased, sector-based demand increases or decreases regardless of the total. Sectoral consumption data for Turkey’s electricity market is published as monthly sector reports by Energy Market Regulatory Authority (EMRA). Since there are no more detailed data such as hourly and daily demand in the published reports, monthly data is studied. Monthly electricity consumption data is collected from the monthly sector reports for the last 5 years (until June 2020), and sector-based increases and decreases are examined (EMRA, 2020). The total consumption is divided into five sectors in the published reports, namely, lighting, residential, industrial, agricultural irrigation, and commercial. The residential, industrial, and commercial sectors, which are most affected during the pandemic, are examined. The consumption data of these three sectors observed in the January-June periods of the last 5 years are given in Figures 3-5.

**Figure 1:** COVID-19 restrictions and total electricity demand in Turkey (March-June, 2020) (Yukseltan et al., 2021)

**Figure 2:** Total electricity demand in Turkey 2019-2020 (EPIAS, 2020)
Figure 3: Commercial consumption between January-June 2016-2020

As seen in Figure 3, in March, April, and May 2020, the commercial consumption has gone out of the trend seen in the previous years and has shown a severe decrease. Despite the normalization that started in June, commercial electricity demand remained below the previous 4 years.

As seen in Figure 4, industrial consumption has shown a sharp decline in April and May 2020. Consumption in March 2020 remained above the consumption of March 2019. The reason for this is that no restriction decision was taken for the industrial sector in March 2020. In June 2020, a serious increase was observed in contrast to the trend seen in previous years. This is because companies that had to suspend production due to the restrictions in April and May, with the start of normalization as of June, are assumed to have produced above their regular routines to meet the demand that they could not meet before.

As seen in Figure 5, residential consumption leaped out of the trend observed in previous years in April 2020 and followed a similar path to the trend of previous years in the other months of 2020. Based on this, it can be said that there is no serious increase in residential consumption except in April. The reason for this can be shown as the fact that the curfews in Turkey do not cover long periods but short periods such as weekends, and the first curfew was implemented in April.

Finally, taking the year 2019 as a reference, the way the demand changed in March, April, May, and June 2020 is calculated for all three sectors. Percentage changes obtained as a result of these calculations are presented in Table 2.

As seen in the consumption change percentages given in Table 2, commercial consumption is most affected by the pandemic restrictions. Even in June, when the normalization started, a sharp decrease of approximately 21% was observed in commercial consumption compared to the previous year. Industrial consumption is the most affected sector after commercial consumption. Industrial consumption was not affected by the restrictions in March, showed a significant decrease in April and May. Afterward, a recovery effect with a significant increase of 17% is observed with normalization in June. Household consumption showed no unusual increase except in April but increased in all months compared to the previous year. As a result, it is understood that the significant decrease observed in total demand originates from industrial and commercial consumption, and even though household consumption has increased, these two sectors have dominated this increase.

3. FORECASTING DEMAND USING MULTIPLE LINEAR REGRESSION

Regression models are widely used in the literature for electricity demand forecasting since they are easy to implement and interpret (Kuster et al., 2017; Singh and Khatoon, 2013). In this study, a multiple linear regression model is developed to forecast the future demand and to describe the effects of nation-wide restrictions on the electricity demand of Turkey. To this end, hourly electricity consumption data from January 1, 2018 to June 30, 2020 is collected. To model the effects of restrictions, binary variables for
each time period are employed. When a restriction is present, its value is set to 1 and 0 otherwise. Similarly we employ \( R_i \), \( i = 1, 2, 3 \) for \( R_i \), the curfew for age groups young and old citizens (ages under 20 and over 65), \( R_i \), the is full curfew, \( R_i \), is the intercity transportation ban, and \( c \) is the residual term. As full curfew includes all age groups, values of other restriction variables are set to 0 when \( R_i \) is 1. Likewise, an intercity transportation ban is a more comprehensive ban than an age-group-specific curfew. Hence, \( R_i = 0 \) whenever \( R_i = 1 \). These variable configurations yield orthogonal regressors in our model.

