Energy Policy 90 (2016) 92–101

Contents lists available at ScienceDirect

Energy Policy

journal homepage: www.elsevier.com/locate/enpol

Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of Turkey

M. Erdem Günay

Department of Energy Systems Engineering, Istanbul Bilgi University, 34060 Eyup, Istanbul, Turkey

HIGHLIGHTS

- Electricity demand of Turkey increased from 15.6 to 246.4 TW h in 1975–2013 period.
- Population, GDP per capita, inflation and average summer temperature influence demand.
- Future values of descriptor variables can be predicted by time series ANN models.
- ANN model simulated by the predicted values of descriptors can forecast the demand.
- Demand is forecasted to be doubled reaching about 460 TW h in the year 2028.

ARTICLE INFO

Article history:
Received 13 August 2015
Received in revised form
1 December 2015
Accepted 17 December 2015
Available online 22 December 2015

Keywords:
Artificial neural networks
Time series
Electricity demand forecasting
Population
Economic indicators
Average ambient temperature

ABSTRACT

In this work, the annual gross electricity demand of Turkey was modeled by multiple linear regression and artificial neural networks as a function population, gross domestic product per capita, inflation percentage, unemployment percentage, average summer temperature and average winter temperature. Among these, the unemployment percentage and the average winter temperature were found to be insignificant to determine the demand for the years between 1975 and 2013. Next, the future values of the statistically significant variables were predicted by time series ANN models, and these were simulated in a multilayer perceptron ANN model to forecast the future annual electricity demand. The results were validated with a very high accuracy for the years that the electricity demand was known (2007–2013), and they were also superior to the official predictions (done by Ministry of Energy and Natural Resources of Turkey). The model was then used to forecast the annual gross electricity demand for the future years, and it was found that, the demand will be doubled reaching about 460 TW h in the year 2028. Finally, it was concluded that the approach applied in this work can easily be implemented for other countries to make accurate predictions for the future.

1. Introduction

The world electricity demand has increased extremely in the recent years as the world has become more populated and as the electricity consuming devices and appliances have become more common in the daily lives of people. It is vital for a country to be able to supply the electricity exactly equal to the demand. If the electricity generation capacity of a country is lower than the gross demand, electricity dependent industry is affected negatively and blackouts occur; on the other hand, a higher electricity generation capacity than the demand leads for the power plants to work with idle capacity, which is a waste in economic resources. Hence, accurate prediction of the electricity demand for the future is very important to correctly plan and develop new electricity generation investments for maintaining the electricity demand–supply balance.

In order to forecast the electricity demand with a good precision, one must correctly determine the variables which may influence the electricity demand in that country. Population is one of the key factors that are highly correlated with the electricity demand (more people consume more electricity). However, population alone is not sufficient to explain the changes in the electricity demand through years.

It is also quite common to consider some economic indicators in correlation with the electricity demand (Askarzadeh, 2014); one factor that can be used for this purpose is the gross domestic product (GDP) per capita, which is an indicator of the wealth of the

E-mail address: erdem.gunay@bilgi.edu.tr

http://dx.doi.org/10.1016/j.enpol.2015.12.019
0301-4215/

people living in a country (Kucukali and Baris, 2010). As the GDP per capita increases the living standards of people get better and their lifestyles become more dependent on energy consuming devices and appliances. In addition to GDP per capita, employment and the inflation rates are two other economic factors that may affect the electricity demand (Zahedi et al., 2013).

Another factor that has a possible effect on electricity consumption is the electricity price (Inglesi, 2010; Kialashaki and Reisel, 2014; Nawaz et al., 2014). If there are available alternatives to electricity consumed in a country, the electricity demand is expected to be price elastic (an increase in price causes significant decline in the demand); otherwise, the demand is expected to be price inelastic (an increase in price causes only minor decline in the demand) (Nawaz et al., 2014).

The consumption of electricity may also depend on climatic conditions such as the average summer and winter temperatures. The hotter it is in the summer, more electricity is consumed for residential cooling, refrigeration and irrigation; on the other hand, the colder it is in the winter more electricity is consumed due to the electricity based heating of the residents (De Felice et al., 2013; Ekonomou, 2010).

For many years, many different forecasting tools based on data mining were applied to predict the future electricity or energy demand. Multiple linear regression (MLR) (Bianco et al., 2009, 2013; Ekonomou, 2010; Geem and Roper, 2009; Kialashaki and Reisel, 2014), methodologies based on fuzzy logic (Azadeh et al., 2010; Kucukali and Baris, 2010; Zahedi et al., 2013), autoregressive forecasting methods (García-Ascanio and Maté, 2010; Kheirkhah et al., 2013; Nawaz et al., 2014), support vector regression models (Ekonomou, 2010; Kavaklioglu, 2011) and artificial neural network (ANN) based models (Ekonomou, 2010; Geem and Roper, 2009; Kankal et al., 2011; Kavaklioglu et al., 2009; Kheirkhah et al., 2013; Kialashaki and Reisel, 2014; Pao, 2009; Zahedi et al., 2013) have been widely applied for this purpose. Various papers published in the last five years to forecast the electricity or energy demand for different countries with a variety of methods, together with the input variables considered and the forecasting time period are reviewed and summarized in the Supplementary Material. In some of these studies the demand was modeled as a function of the past data of the demand (time series approach) to forecast the possible demand in the future (Azadeh et al., 2010; Bianco et al., 2013, 2010; García-Ascanio and Maté, 2010; Hamzacebi and Es, 2014; Kheirkhah et al., 2013; Pao, 2009). Although, accurate results can be achieved when this method is employed, the factors leading to an increase or decrease in the demand cannot be analyzed this way. One other type of approach commonly implemented is to adopt the demand as a function of some descriptor variables (Askarzadeh, 2014; Behrang et al., 2011; Ekonomou, 2010; Geem and Roper, 2009; Ghanbari et al., 2013; Inglesi, 2010; Kavaklioglu, 2011; Kavaklioglu et al., 2009; Kialashaki and Reisel, 2014; Kıran et al., 2012; Kucukali and Baris, 2010; Nawaz et al., 2014; Zahedi et al., 2013). However, the problem with this approach is that the future values of the descriptor variables are uncertain (i.e. the population in the future is unknown); therefore, the demand is usually forecasted using the predicted values of the descriptor variables based on different scenarios that are in some cases too arbitrary.

