SETSCI Conference Indexing System, Volume 3 (2018), 907-910

ISAS 2018-Winter
2nd International Symposium on Innovative Approaches in Scientific Studies
November 30 – December 2, 2018, Samsun, Turkey

Estimation of Electricity Consumption of Turkey by using ARIMA, Grey Model and Linear Regression Analysis

Zeynep Ceylan[1*+], Hakan Öztürk[1] and Birol Elevli[1]

[1]*Department of Industrial Engineering, Ondokuz Mayıs University, Samsun, Turkey*
[*]*Corresponding author:* zeynep.dokumaci@omu.edu.tr
[+]*Speaker:* zeynep.dokumaci@omu.edu.tr
Presentation/Paper Type: Oral / Full Paper

Abstract – The successful estimation of future electricity consumption has a crucial importance in the energy planning. Because, in order to meet rising energy demand, policy makers should formulate electricity supply policies and make critical decisions and develop new strategies. This study aims to compare prediction capabilities of three different techniques in order to forecast electricity energy consumption of Turkey. These three techniques are; Autoregressive Integrated Moving Average (ARIMA), grey prediction model GM (1,1) and Linear Regression (LR) analysis. Yearly electricity consumption data of Turkey between 1970 and 2017 were obtained from the Turkish Electricity Transmission Company (TEIAS). The future electricity demand for a period of 6 years from 2018 to 2023 has been predicted. ARIMA (1,1,2) model showed best results in terms of highest value of coefficient of determination (R^2 = 99.9 %). The results of the study can help decision makers in planning future applications.

Keywords *– ARIMA, electricity consumption, estimation, grey model, linear regression analysis*

I. Introduction

Electrical energy is the most needed type of energy that does not pollute the environment and can be applied in many technologies. Due to these characteristics, electrical energy has become one of the most important indicators of the development and prosperity of a country. Thus, planning and forecasting of supply-demand, transmission-distribution and pricing is of great importance for the development of the electricity sector. Furthermore, estimation of electricity consumption has an important role in determining future energy policies.

In the literature, different methods are used to estimate electricity consumption in Turkey. They can be classified as statistical methods [1]-[4], meta-heuristic [5] and artificial intelligence [6]-[11] techniques.

The main objective of this study is to compare Autoregressive Integrated Moving Average (ARIMA), grey prediction GM (1,1) models and Linear Regression (LR) analysis each other to estimate future electricity demand of Turkey between 2018 and 2023.

II. Materials and Method

A. Dataset

The used data of Turkey between 1970 and 2017 were obtained from TEIAS [12]. The future electricity demand from 2018 to 2023 was predicted using Minitab 17.0 and MATLAB 2017a software.

The electricity consumption values in Turkey between the years 1970-2017 was plotted in Fig.1.

B. Data Analysis

The data analysis was performed using three methodologies; ARIMA, Grey Model GM (1,1) and LR analysis.

1. Autoregressive Integrated Moving Average (ARIMA)

Autoregressive Integrated Moving Average (ARIMA) method is used to analyze and estimate time series data. It is one of the most widely used linear econometric methods in the literature. The order of an ARIMA model is normally shown as in the form of (p, d, q). Here, "*p*" is the degree of the Autoregressive (AR) Model, "*q*" is the degree of the Moving Average (MA) Model, and "*d"* is the degree of difference. The common form of the ARIMA model is shown in Eq. (1):

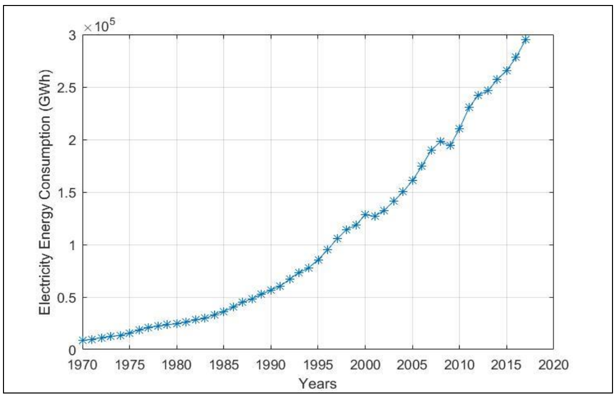

Fig. 1 Turkey net electricity consumption period of 1970-2017

$$y_t = \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q} \quad (1)$$

where y_t and ε_t are the actual value and random error at time period t, respectively. Random error, ε_t, is assumed to be independently and identically distributed with a mean of 0 and a constant variance of σ^2. ϕ_i and θ_j (i=1,2,…,p; j=0,1,2,…,q) are model parameters; where p and q are integers and generally called as the model orders [13].