The resulting model’s coefficient of determination, \( R^2 \), value is 0.453. To increase the model’s ability to explain variability in the dependent variable, more regressors are introduced into the model. Binary variables \( h_i, \ i = 0, \ldots, 22 \) indicates hours of the day, \( h_i = 1 \) if it is the \( i \)-th hour of the day and 0 otherwise. \( wd, j = 1, \ldots, 6 \) represents weekdays from Monday to Saturday respectively, \( wd = 1 \) if it is the \( i \)-th day of the week, 0 otherwise. \( m, j = 1, \ldots, 11 \) stands for months of the year, \( m = 1 \) if it is the \( i \)-th month of the year and 0 otherwise. \( d \) is a day-time indicator which depends on the season. Note that the binary variables \( h_i, wd_i \) and \( m_i \) are presented in a way that does not cover the last period of their own. Presenting the last periods becomes trivial since these variables are mutually exclusive within themselves, and they cover the whole dataset; in other words, they create linear dependency. \( d = 1 \) if it is daytime and 0 otherwise. Time periods assumed as daytime or nighttime to determine the value of \( d \) in different seasons are given in Table 3. \( h_i \), \( I \in \{ p, r \} \) represents public \(( i=p)\) and religious \(( i=r)\) holidays. It is known that during public and religious holidays, electricity demand significantly drops as manufacturing facilities are closed. \( h_i = 1 \) if that day is an \( i \)-th holiday and 0 otherwise. \( ne \) stands for number of cases for a given date. \( pr \) is another binary variable representing post-restriction days, i.e., \( pr = 1 \) if the date is later than May 31.

In order to understand autocorrelation in the electricity demand, the autocorrelation coefficient with a lag of 168 periods is estimated as 0.83 as shown in Figure 6b. The resulting model has an adjusted \( R^2 \) of 0.854, and the F-statistic of the model is 2375 (P-value smaller than 2.2e-16).

To include daily and weekly cycles, a sinusoidal regression model similar to Shah et al., 2019 is considered. In the sinusoidal regression, a time index \( t \in \{ 1, \ldots \} \), where 1 represents the earliest date is created. In addition to these cyclic variables, covariates for their interactions of regressors are added to the model. Specifically, the interaction term between full curfew and daily cycles (daily periodicity), \( \sin \left( \frac{2\pi}{24} t \right) R_i, \cos \left( \frac{2\pi}{24} t \right) R_i \), and post-restriction dates and daily periodicity, \( \sin \left( \frac{2\pi}{24} t \right) pr, \cos \left( \frac{2\pi}{24} t \right) pr \), are added to the model. The resulting model has an adjusted \( R^2 \) of 0.854, and the F-statistic of the model is 2375 (P-value smaller than 2.2e-16).

In order to evaluate the possible nonlinear relationship between dependent variable and regressors, Box-Cox transformation is considered that is given below: \( D' = D^{\frac{\lambda - 1}{\lambda}} \). To calculate the optimal \( \lambda \) value, we employed the boxcox routine from MASS package of R GUI that computes the optimal \( \lambda \) value using the log-likelihood function (Venables and Ripley, 2002). This routine yields 0.22 for the optimal \( \lambda \). Although this \( \lambda \) value does not imply exact log-transformation, log-transformation is chosen since it is close to 0. The summary statistics indicate that taking the natural logarithm increased the adjusted \( R^2 \) to 0.8649, and the model could explain a large proportion of the variation in the dependent variable. The resulting mathematical model of the log-linear multiple regression model is given below:

\[
\ln(D_t) = \beta_0 + \sum_{i=1}^{3} \beta_i R_i + \sum_{i=0}^{22} \beta_i h_i + \sum_{i=0}^{6} \beta_i wd_i + \sum_{i=1}^{11} \beta_i m_i + \beta_d dt + \sum_{i=p,r} \beta_i wd_i + \beta_p nc + \beta_r pr + \beta_{sin} \sin \left( \frac{2\pi}{24} t \right) + \beta_{cos} \cos \left( \frac{2\pi}{24} t \right) + \beta_{sin} \sin \left( \frac{2\pi}{168} \right) + \beta_{cos} \cos \left( \frac{2\pi}{168} \right) + \beta_{sin} \sin \left( \frac{2\pi}{24} \right) + \beta_{cos} \cos \left( \frac{2\pi}{24} \right) + \beta_{sin} \sin \left( \frac{2\pi}{24} \right) + \beta_{cos} \cos \left( \frac{2\pi}{24} \right) + \epsilon #(1)
\]

The F-statistic of the model is 2597 (P-value smaller than 2.2e-16). The summary statistics indicate strong overall significance for the model. In this model, there are 56 regressors. To simplify the model, the stepwise regression method is employed.