In this study, both approaches are combined in a way that, the future values of the descriptor variables (i.e. population) were predicted using the past values of these variables by time series ANN models, and the future values of the electricity demand were forecasted using the predicted values of the descriptor variables by multilayer perceptron ANN models. Inspired from the biological nervous systems, ANNs mimic the learning that happens in humans; they have a great ability to approximate any nonlinear relationship that exists between a set of input variables and an output variable (Günay and Yildirim, 2011; Kialashaki and Reisel, 2014). Moreover, they are available as a toolbox in many of the programming environments (i.e. MATLAB, Weka, etc.); and thus, they are easy to implement for any purpose. They are also proven to be very successful for time series modeling in which the future values of a variable is determined using its past values (Kheirkhah et al., 2013).

In order to test the success of the approach, Turkey was chosen a case study since the rise in the energy demand of Turkey through years is even sharper than most of the other countries; such that, the annual gross electricity demand of Turkey increased from 15.6 TW h in the year 1975 to 246.4 TW h in the year 2013 (multiplied by a factor of 15.7 in this time interval) (TEIAS, 2013). Fossil fuel (petroleum, coal and natural gas) based electricity generation has been the primary way to meet this enormous increase of electricity demand. However, due to the fact that the fossil fuel resources are too limited in Turkey, they are mostly imported from other countries; consequently, most of the electricity demand of Turkey is supplied by foreign sources, which makes the accurate prediction of the future electricity demand even more crucial.

In this work, a database (covering the years between 1975 and 2013) was constructed including population, GDP per capita, inflation percentage, unemployment percentage, average summer temperature and average winter temperature to forecast the annual gross electricity demand of Turkey. First, the future values of the statistically significant descriptor variables were predicted from their historical values using time series ANN models, and possible future trends of these variables were analyzed. Then, the results were validated for the years that the electricity demand was known (2007–2013). Finally, the future electricity demand for Turkey for the years between 2014 and 2028 were forecasted by multilayer perceptron ANN models using the predicted values of the descriptor variables, and the results were also compared with the official predictions (done by Ministry of Energy and Natural Resources of Turkey).

2. Computational details

2.1. Data

Historical data for population, GDP per capita, inflation percentage, unemployment percentage, average summer temperature and average winter temperature (input variables), and the electricity demand (output variable) were collected from different sources for the years between 1975 and 2013. The gross electricity demand data (including the losses related to electricity transmission and distribution) was taken from the 2013 dated report of Turkish Electricity Transmission Company (TEIAS, 2013) while the past population data was extracted from Turkish Statistical Institute (TurkStat, 2013). Among the economic indicators, GDP per capita (based on purchasing power parity in international dollars) was calculated from the statistical data supplied by Organization for Economic Cooperation and Development (OECD, 2015), historical inflation percentage (based upon the consumer price index) was taken from World Wide Inflation Database (WWID, 2014), and the unemployment percentage was taken from International Monetary Fund World Economic Outlook Database (IMF, 2013). Finally, the average summer (June, July and August) and winter (December, January and February) temperatures were mined out from the report of Turkish State Meteorological Service based on the measurements made on 130 meteorological stations all over the country (MGM, 2014). The missing data points in the database were determined by linear interpolation or extrapolation using the closest values.

It should be noted that, the price of electricity was not used as a descriptor input variable in this work since the electricity demand

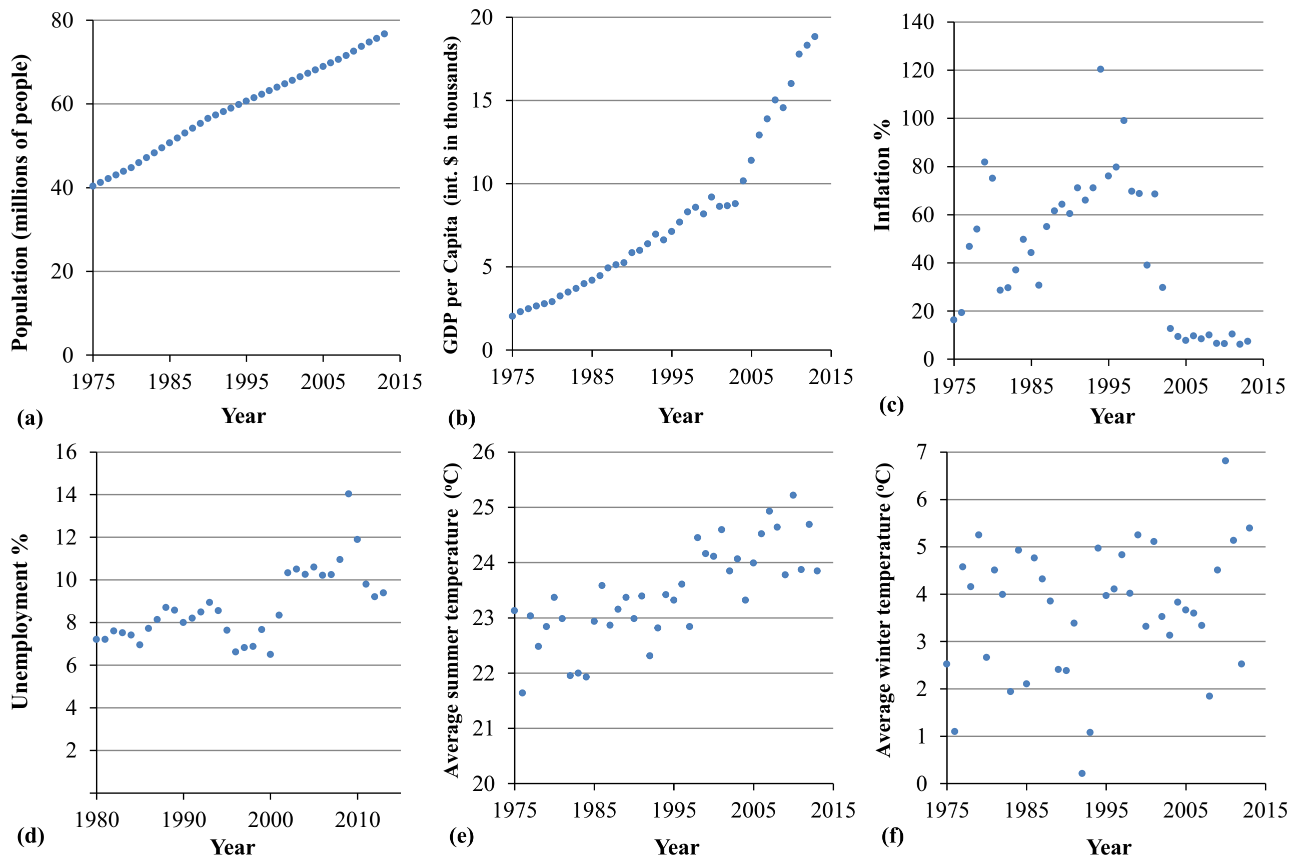

Fig. 1. Data used in this study: (a) population, (b) GDP per capita, (c) inflation percentage, (d) unemployment percentage, (e) average summer temperature, and (f) average winter temperature.