ARIMA statistical analysis consists of four stages: time series stationary, parameter estimation, model evaluation and prediction [14]. In the first stage, the stability analysis of the variables is made since the ARIMA model needs stationary time series. In the second stage, ARIMA parameters are estimated by autocorrelation function (ACF) graph and partial autocorrelation (PACF) graph. In the third stage, different ARIMA models are checked to determine whether the predicted model is the best model among other possible ARIMA models. Finally, the prediction is made using best ARIMA model.

2. Linear Regression Analysis

Linear regression is a basic and widely used predictive analysis [15]. In this analysis, a mathematical model (known as regression model) is used to explore the relationship between two (simple regression) or more variables (multiple regression). In the regression analysis, the variables that affect are called (descriptive variables) independent variables and the variables affected are called the dependent variables. Regression models can be classified as linear or nonlinear models. Eq. 2-3 show linear and quadratic regression models with one independent variable, respectively.

$$Y = b_0 + b_1 X + \varepsilon \quad (2)$$

$$Y = b_0 + b_1 X + b_2 X^2 + \varepsilon \quad (3)$$

where ε is a random error term and ND (0, σ^2) means normally and independently distributed with mean 0, and constant variance σ^2. X is the independent variable vector, the Y is dependent variable vector, b_0 is the fixed value, and b_1, b_2 are the coefficients for the corresponding variables.

3. Grey Model GM (1,1)

Grey model GM (1,1) is a time series alternative forecasting model and for the systems whose structure is complicated, uncertain and incomplete information. The main advantage of this method is that it needs some simple mathematical operations with few data for forecasting models when compared with other time series techniques. GM (1,1) grey prediction model has been used commonly in various fields such as finance, physical control, engineering, economics, etc. and demonstrated satisfactory results [16]. The steps of the GM (1,1) method are shown as follow [1]:

Step 1. *The nonnegative original data sequence and Accumulated Generation Operation (AGO) time series with n samples are expressed in Equation 4 and 5, respectively:*

$$x^0 = \{x_1^0, x_2^0, x_3^0, \dots, x_n^0\}, (x_t^0; t = 1,2,3,\dots,n; n \geq 4) \quad (4)$$

$$x^1 = \{x_1^1, x_2^1, x_3^1, \dots, x_n^1\}, (x_t^1; t = 1,2,3,\dots,n; n \geq 4) \quad (5)$$

where,

$$x_k^1 = \left\{\sum_{t=1}^{k} x_t^0,\ \ t,k = 1,2,3,\dots,n\right\} \quad (6)$$

Step 2. *A first order grey differential equation is established to construct the GM (1,1) prediction model:*

$$x_t^0 + aZ_t^1 = b,\ \ t = 2,3,\dots \quad (7)$$

where,

$$Z_t^1 = \theta x_t^1 + (1-\theta)x_{t-1}^1 \quad t = 2,3,\dots,n \quad (8)$$

t is time point, θ is horizontal adjustment coefficient and takes value between (0-1). Accurate selection of θ provides the lowest estimation error. In Eq. (9), a and b are the development coefficient and control variable, respectively. These are two parameters of GM (1,1) model. [a,b]T can be estimated by using least mean square method coefficient;

$$A = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T\ B)^{-1} B^T Y_n \quad (9)$$

where,

$$B = \begin{bmatrix} -z^1 2 & 1 \\ -z^1 3 & 1 \\ \dots & 1 \\ -z^1 n & 1 \end{bmatrix}, Y = BA = \begin{bmatrix} x_2^0 \\ x_3^0 \\ \dots \\ x_n^0 \end{bmatrix} \quad (10)$$

The whitenization was showed in Eq. (11):

$$\frac{dx_t^1}{dt} + ax_t^1 = b \quad (11)$$

Step 3: *After calculating a and b parameters values, AGO grey prediction model is done:*

$$\hat{x}_{t+1}^1 = \left[x_1^0 - \frac{b}{a}\right]e^{-at} + \frac{b}{a}\ ,\ t = 0,1,2,\dots \quad (12)$$

where $\hat{x}$ denotes AGO prediction of x at time t point.