Stepwise regression is the process of repetitively introducing or eliminating regressors by obtaining a subset of predictors resulting in the best performing model (Kassambara, 2017). It is recognized as an application of feature selection in the machine learning
literature (Alpaydin, 2020). There are three strategies of the stepwise regression method: forward selection, backward selection, and stepwise selection (Bruce and Bruce, 2017; James et al., 2013). For all strategies, different criteria, such as $R^2$, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), are suggested for addition or elimination operations on regressors. AIC and BIC are extended versions of the log-likelihood statistic of the regression model. They include penalizing terms in favor of model simplification (Alpaydin, 2020). Chakrabarti and Ghosh, 2011 state that the BIC is more appropriate in choosing the correct model, while the AIC is more suitable for choosing the best model to forecast future observations. In our study, thus, the backward selection with the AIC criterion is employed to simplify our regression model with 56 regressors. To implement, the step routine in R Gui (R Core Team, 2020) is utilized. The backward selection of this routine initially takes all variables into account. Then, repeatedly extracts the predictor that made the least contribution to the model in terms of AIC. The program routine stops when removing any regressor does not improve (decrease) the AIC any further. After applying the stepwise regression with backward selection, the following regression model is obtained:

$$
\ln (D_i) = \beta_0 + \sum_{i=1}^{12} \beta_i R_{\text{th}} + \sum_{i=4}^{22} \beta_{i, h} h_i + \sum_{i=11}^{46} \beta_{i, w} w_d_i + \sum_{i=13,5,6,7,8,9,10,11} \beta_{i, m} m_i + \beta_{d, t} t_i + \sum_{i=12,\sigma} \beta_{i, h} h_i + \beta_{i, n} n_i + \beta_{i, p} p_i \
$$

$$+ \beta_{i, \sin} \sin \left( \frac{2 \pi i}{168} t \right) + \beta_{i, \cos} \cos \left( \frac{2 \pi i}{168} t \right) + \beta_{2i} \sin \left( \frac{2 \pi i}{24} t \right) + \beta_{2i} \cos \left( \frac{2 \pi i}{24} t \right) + \epsilon \in \#(2)
$$

The adjusted $R^2$ value of the model is 0.865, meaning that the model could explain a large proportion of the variation in the dependent variable. Although the stepwise regression improved the adjusted $R^2$ value marginally, it simplified the initial model with 56 regressors to the stepwise model with 52 regressors. Note that stepwise regression removed the statistically insignificant regressors only. The residual standard error is 0.058 which shows that the model fits the data very well. The F-statistic is 2697 with a P-value smaller than 2.2e-16, indicating strong overall significance for the model. The estimates of regression coefficients are given in Table 4 and all the covariates in this table are statistically significant with P-values of zero.

To visualize the efficiency of the regression model, a line plot depicting actual and fitted electricity demand is provided in Figure 7. As can be seen from our regression fit plot, the model captures cyclic demand patterns as well as sudden drop due to COVID-19 restrictions starting from the middle of March, 2020.

### 4. APPLICATIONS OF THE REGRESSION MODEL

In this section, two applications of the regression model presented above are provided: First, the details of the forecasting study and tests for measuring the predictive power of the model are provided. Second, a discussion on the importance ranking of independent variables for explaining the electricity demand is presented.

#### 4.1. Demand Forecasting Under COVID-19 Restrictions

In order to forecast hourly electricity demand under COVID-19 restrictions, the model suggested above is subjected to tests for its prediction accuracy. First, the error terms are analyzed for the existence of any bias from 0. As can be seen from Figure 8, the error terms of the model have a symmetric distribution around 0 in the periods before and after COVID-19. This indicates the success of the regression equation for modeling the hourly electricity demand before COVID-19 period and the efficiency of binary variables to explain the effect of COVID-19 on hourly electricity demand.