in Turkey is price inelastic (the demand is not significantly affected by increase or decrease in price), which is because of lack of alternatives for electricity as discussed elsewhere (Erdogdu, 2007). A similar result was also reported for Italy (Bianco et al., 2009) and Pakistan (Nawaz et al., 2014) while an opposite situation was reported for South Africa (Inglesi, 2010).

The changes of all the variables through years are given in Fig. 1. The figure indicates that the population of Turkey has reached 76.7 million in the year 2013 while it was only 40.3 million in the year 1975, which means that it has been multiplied by a factor of 1.9 in this time interval. The GDP per capita of Turkey is in an increasing trend in the recent years; such that, it has been multiplied by a factor of 9.3 from the year 1975 to 2013 reaching above $18.000. For the unemployment rate, the lowest rate was 6.5% in the year 2000 while the highest was 14% in the year 2009. Although the unemployment rate was found to fluctuate between these percentages, no particular trend was observed for this variable in this time period. Inflation had been in the 15–20% interval in the early 1970s while it reached a peak value around 120% in mid 1990s, and has become almost stable in the 6–10% interval in the last ten years for Turkey. The average summer temperature has increased by almost 2 °C from about 22.5 °C to 24.5 °C in the 1975–2013 period; in contrast, no particular trend was observed for the average winter temperature.

2.2. Methods

In this work, all the computational models were created by writing computer codes in MATLAB 8.2 (R2013b) environment. Population, GDP per capita, inflation percentage, unemployment percentage, average summer temperature and average winter temperature were used as the descriptor variables to determine the electricity demand. First, a multiple linear regression (MLR) model applied on the entire available data was used to determine whether a particular descriptor variable was significant or not. For this purpose, the coefficients of the multiple regression, the standard errors of the coefficients, the corresponding t and two tailed p values were calculated using the "regstats" function of MATLAB, and the variables with a two tailed value smaller than 0.05 were considered as statistically significant with a 95% confidence interval (Larose, 2006; Walpole et al., 2012). The computational details of the MLR model is given in the Supplementary material.

Artificial neural network (ANN) models were used for two different purposes; for the ANN-Type I (multilayer perceptron), the electricity demand was modeled as a function of the descriptor variables as shown in Fig. 2 (only the statistically significant variables were used as input variables in the final model). Next, in order to simulate this ANN model for the future years, the future values of the statistically significant descriptor variables (i.e. population) was determined by using time series ANN models (ANN-Type II), in which one particular variable in the year "t" was modeled as a function of its values in the past years. The conceptual ANN model of this type is shown in Fig. 3 (for the population as an example).

ANN models were trained by the "trainbr" function of MATLAB, which is a network training function that updates weight and bias values according to Levenberg–Marquardt optimization based on gradient descent and Gauss–Newton iteration (Wilamowski and Chen, 1999). This training function uses Bayesian regularization and was reported to be very suitable for obtaining a high generalization accuracy (extrapolating the results for the uncertain

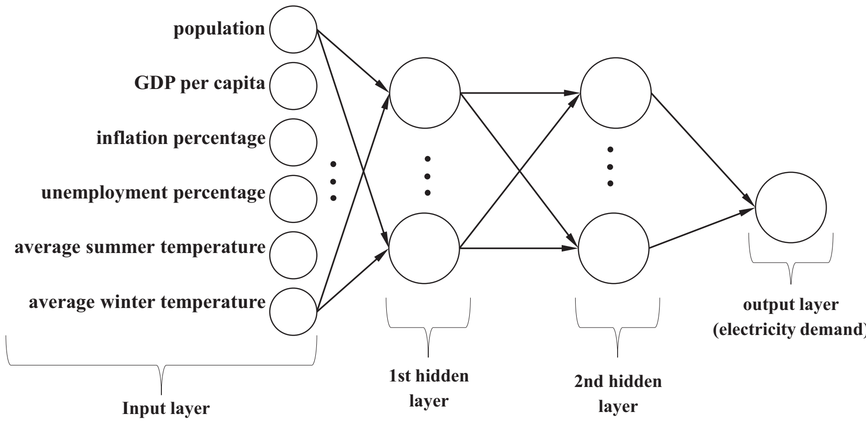

Fig. 2. ANN-Type 1 (multilayer perceptron) to determine the electricity demand from descriptor variables.

conditions) (Pan et al., 2013). Two hidden layers were used in the ANN architectures to capture any possible nonlinearity in the data more effectively. For each different model, various numbers of neurons with the combination of hyperbolic tangent sigmoid and identity transfer functions were tried in the layers of the ANNs, and the sum of square error (SSE) was employed as the network performance function to determine the most accurate model. For the time series ANN models, "narnet" function of MATLAB was implemented; first, the networks were trained in open loop mode (Fig. 3), where the value of a particular variable in the year "*t*" was modeled as a function of its literature values in the past years. Then, for determining the forecasting ability of the networks, a closed loop simulation was implemented; such that, the first value of the variable in the year "*t*" was determined as a function of its past values, then, the predicted value was fed back to the network to determine another value corresponding to one step in the future (Fig. 4) (Kheirkhah et al., 2013). Repeating this procedure year by year, it was possible to get into the future years, where no literature data was available; hence, the corresponding future values of the variable could be determined by its past predicted values.