Step 4: *In order to reverse the forecasting value, Inverse Accumulated Generation Operation (IAGO) can be used. Because, the grey forecasting model is formulated using the data of AGO rather than original data.*

$$x_{t+1}^0 = x_{t+1}^1 - x_t^1 \quad (13)$$

Eq. (14) is formed using Eq. (13) and Eq. (12);

$$\hat{x}_{t+1}^0 = (1 - e^a)\left[x_1^0 - \frac{b}{a}\right]e^{-at}\ ,\ t = 0,1,2,\dots \quad (14)$$

C. Evaluation of Method Performances

In the literature, in order to determine the prediction accuracy of the models, different performance measures have been proposed. In this study, the performances of methods were evaluated according to coefficient of determination (R^2)

performance criteria. The related equation for calculation of R^2 is given in Eq. (15) as follows:

$$R^2 = 1 - \frac{\sum_{k=1}^{n}(t^k - \hat{y}^k)^2}{\sum_{k=1}^{n}(\bar{t}^k - \hat{y}^k)^2} \quad (15)$$

where "n" is the number of measured data points t_k, $\bar{t}$ and $\hat{y}_k$ actual target output, the mean of the actual output, and model output, respectively.

III. Electricity Consumption Forecasting

Results should be clear and concise. The most important features and trends in the results should be described but should not interpreted in detail.

A. Forecasting with Linear Regression Model

The annual electricity energy consumption dataset from 1970 to 2017 was used to create the linear regression model. As shown Fig. 2a, the linear regression model was fitted to the electricity energy consumption dataset using MATLAB 2017a. The R^2 calculated as 0.924. Then, residual plot was obtained after a regression equation is determined (Fig. 2b).

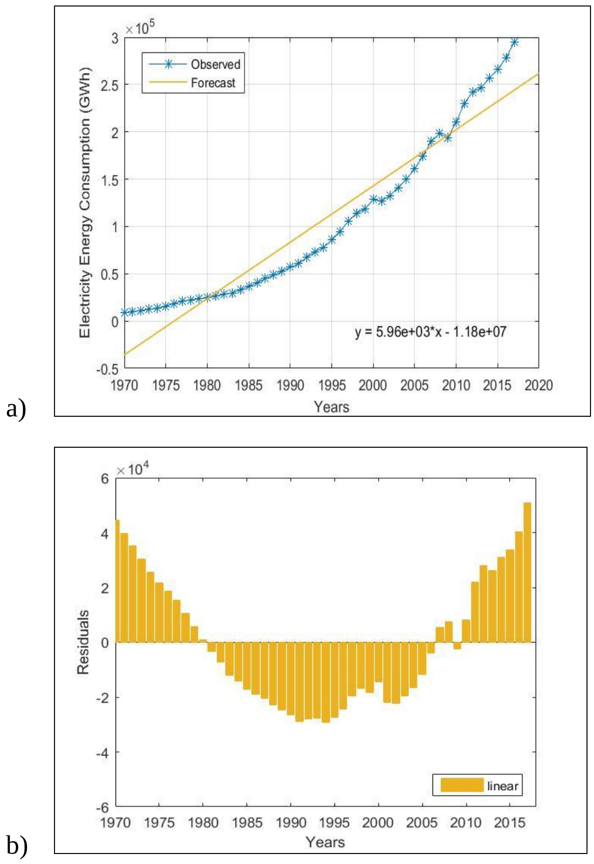

Fig. 2. a) Linear regression model and b) residual plot

B. Forecasting with ARIMA

In this study, the normality test of the dataset was made in Minitab 17.0 and results showed that the annual electricity consumption dataset was normally distributed and statistically significant (p <0.005). Autocorrelation (known as serial correlation) is the correlation of time series itself with lags. In order to understand the stability of the series, the ACF plot should be examined. As shown from ACF plot in Fig. 3, the series is not stationary, so the 1st order differences must be taken.