In order to test the predictive power of the model, two types of tests are performed. First, k-fold cross-validation is implemented on the whole time series, including the pandemic period. In the k-fold cross-validation, first, the data set is split into k random subsets. One of the k subsets is chosen as a test set and the model is trained on the remaining parts of the data. The trained model is used to generate predictions for the test set and the prediction error is computed. These steps are repeated until each of the k subsets is used as a test set. As an accuracy measure, the average of the prediction errors of k test sets is calculated. In this study, k is set to...
Table 4: The estimates of the coefficients of each independent variable obtained from stepwise regression (P-values are omitted as they are all equal to 0)

| Variable | Estimate | Variable | Estimate | Variable | Estimate |
|----------|----------|----------|----------|----------|----------|
| $\beta_0$ | 10.3800  | $h_{13}$ | 0.0501   | $m_3$    | -0.0481  |
| $R_1$    | -0.0548  | $h_{14}$ | 0.0709   | $m_4$    | -0.0732  |
| $R_2$    | -0.1416  | $h_{15}$ | 0.0717   | $m_5$    | -0.0757  |
| $R_3$    | -0.0896  | $h_{16}$ | 0.0825   | $m_6$    | -0.0211  |
| $h_0$    | -0.0589  | $h_{17}$ | 0.0955   | $m_7$    | 0.0986   |
| $h_1$    | -0.1108  | $h_{18}$ | 0.1020   | $m_8$    | 0.0898   |
| $h_2$    | -0.1480  | $h_{19}$ | 0.1039   | $m_9$    | -0.0110  |
| $h_3$    | -0.1733  | $h_{20}$ | 0.0912   | $m_{10}$ | -0.1023  |
| $h_4$    | -0.1846  | $h_{21}$ | 0.0679   | $m_{11}$ | -0.0567  |
| $h_5$    | -0.1868  | $h_{22}$ | 0.0417   | $dt$     | -0.0191  |
| $h_6$    | -0.1781  | $wd_4$  | 0.1113   | $hdr$    | -0.1056  |
| $h_7$    | -0.1321  | $wd_5$  | 0.0967   | $R_t$    | -0.3068  |
| $h_8$    | -0.0215  | $wd_6$  | 0.0594   | $nc$     | 0.0000   |
| $h_9$    | 0.0372   | $wd_7$  | 0.0574   | $pr$     | -0.0115  |
| $h_{10}$ | 0.0550   | $wd_8$  | 0.0813   | $dt$     | -0.0191  |
| $h_{11}$ | 0.0698   | $wd_9$  | 0.0769   | $h$      | 0        |
| $h_{12}$ | 0.0396   | $m_2$   | 0.0146   | $pr$     | -0.0115  |
| $n_c$    | 0.0000   | $pr$    | -0.0545  | $R_t$    | 0.0113   |

$\sin\left(\frac{2\pi}{168} t\right)$

$\cos\left(\frac{2\pi}{168} t\right)$

$R_t \sin\left(\frac{2\pi}{24} t\right)$

$\cos\left(\frac{2\pi}{24} t\right) pr$

$0.1511$

The mean MAPE value of the 14 steps is 7.6. As seen in Figure 9, however, the MAPE value of the first step, 27.3, is too high compared to others and raises the average. This is because there is no data regarding the pandemic period in the first training set. Thus, it may be unfair to include the MAPE value of the first step in the average. If we exclude it, the mean MAPE value of 27, 2020, and the second test set range is March 28, 2020-April 3, 2020. In this manner, with 1-week incremental expansions, 14 different training and test sets are formed. The Mean Absolute Percentage Error (MAPE) values for each of the 14 steps are given below:

Second, the prediction performance of the proposed model is tested using a rolling horizon approach. Specifically, data until the start of the COVID-19 pandemic is taken as a training set, which consists of data starting from January 1, 2019 until the first pandemic restriction is applied. The test set consists of data starting from the start of the pandemic restrictions. Demand forecasts are calculated for the 1st week of the test set using the trained model. Afterward, the training set is extended with the actual demand of the forecasted week, and after training, the regression model is used to forecast the 1st week of the test set. The first restriction (curfew for people over 65) is applied on March 21, 2020 at 0:00. Therefore, the first training set range is January 1, 2019-March 20, 2020, and the range of the first test set is March 21, 2020- March 27, 2020. The second training set range is January 1, 2019- March 27, 2020, and the second test set range is March 28, 2020-April 3, 2020. In this manner, with 1-week incremental expansions, 14 different training and test sets are formed. The Mean Absolute Percentage Error (MAPE) values for each of the 14 steps are given below:
the 13 steps reduces to 6.09. Note that before calculating MAPE, the predicted values are transformed to their actual magnitude by taking natural exponential.