The relative significance of each variable was found by applying the "change of root mean square error (RMSE) method", which was reported to be quite successful for this purpose (Günay and Yildirim, 2011). First, an ANN model was trained with all the input variables and the resulting RMSE was recorded as the original value. Then, one variable was removed from the inputs, the network was trained again, and the new RMSE was recorded. This procedure was repeated for all the variables and the differences between the original RMSE of the network and the RMSE values found in the absence of the variables were used as an indicator of the significance of each variable. The higher the RMSE became in the absence of a particular variable, the higher its relative significance was for determining the output variable (Günay and Yildirim, 2011).

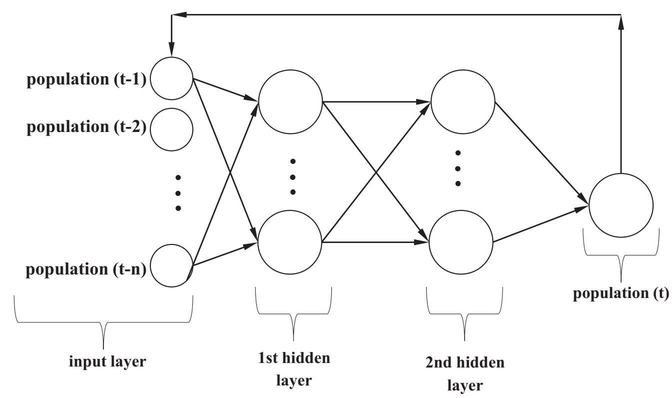

Fig. 4. ANN-Type 2 (time series) to determine the population in the future from its past values (in closed loop mode).

3. Results and discussion

This section is divided in three parts; first, the statistically significant variables were determined by applying a MLR model to the entire database consisting of six descriptor variables (population GDP per capita, inflation percentage, unemployment

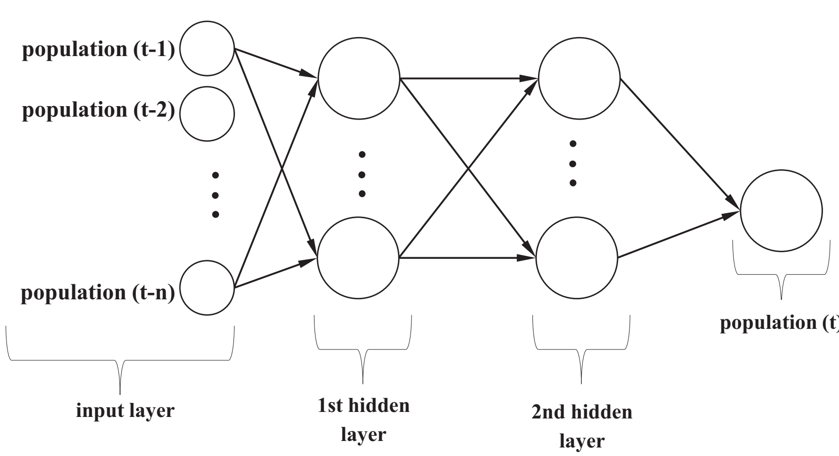

Fig. 3. ANN-Type 2 (time series) to determine the population in the future from its past values (in open loop mode).

percentage, average summer temperature and average winter temperature). Then, an ANN model was constructed to find the relative significances of the statistically important variables by applying the "change of RMSE method".

In the second part, the data between 2007 and 2013 were excluded from the dataset, and the values of the statistically significant descriptor variables in these years were predicted by time series ANN models trained by the data between 1975 and 2006. Then, a MLR and an ANN model were constructed to simulate the predicted values of these variables to forecast the demand between 2007 and 2013. The results were compared with the real electricity demand as well as the official predictions done by the done by Ministry of Energy and Natural Resources of Turkey. This way, the success and the reliability of the forecasting approach applied in this work was validated.

In the third part, the data of 1975–2013 was used to forecast the electricity demand between the years 2014 and 2028. The procedure applied was similar to the one employed in the second part. First, the future values of the descriptor variables were found by time series ANN models, and then, another ANN model with the descriptor variables as the inputs was used to forecast the future electricity demand.

3.1. Determining the statistically significant variables for electricity demand

The data between the years 1975 and 2013 was used to construct a MLR in the form as shown in Eq. (1), where electricity demand is the dependent variable (with an estimated value of $\hat{y}$) while population in millions (x_1), GDP per capita in thousand dollars (x_2), inflation percentage (x_3), unemployment percentage (x_4), average summer temperature in Celsius (x_5) and average winter temperature in Celsius (x_6) were the independent (descriptor) variables.

$$\hat{y} = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \beta_3 \cdot x_3 + \beta_4 \cdot x_4 + \beta_5 \cdot x_5 + \beta_6 \cdot x_6 \quad (1)$$

The final equation for the MLR model is given in Eq. (2) while the regression coefficient (β) of each variable, the standard errors, the t values and the two tailed p values of the coefficients are summarized in Table 1. The statistical significance of each variable was determined by inspecting on their p values; those variables with a two tailed p value lower than 0.05 was accepted as statistically significant (with a 95% confidence interval) (Larose, 2006). It was found that unemployment percentage (x_4) and average winter temperature (x_5) were not statistically significant according to this criterion. This was an expected result considering that no particular trends were observed for these two variables through the years as shown in Fig. 1. It should be noted here that the significant variables affecting the demand may differ for different countries, in other words, a variable in correlation with the electricity demand in a country may not have any significant correlation in another country. Hence, it is recommended to determine the statistically significant variables affecting the electricity demand of a particular country before performing future predictions.

Table 1
Parameters of MLR model for estimating annual electricity demand (full model).

Variables	β	**SE(β)**	***t* Value**	***p* Value**
Constant	−188.1	49.58	−3.794	0.00062
Population (millions of people)	1.830	0.440	4.154	0.00023
GDP per capita (1000$)	9.604	0.968	9.924	0.00000
Inflation %	−0.195	0.063	−3.088	0.00415
Unemployment %	−0.736	1.229	−0.599	0.55343
Average summer temperature (°C)	4.845	2.360	2.053	0.04833
Average winter temperature (°C)	1.076	1.001	1.075	0.29034

Table 2
Parameters of MLR model for estimating annual electricity demand (reduced model).