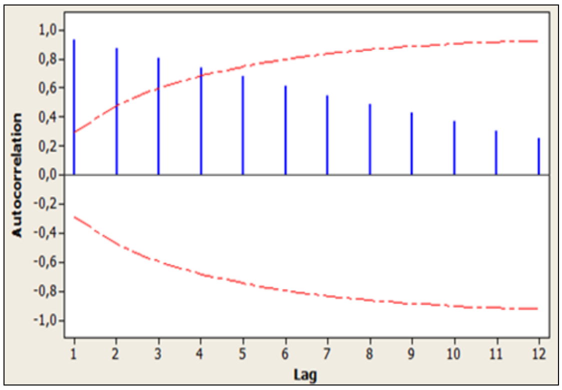

Fig. 3. ACF (Autocorrelation Function) plot

Then, ACF and PACF plots (Fig. 4a and 4b, respectively) were created after the first differential. According to these graphs, the series is stationary and thus no further differencing is required.

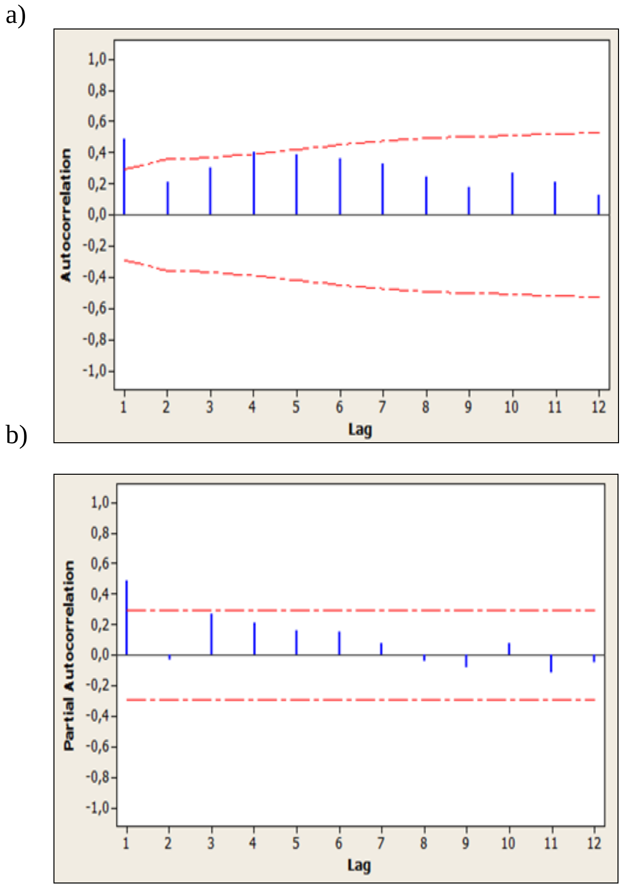

Fig. 4. Plots for a) ACF and b) PACF

Coefficient of determination (R^2) performance criteria was used to determine the best model among possible ARIMA models. Results showed that, the ARIMA (1,1,2) model is the best model with highest R^2 value ≅ 0.999.

C. Forecasting with Grey Model GM (1,1)

Because of prediction capability, grey model GM (1,1) was also applied on dataset using MATLAB 2017a. The development coefficient α and control variable b were obtained as *-0,0626* and *22741.65*, respectively. It is important to select the value of horizontal adjustment coefficient (θ) to ensure the lowest estimation error. Thus, the coefficient of determination (R^2) values were calculated using different θ values. The best value of θ to ensure the lowest estimation error. Thus, was found as 0.99 and the highest R^2 value ≅ 0.986 was obtained. Fig. 5 presents the forecasting results using the GM (1, 1) model.

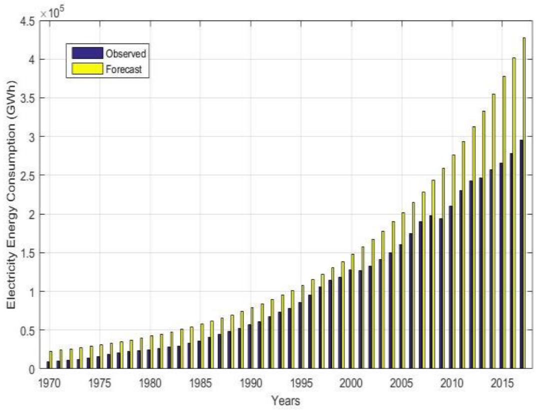

Fig. 5. Grey Model GM (1,1) prediction results and observed data between 1970-2017

D. Comparison of Methods

Table 1 presents the performance measure obtained using the ARIMA, GM (1, 1) model and the linear regression analysis. It is apparent that ARIMA (1,1,2) outperforms all other methods in terms of the error measure.