The effects of restrictions are measured for weekdays and weekends separately, as well as for all days, using the proposed model, and these are compared with the empirical effects as presented in Table 5. These results show that age-specific restrictions (Level 1) have a larger effect on weekdays whereas complete lockdown (Level 3) has a larger effect on weekends. The relative effects of restrictions on all days and the total demand losses are closely estimated by the proposed model. The only important difference between model fit and the empirical results is the effect of Level 3 lockdown on weekends. A closer look at the results reveals that the main difference between the empirical results and the model partly stems from the transition periods from Level 3 lockdown to normalization. In addition, Turkey applied a softer version of Level 3 restrictions in the second half of May 2020 which altered the hourly consumption pattern. We find that the model’s fit from May to July gets significantly poorer compared to April, the beginning of Level 3 restrictions.

The effects of restrictions on the electricity demand over hours of the day are also shown in Figure 10, where the significant impact of complete lockdown and industrial shutdown (Level 3) is clearly observed, especially during the daytime.

4.2. Importance Ranking of Regressors

In the regression model, electricity demand is forecasted based on the future values of covariates in the model. However, future values of independent variables are also subject to uncertainty and their actual values can be different than the anticipated values at the time of estimation. Therefore, in regression modeling, it is essential to determine the sensitivity of the estimated value to small changes in the regressors. This sensitivity information is analyzed through different regression statistics in the literature. One of the common ways to do so is using standardized regression coefficients.

Coefficients of a regression model are affected by the magnitudes of the values of each covariate. Therefore, it is impossible to assess the importance of a regression variable within the model by just looking at its regression coefficient. This problem can be eliminated by applying Z-transform to each covariate before including it into the regression model. The resulting regression coefficient is called standardized regression coefficients and they are recognized to be an efficient tool for analyzing the sensitivity of the estimated value to each regressor. Hekimoğlu and Barlas (2016) utilize standardized regression coefficients’ absolute values to calculate the most important regressor in regression models. In addition to its magnitude, the sign of a standardized regression coefficient indicates if the estimate will increase when the value of the regressor is found to be higher. In this study, their approaches are followed to obtain an importance ranking of regressors for the estimated electricity demand. To this end, normalized values of each variable and the standardized regression coefficients are calculated. Then, all regressors are ordered with respect to the magnitude of the regression coefficient and indexed starting from 1. Standardized regression coefficients and the importance ranking of all regressors are given in Table 6.

Our results indicate that $hd \cdot \cos\left(\frac{2\pi t}{168}\right)$, $wd$, are the most important regressors in the model. $hd$ represents the religious holiday within a year and the electricity demand severely drops during those days since almost all industrial and commercial facilities are shut down. As expected, its coefficient is found to be negative. Interestingly, the effect of public holidays on the demand estimate is found to be much smaller (ranked 27th) in the model. This can be attributed to the fact that some facilities might still be running and consuming electricity during public holidays without much resistance from their employees. The second most important regressor is the cosines function reflecting the importance of the weekly cycles within the model. The third most important regressor in the model is the binary variable representing the Mondays in the model. This can be attributed to the fact that all facilities start working on the 1st day of the week; hence, on that day the electricity consumption is significantly elevated compared to the other days of the week. Interestingly, the second most important weekday is Tuesday in our model (ranked 8th). The combined effect of these 2 days is dubbed start-of-the-week effect in our analysis.