Variables	β	SE(β)	t Value	p Value
Constant	−200.9	47.97	−4.19	0.00019
Population (millions of people)	1.660	0.402	4.13	0.00022
GDP per capita (1000$)	9.860	0.936	10.53	0.00000
Inflation %	−0.163	0.054	−3.04	0.00448
Average summer temperature (°C)	5.573	2.258	2.47	0.01878

$$\hat{y} = -188.1 + 1.830 \cdot x_1 + 9.604 \cdot x_2 - 0.195.\ x_3 - 0.736 \cdot x_4 + 4.845 \cdot x_5 + 1.076 \cdot x_6 \quad (2)$$

Next, a reduced model (Eq. (3)) excluding the insignificant variables were constructed, using population in millions (x_1), GDP per capita in thousand dollars (x_2), inflation percentage (x_3), average summer temperature in Celcius (x_5) as the independent variables. As it is shown in Table 2, the two tailed p values of all the variables were found to be lower than 0.05 meaning that all of them were statistically significant. After that, the coefficients of MLR model were examined and the results were found to be consistent with what was expected; such that, population, GDP per capita, and average summer temperature were directly proportional with the electricity demand (positive coefficients) while inflation percentage was inversely proportional with the electricity demand (negative coefficient).

$$\hat{y} = -200.9 + 1.660 \cdot x_1 + 9.860 \cdot x_2 - 0.163 \cdot x_3 + 5.573 \cdot x_5 \quad (3)$$

Then, Pearson correlation coefficients (R) were analyzed to determine the degree of linear relationship between the independent variables and the electricity demand for the reduced model, and between the independent variables themselves (Table 3). It was found that population, GDP per capita and the average summer temperature are highly positively correlated with the demand ($R > 0.7$) while inflation is mildly negatively correlated with the demand ($-0.7 < R < -0.33$) (Larose, 2006). As for the correlation between the dependent variables, the highest degree of correlation was found between population and GDP per capita (R=0.92). Then, the corresponding variance inflationary factor (VIF) of each independent variable was calculated, and the values were found as 9.3, 9.8, 1.7, 2.5 for population, GDP per capita, inflation and the average summer temperature respectively. As a rule of thumb, in order to avoid multicollinearity the value of the VIF factor is recommended to be less than 10 (although this limit may be higher for some cases) as argued elsewhere (O'brien, 2007). In the current MLR model, the VIF values of inflation and the average summer temperature were found to be well below 10 while the VIF values for population and GDP per capita were slightly smaller than 10. Another rule of thumb that can be used in

Table 3
Matrix of correlation coefficients for the variables considered.

	Population	GDP per capita	Inflation	Average summer temperature	Annual electricity demand
Population	1.00	0.92	−0.32	0.77	0.94
GDP per capita	0.92	1.00	−0.51	0.73	0.99
Inflation	−0.32	−0.51	1.00	−0.31	−0.52
Average summer temperature	0.77	0.73	−0.31	1.00	0.77
Annual electricity demand	0.94	0.99	−0.52	0.77	1.00

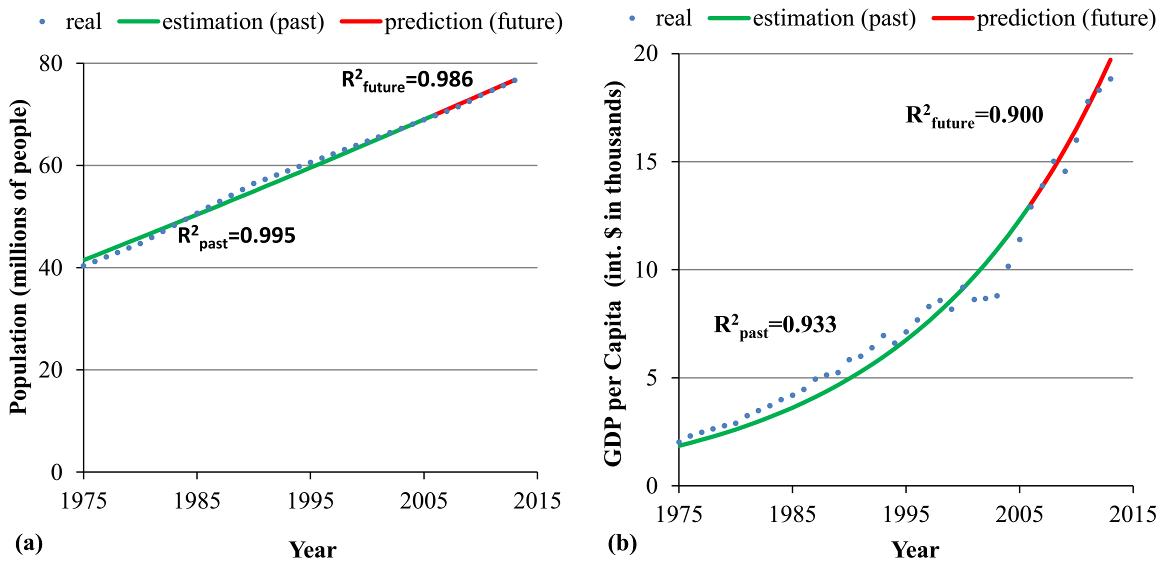

Fig. 5. Forecasting (a) population and (b) GDP per capita for the years 2007–2013 using time series ANN models.

most cases is to ignore multicollinearity if the t values of all the independent variables are higher than 2 or less than -2 (Kennedy, 2008); and in the current case all the variables also obeyed to this rule as shown in Table 2; hence, all the variables were held in the MLR model.

Although MLR model determined the statistically significant variables, it did not give the relative significances of the variables compared to each other. Hence, for this purpose, a multilayer perceptron ANN model was applied to the entire dataset trained by population, GDP per capita, inflation percentage and the average summer temperature. After trying different neural network topologies, the optimal neural network topology was found as 4-2-2-1 (4 inputs, 2 neurons each in the first and second hidden layers and 1 output; with the activation functions of hyperbolic tangent sigmoid function for the input layer, identity function for the first and second hidden layers), and the "change of RMSE method" was employed to determine the relative significances of the variables. It was found that population and GDP per capita had almost the same relative significances on the electricity demand, and their significances were much higher than those of inflation percentage and average summer temperature (Table 4). However, all the variables had a degree of contribution to accurately estimate the electricity demand; hence, the forecasting models in the rest of this work were constructed by using these four variables.