Table 1. Performances of different prediction models

Performance measure	**ARIMA (1,1,2)**	**GM (1,1)**	**LR**
R^2	0.999	0.986	0.924

E. Forecasting the future electricity demand using ARIMA (1,1,2) model

Because of higher forecasting accuracy than existing models, ARIMA (1,1,2) was applied to estimate the electricity consumption from 2018 to 2023 of Turkey. The predicted values were shown on Table 2.

Table 2. The predictive values of Turkey's electricity consumption from 2018 to 2023

Years	**Electricity Demand (GWh)**
2018	306484.89
2019	315939.89
2020	325674.30
2021	335555.07
2022	345468.49
2023	355318.25

It is seen that from Table 2, the total electricity consumption of Turkey will show a relatively stable rising trend in the following six years, and will reach approximately 355318.25 GWh by the year 2023.

IV. CONCLUSION

Electricity consumption prediction plays a significant role in economy development and electricity management. Because, accurate forecasting results can ensure effective implementation of electricity supply policies. They also help prevent economic losses because of insufficient electricity and reduce risks of economy and operating costs. Although, there are many forecasting techniques in the literature, the selection of the most suitable technique is importance. The main purpose of this study is to compare of ARIMA, grey prediction and linear regression models each other to forecast the future electricity consumption of Turkey. The results of study showed that, ARIMA (1,1,2) model estimates have more accurate results than other forecasting methods in the comparison of electricity consumption.

REFERENCES

[1] Hamzacebi, C., & Es, H. A. (2014). Forecasting the annual electricity consumption of Turkey using an optimized grey model. Energy, 70, 165-171.

[2] Kavaklioglu, K. (2011). Modeling and prediction of Turkey's electricity consumption using Support Vector Regression. Applied Energy, 88(1), 368-375.

[3] Kavaklioglu, K. (2014). Robust electricity consumption modeling of Turkey using singular value decomposition. International Journal of Electrical Power & Energy Systems, 54, 268-276.

[4] Kaytez, F., Taplamacioglu, M. C., Cam, E., & Hardalac, F. (2015). Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. International Journal of Electrical Power & Energy Systems, 67, 431-438.

[5] Ozturk, H. K., Ceylan, H., Canyurt, O. E., & Hepbasli, A. (2005). Electricity estimation using genetic algorithm approach: a case study of Turkey. Energy, 30(7), 1003-1012.

[6] Hamzaçebi, C. (2007). Forecasting of Turkey's net electricity energy consumption on sectoral bases. Energy policy, 35(3), 2009-2016.

[7] Kavaklioglu, K., Ceylan, H., Ozturk, H. K., & Canyurt, O. E. (2009). Modeling and prediction of Turkey's electricity consumption using artificial neural networks. Energy Conversion and Management, 50(11), 2719-2727.

[8] Sözen, A., Akçayol, M. A., & Arcaklioğlu, E. (2006). Forecasting net energy consumption using artificial neural network. Energy Sources, Part B, 1(2), 147-155.

[9] Sözen, A., & Arcaklioglu, E. (2007). Prediction of net energy consumption based on economic indicators (GNP and GDP) in Turkey. Energy policy, 35(10), 4981-4992.

[10] Sözen, A., Arcaklioğlu, E., & Özkaymak, M. (2005). Turkey's net energy consumption. Applied Energy, 81(2), 209-221.

[11] Kankal, M., Akpınar, A., Kömürcü, M. İ., & Özşahin, T. Ş. (2011). Modeling and forecasting of Turkey's energy consumption using socio-economic and demographic variables. Applied Energy, 88(5), 1927-1939.

[12] TEIAS electricity consumption in Turkey between 1970 and 2017, (2017). Retrieved from www.teias.gov.tr/sites/default/files/2017-06/Önsöz.xls.

[13] Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 159-175.

[14] Gujarati, D. N. (2004). Basic Econometrics. (4 th edtn) The McGraw-Hill Companies.

[15] Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introduction to linear regression analysis (Vol. 821). John Wiley & Sons.

[16] Tseng, F. M., Yu, H. C., & Tzeng, G. H. (2001). Applied hybrid grey model to forecast seasonal time series. Technological Forecasting and Social Change, 67(2-3), 291-302.