Table 5: Effects of restrictions on electricity demand for weekends and weekdays, empirical vs. modelled results

| Restrictions | Model Fit | Empirical |
|--------------|-----------|-----------|
|              | Weekdays  | Weekends  | All Days | Weekdays  | Weekends  | All Days |
| Level 1      | 8.1%      | 8.2%      | 7.6%     | 7.4%      | 5.6%      | 6.4%     |
| Level 2      | 4.8%      | 5.5%      | 2.8%     | 5.6%      | 5.0%      | 3.9%     |
| Level 3      | 5.3%      | 5.4%      | 13.6%    | 4.7%      | 11.9%     | 14.5%    |
| Total loss   | 18.2%     | 19.2%     | 24.0%    | 17.8%     | 22.5%     | 24.8%    |
Table 6: Standardized regression coefficients of the regression model and their importance ranking

| Variable | Standardized Regression Coeff. | Importance Ranking | Variable | Standardized Regression Coeff. | Importance Ranking |
|----------|--------------------------------|--------------------|----------|--------------------------------|--------------------|
| $R_1$    | -0.0600                        | 39                 | $h_{21}$ | 0.0851                        | 32                 |
| $R_2$    | -0.1410                        | 18                 | $h_{22}$ | 0.0525                        | 42                 |
| $h_3$    | -0.0748                        | 36                 | $wd_1$  | 0.2451                        | 3                  |
| $h_4$    | -0.1410                        | 17                 | $wd_2$  | 0.2102                        | 8                  |
| $h_5$    | -0.1891                        | 9                  | $wd_3$  | 0.1264                        | 21                 |
| $h_6$    | -0.2223                        | 7                  | $wd_4$  | 0.1239                        | 23                 |
| $h_7$    | -0.2379                        | 5                  | $wd_5$  | 0.1795                        | 10                 |
| $h_8$    | -0.2421                        | 4                  | $wd_6$  | 0.1713                        | 12                 |
| $h_9$    | -0.2325                        | 6                  | $wd_7$  | 0.0279                        | 47                 |
| $h_{10}$ | -0.1756                        | 11                 | $wd_8$  | -0.0887                       | 31                 |
| $h_{11}$ | -0.0366                        | 46                 | $wd_9$  | -0.1385                       | 19                 |
| $h_{12}$ | 0.0370                         | 45                 | $wd_{10}$| -0.1453                       | 15                 |
| $h_{13}$ | 0.0588                         | 40                 | $wd_{11}$| -0.0400                       | 43                 |
| $h_{14}$ | 0.0773                         | 35                 | $wd_{12}$| 0.1576                        | 14                 |
| $h_{15}$ | 0.0391                         | 44                 | $wd_{13}$| 0.1435                        | 16                 |
| $h_{16}$ | 0.0529                         | 41                 | $wd_{14}$| -0.0172                       | 48                 |
| $h_{17}$ | 0.0801                         | 34                 | $wd_{15}$| -0.1633                       | 13                 |
| $h_{18}$ | 0.0822                         | 33                 | $wd_{16}$| -0.0892                       | 30                 |
| $h_{19}$ | 0.0973                         | 28                 | $dt$    | -0.0604                       | 38                 |
| $h_{20}$ | 0.1138                         | 25                 | $hd_p$  | -0.0981                       | 27                 |
| $h_{21}$ | 0.1152                         | 24                 | $hd_r$  | -0.2986                       | 1                  |
| $h_{22}$ | 0.1249                         | 22                 | $nc$    | -0.0963                       | 29                 |
| $h_{23}$ | 0.1287                         | 20                 | $pr$    | -0.0130                       | 49                 |
| $h_{24}$ | 0.0663                         | 37                 | $sin\left(\frac{2\pi}{168}t\right)$ | -0.2468          | 2                  |
| $h_{25}$ | 0.0801                         | 34                 | $cos\left(\frac{2\pi}{168}t\right)$ |                      |                    |

The least important regressors in the model are the binary variables for daytime hours. Specifically, variables $h_i$, $i=9,10,\ldots,16$ are ranked between 28th to 45th with positive regression coefficients. Positivity of the regression coefficients is expected as electricity consumption significantly increases during the daytime. On the other hand, low ranks imply that there is not much difference between hours of a day and other effects are more important for the electricity demand estimation. Note that the sensitivity of estimated demand to regression covariates not only provides intuition into the regression model but also indicates the consumption routines and lifestyles of Turkish electricity consumers.