3.2. *Forecasting the electricity demand for the years between 2007 and 2013*

The main goal of this work is to forecast the electricity demand for the future years; however, before that, forecasting abilities of both MLR and ANN models were compared for the years whose data is available with a procedure very similar to the one that would have been applied for future prediction. This way, the success of the current approach was tested; if the model was able to predict the demand in this time interval with a good accuracy, it could be used for forecasting the future demand with a high reliability. Thus, the data corresponding to the years between 2007 and 2013 were excluded from the dataset, and the statistically significant descriptor variables (population, GDP per capita, inflation percentage and the average summer temperature) were predicted for the years 2007–2013 from their past data by using time series ANN models. Population in the year "t" was modeled as a function of the population in the years "$t-1$" to "$t-5$" (Eq. (4)), likewise GDP per capita, inflation percentage and average summer temperature were modeled as functions of the past values of themselves as shown in Eqs. (5)–(7). It should be noted that time series modeling of the average summer temperature was more difficult than the others due to its fluctuations over years; hence, the temperature in each year was modeled as a function of the last ten years to get a successful trend.

$$\text{population}_t = f(\text{population}_{t-1}, \text{population}_{t-2}, \text{population}_{t-3}, \text{population}_{t-4}, \text{population}_{t-5}) \quad (4)$$

$$\text{GDP}pc_t = f(\text{GDP}pc_{t-1}, \text{GDP}pc_{t-2}, \text{GDP}pc_{t-3}) \quad (5)$$

$$\text{inflation}_t = f(\text{inflation}_{t-1}, \text{inflation}_{t-2}) \quad (6)$$

$$\text{summer}T_t = f(\text{summer}T_{t-1}, \text{summer}T_{t-2}, \text{summer}T_{t-3}, \ldots \text{summer}T_{t-10}) \quad (7)$$

Fig. 5 shows the closed loop performance of the time series ANN model for forecasting the population and GDP per capita. The ANN model successfully estimated the values corresponding to the past (1975–2006) and predicted the values of the future (2007–2013) for both the population and the GDP per capita as indicated by the high coefficient of determination (R^2) values. In addition to these, a decrease in inflation (especially in the last years) and an increase in average summer temperature were observed through years; these trends (given in the Supplementary material) were also successfully captured by the time series ANN models.

Table 4
Relative significance of input variables found by the ANN model.

Variables	RMSE found in the absence of the variable	RMSE difference	Relative input significance%
Population (millions of people)	3.44	1.87	44.0
GDP per capita (1000 $)	3.54	1.98	46.4
Inflation %	1.90	0.34	7.9
Average summer temperature (°C)	1.78	0.21	5.0
Original RMSE	1.56		

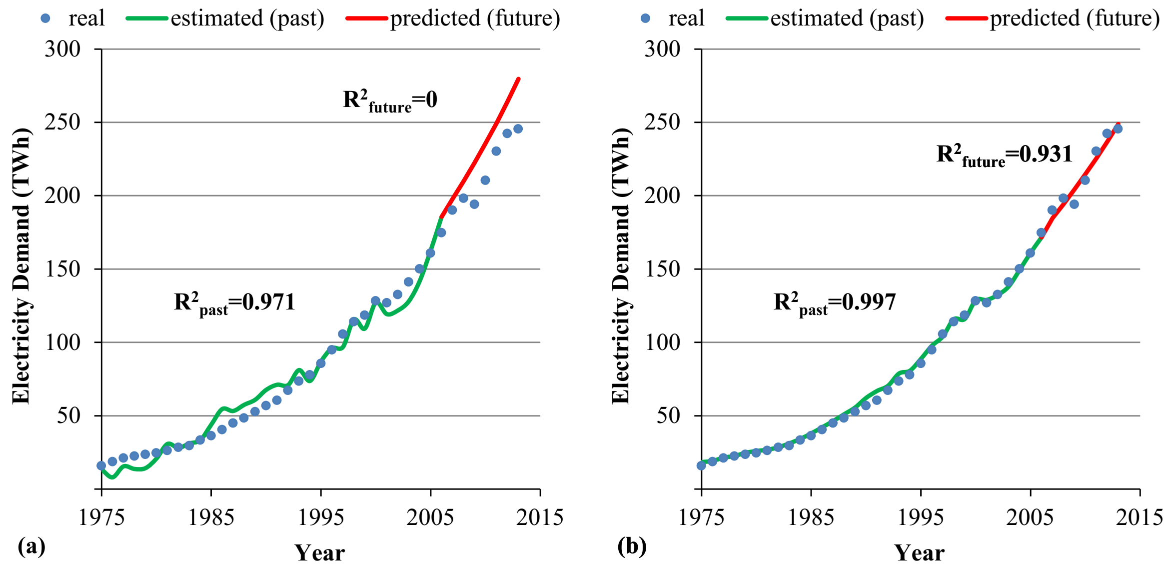

Fig. 6. Forecasting the electricity demand for the years between 2007 and 2013 by (a) MLR model and (b) ANN model.

Then, the predicted values of the population, GDP per capita, inflation percentage and the average summer temperature for the years 2007–2013 were simulated by MLR and multilayer perceptron ANN models to forecast the electricity demand for these years. It should be noted that the models constructed in this part were trained by the data corresponding to the 1975–2006 period, and no real data was introduced to the models for the simulated years. The predictions of the MLR and ANN model were compared with the real values and with each other in Fig. 6, which indicates that the forecasting ability of ANN was superior to MLR. Although, the MLR model had good accuracy for estimating the demand in the past, it failed to forecast the demand in the future (predicted the results were higher than the real values). As a result, the ANN model had a higher ability to forecast the future data; hence, the rest of the work was done by ANN modeling.

The results obtained by the ANN model employed in this work were also compared with the official predictions done by Ministry of Energy and Natural Resources of Turkey (TEIAS, 2007). The forecasting of the gross annual electricity demand by the official model was done by using a simulation model MAED (Model for Analysis of Energy Demand) based on high demand and low demand scenarios considering the data related to social, economic and demographical structure of the country (Hamzaçebi, 2007). However, the results related to both scenarios were found to be higher than the real values as indicated by Fig. 7; which was an expected result as argued elsewhere (Kavaklioglu, 2011; Kucukali and Baris, 2010). Even though the low demand scenario predicted the electricity demand with an R^2 value of 0.706, the predictions of the ANN model applied in this study was found to be superior with an R^2 of value of 0.931.

The accuracy of the method applied in this work was also compared with different methods performed by different researchers as summarized in Table 5. Forecasted gross electricity demands found by ARIMA modeling (Erdogdu, 2007) and gray prediction with rolling mechanism (Akay and Atak, 2007) resulted RMSE values 22.5 and 25.7 respectively; while the method applied in this work had a RMSE value of 5.7.

3.3. Forecasting the electricity demand for the years between 2014 and 2028

In this part of the work, first, the future values of population, GDP per capita, inflation percentage and the average summer temperature in the years between 2014 and 2028 were predicted by time series ANN models with the same neural network parameters that were used in Section 3.2. The change of population and GDP per capita through years is shown in Fig. 8 while the values related to inflation percentage and average summer temperature can be found in the Supplementary material. The figure indicates that the population is expected to reach over 90 million in the year 2028, while the GDP per capita is predicted to reach over $40,000 in the same year. In addition to these, the inflation is expected to decrease slightly, and the average summer temperature is forecasted to increase by about 1 °C in this 15 year period.

Finally, the data from the year 1975 to 2013 were used to train a multilayer perceptron ANN model and the predicted values of population, GDP per capita, inflation percentage and the average summer temperature in the years between 2014 and 2028 were used to simulate the ANN model to forecast the electricity demand

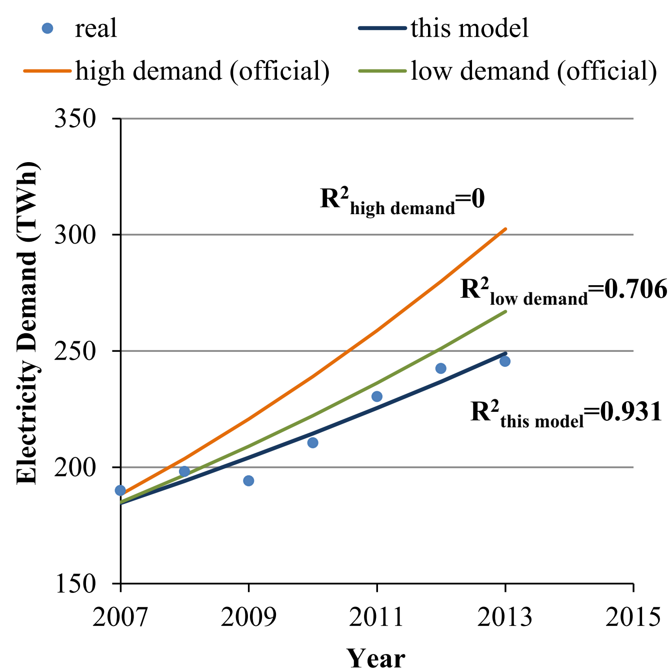

Fig. 7. Comparison of predictions by the ANN model applied in this work with the official predictions between the years 2007 and 2013.

Table 5
Comparison of predictions by different methods between the years 2007 and 2013.

Year	Real annual gross electricity demand (TW h)	Predicted Annual gross electricity demand (TW h)				
		Current model	Low demand scenario (TEIAS, 2007)	High demand scenario (TEIAS, 2007)	ARIMA modeling (Erdogdu, 2007)[a]	Gray prediction (Akay and Atak, 2007)[a]
2007	190.0	184.6	185.0	188.3	177.8	180.2
2008	198.1	194.2	196.7	203.8	188.4	199.5
2009	194.1	204.1	209.1	220.7	186.8	214.9
2010	210.4	214.6	222.3	239.0	200.3	231.7
2011	230.3	225.5	236.3	258.9	200.8	252.3
2012	242.4	236.9	251.1	280.1	203.5	271.6
2013	245.5	248.9	267.0	302.5	217.8	293.4
R^2 of prediction		0.931	0.706	0	0	0
RMSE of prediction		5.7	11.7	31.7	22.5	25.7

[a] Net electricity demand was reported in these works. Net electricity demand was assumed to be 77.7% of the gross electricity demand due to losses in the network (Erdogdu, 2007)

in this time period. Fig. 9 shows that the electricity demand will almost be doubled in the next 15 years interval reaching over 460 TW h in the year 2028.

The forecasted results of the gross annual electricity demand obtained by the ANN model employed in this work was compared with the official predictions based on high demand and low demand scenarios for the 2014–2021 time period (TEIAS, 2012). Similar to the comparison done for the years between 2007 and 2013, the official predictions were found to be much higher than the predictions of the ANN model as indicated by Fig. 10.

4. Conclusion and policy implications

In this work, the annual gross electricity demand of Turkey was modeled by MLR and ANN models using population, GDP per capita, inflation percentage, unemployment percentage, average summer temperature and average winter temperature as the descriptor variables. Population and GDP per capita were found to be the major factors affecting the demand while inflation percentage and average summer temperature had minor influence. Among these variables the unemployment percentage and the average winter temperatures were found to be insignificant for determining the demand for the years between 1975 and 2013.

Next, the data between 2007 and 2013 were excluded from the

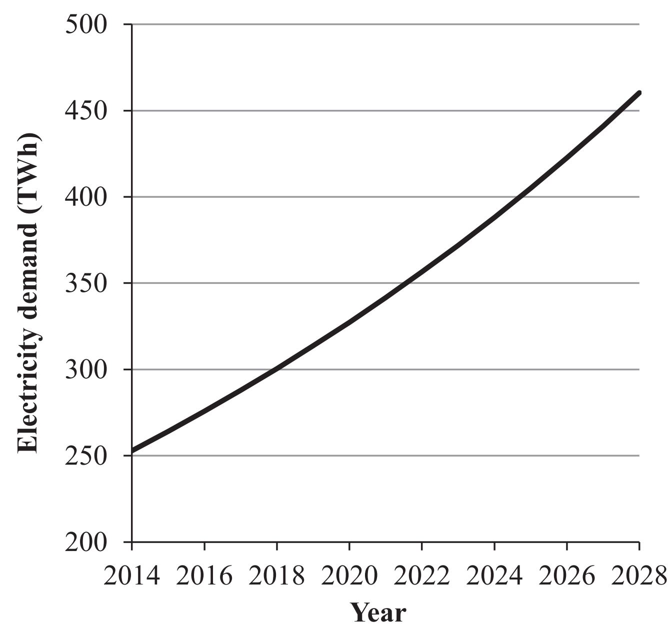

Fig. 9. Forecasting the electricity demand for the years between 2014 and 2028 by ANN model.