5. RESULTS AND DISCUSSION

In this study, the problem of forecasting hourly electricity demand under COVID-19 restrictions is analyzed using a sinusoidal regression model extended with a binary covariate for each type of restriction. The results indicate that the regression model has a forecast accuracy of 6% for the hourly electricity demand. The forecast accuracy is found to be low (27%) in the 1st week of the COVID-19 restrictions. Then the regression model learns the effect of restrictions on the electricity demand and forecast errors fluctuate around 6.07% for the rest of the COVID-19 restrictions.

In addition to this significant forecast accuracy, it is found that the error terms of our regression model have a symmetric distribution around 0 which is the same as the pre-COVID-19 period. This indicates that the binary variables are successful and efficient for grasping the effect of COVID-19 restrictions on electricity demand.

In addition to its predictive accuracy, the regression model is utilized to obtain an importance ranking between different independent covariates. It is found that the effect of the religious holiday, weekly cycles, and Monday effect create the largest impact on the hourly electricity demand of Turkey. The importance of religious holidays on the hourly electricity demand is easy.
to interpret as almost all commercial and industrial activities stop during those days, significantly lowering the electricity consumption of the country. The importance of Mondays stems from the fact that the 1st day of the week is the busiest one for all business activities driving the electricity consumption in all cities.

In addition to these important variables, it is found that binary variables for hours of the day between 9 am and 4 pm are the least important regressors in the equation although they are significant in the model. This is mainly due to the fact that electricity demand during work hours is mainly driven by day-level factors, such as religious holidays, day of a week, etc., and consumption during the work hours stay stable until the end of the daytime shift.

It is argued that findings are not only meaningful and important for national electricity demand forecasting, but they can also be used for forecasting regional electricity demand as they can easily be applied to different regions, e.g., city, town, neighborhood, etc., of the country. Therefore, these findings are also important and provide novel approaches for network management studies focusing on regions with smaller grid sizes.

6. CONCLUSION

Electricity is one of the main drivers of the modern economy. Due to limitations to store electricity, it needs to be consumed right after its generation. This makes the forecasting of electricity demand a very critical input for planning operations of utility companies and power plants. The electricity demand of a country comes from the household, industrial and commercial consumers, hence, is heavily affected by their consumption behaviors and routines.

Recently, COVID-19 has been the main determinant of behavior changes of consumers which leads to shifting consumption patterns in all countries. As governments apply various restrictions to control the spread of the infection in their countries, citizens adapted their lifestyles to this new normal. This adaptation shifts the electricity consumption pattern of all countries and it is important to analyze these pattern changes for forecasting electricity demand more efficiently.

In this study, a log-linear, sinusoidal regression model is proposed for forecasting the electricity demand of Turkey during the COVID-19 restrictions between March and June 2020. It is found that the model, including specific binary variables for three types of restrictions in Turkey, forecasts electricity demand with a MAPE of 6% during the COVID-19 period. It is found that the total demand falls 7.6%, 2.8%, and 13.6% due to Level 1, 2, and 3 restrictions, respectively, while the total decrease in demand becomes 24% in complete lockdown. In addition, it is shown that age-specific restrictions (Level 1) have a larger effect on weekdays whereas complete lockdown (Level 3) has a larger effect on weekends. It is found that the model is mostly successful in explaining our empirical observations presented in Section 3. Furthermore, standardized regression coefficients of the model are utilized to obtain an importance ranking among independent variables in the regression model. It is shown that religious holidays, weekly cycles, and the Monday effect are the most important observations for forecasting the hourly electricity demand in Turkey.

This study aims to explain changes in electricity demand due to COVID-19 restrictions in Turkey. Although most of our findings are limited to Turkey, the proposed approach to importance ranking among independent variables can be generalized to other countries’ electricity demand forecasting problems.

7. ACKNOWLEDGEMENT

This work was funded by TUBITAK 1001 Program, Grant Number 120M251.

REFERENCES

Abu-Rayash, A., Dincer, I. (2020), Analysis of the electricity demand trends amidst the COVID-19 coronavirus pandemic. Energy Research and Social Science, 68, 101682.
Alpaydin, E. (2020), Introduction to Machine Learning. United States: MIT Press.
AlRashidi, M.R., EL-Naggar, K.M. (2010), Long term electric load forecasting based on particle swarm optimization. Applied Energy, 87(1), 320-326.
Andersen, F.M., Larsen, H.V., Gaarderstrup, R.B. (2013), Long term forecasting of hourly electricity consumption in local areas in Denmark. Applied Energy, 110, 147-162.
Apadula, F., Bassini, A., Elli, A., Scapin, S. (2012), Relationships between meteorological variables and monthly electricity demand. Applied Energy, 98, 346-356.
Azadeh, A., Ghaderi, S.F., Tarverdian, S., Saberi, M. (2007), Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption. Applied Mathematics and Computation, 186(2), 1731-1741.
Bahmanyar, A., Estebarsi, A., Ernst, D. (2020), The impact of different COVID-19 containment measures on electricity consumption in Europe. Energy Research and Social Science, 68, 101683.
Bašta, M., Helman, K. (2013), Scale-specific importance of weather variables for explanation of variations of electricity consumption: The case of Prague, Czech Republic. Energy Economics, 40, 503-514.
Bruce, P., Bruce, A. (2017), Practical Statistics for Data Scientists. United States: O'Reilly Media.
Çakmaklı, C., Demiralp, S., Şebnem, K.O., Yeşiltas, S., Yıldırım, M.A. (2021), The Economic Case for Global Vaccinations: An Epidemiological Model with International Production Networks. United States: NBER Working Paper.
Chakchhoulk, Y., Paniati, P., Mili, L. (2011), Electric load forecasting based on statistical robust methods. IEEE Transactions on Power Systems, 26(3), 982-991.
Chakrabarti, A., Ghosh, J.K. (2011), AIC, BIC and Recent Advances in Model Selection. In: Philosophy of Statistics. ResearchGate: North-Holland. p583-605.
Chen, C., Zarazua de Rubens, G., Xu, X., Li, J. (2020), Coronavirus comes home? Energy use, home energy management, and the social-psychological factors of COVID-19. Energy Research and Social Science, 68, 101688.
Clements, A.E., Hurn, A.S., Li, Z. (2016), Forecasting day-ahead electricity load using a multiple equation time series approach. European Journal of Operational Research, 251(2), 522-530.
Conejo, A.J., Contreras, J., Espinola, R., Plazas, M.A. (2005), Forecasting electricity prices for a day-ahead pool-based electric energy market.
International Journal of Electrical Power and Energy Systems, 39(1), 48-55.

Wang, A.I., Ramsay, B. (1998), A neural network based estimator for electricity spot-pricing with particular reference to weekend and public holidays. Neurocomputing, 23(1-3), 47-57.

Wang, J., Li, L., Niu, D., Tan, Z. (2012), An annual load forecasting model based on support vector regression with differential evolution algorithm. Applied Energy, 94, 65-70.

Wang, Y., Wang, J., Zhao, G., Dong, Y. (2012), Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: A case study of China. Energy Policy, 48, 284-294.

Yucekaya, A., Bilge, A.H., Hekimoğlu, M., Agca Aktunc, E., Yukseltan, E., Kok, A. (2020), TUBITAK 1001-COVID19 project (120M251): Managing Electricity Supply, Generation Resources, and Demand During the COVID-19 Pandemic. Ankara, Turkey: TUBITAK.

Yukseltan, E., Kok, A., Yucekaya, A., Bilge, A.H., Agca Aktunc, E., Hekimoğlu, M. (2021), The Impact of COVID-19 Pandemic and Restrictions on the Electricity Demand and Daily Demand Curve in Turkey, Working Paper. Paris, France: International Energy Agency.

Yukseltan, E., Yucekaya, A., Bilge, A.H. (2017), Forecasting electricity demand for Turkey: Modeling periodic variations and demand segregation. Applied Energy, 193, 287-296.

Yukseltan, E., Yucekaya, A., Bilge, A.H. (2020), Hourly electricity demand forecasting using Fourier analysis with feedback. Energy Strategy Reviews, 31, 100524.

Zhang, B.L., Dong, Z.Y. (2001), An adaptive neural-wavelet model for short term load forecasting. Electric Power Systems Research, 59(2), 121-129.

Zhu, S., Wang, J., Zhao, W., Wang, J. (2011), A seasonal hybrid procedure for electricity demand forecasting in China. Applied Energy, 88(11), 3807-3815.