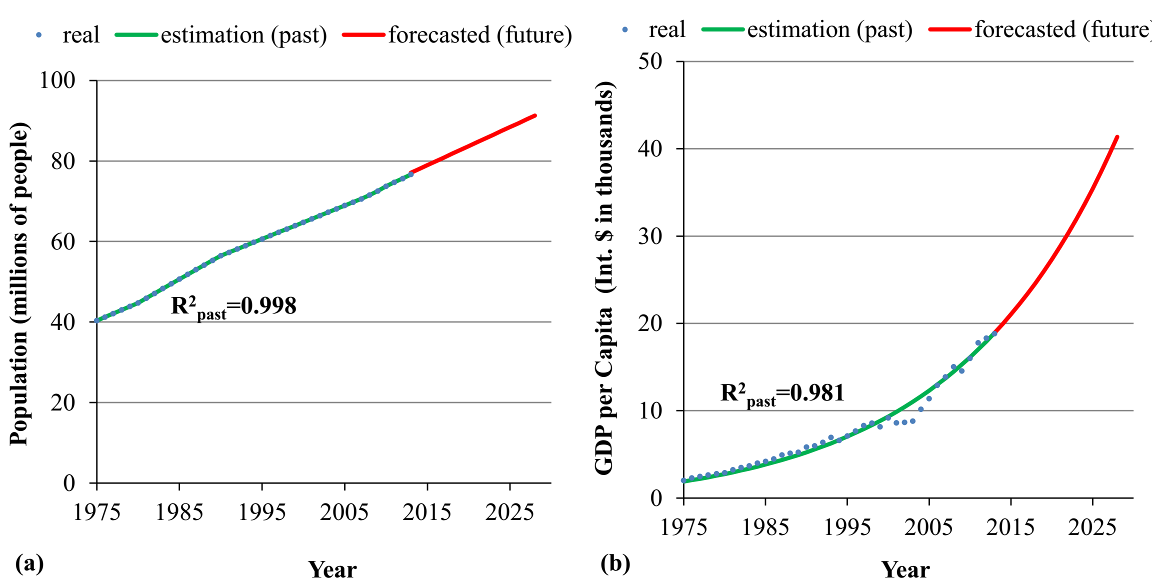

Fig. 8. Forecasting (a) population and (b) GDP per capita for the years 2014–2028 using time series ANN model.

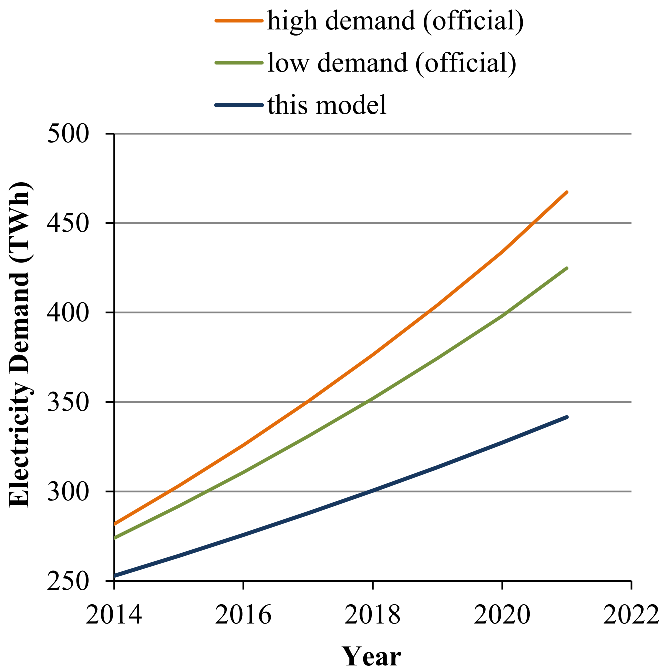

Fig. 10. Comparison of predictions by the ANN model applied in this work with the official predictions for the years between 2014 and 2021.

dataset, and the values of the statistically significant descriptor variables in these years were predicted by time series ANN models trained by the data between 1975 and 2006. Then, a MLR and a multilayer perceptron ANN model were constructed to simulate the predicted values of these variables to forecast the demand between 2007 and 2013. It was found that the ANN model forecasted the electricity demand with a very high accuracy, and the results were superior to the official predictions (done by Ministry of Energy and Natural Resources of Turkey); in contrast, the MLR model failed to forecast the demand with an acceptable accuracy. This way, the ANN modeling approach applied in this work was validated successfully.

Then, a similar procedure was applied for forecasting the electricity demand for the years between 2014 and 2028 using the predicted values of the statistically significant descriptor variables. Finally, the predicted values of these variables in these years were used to simulate a multilayer perceptron ANN model to forecast the electricity demand for the future years; it was found that the electricity demand will be doubled in the next 15 years reaching over 460 TW h in the year 2028. To conclude, after determining the statistically significant descriptor variables influencing the electricity demand, the approach applied in this work can be implemented for other countries to make accurate predictions for the future.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.enpol.2015.12.019.

References

Akay, D., Atak, M., 2007. Grey prediction with rolling mechanism for electricity demand forecasting of Turkey. Energy 32, 1670–1675.
Askarzadeh, A., 2014. Comparison of particle swarm optimization and other metaheuristics on electricity demand estimation: a case study of Iran. Energy 72, 484–491.
Azadeh, A., Saberi, M., Seraj, O., 2010. An integrated fuzzy regression algorithm for energy consumption estimation with non-stationary data: a case study of Iran. Energy 35, 2351–2366.
Behrang, M.A., Assareh, E., Ghalambaz, M., Assari, M.R., Noghrehabadi, A.R., 2011. Forecasting future oil demand in Iran using GSA (Gravitational Search Algorithm). Energy 36, 5649–5654.
Bianco, V., Manca, O., Nardini, S., 2009. Electricity consumption forecasting in Italy using linear regression models. Energy 34, 1413–1421.
Bianco, V., Manca, O., Nardini, S., 2013. Linear regression models to forecast electricity consumption in Italy. Energy Sources, Part B: Econ., Plan., Policy 8, 86–93.
Bianco, V., Manca, O., Nardini, S., Minea, A.A., 2010. Analysis and forecasting of nonresidential electricity consumption in Romania. Appl. Energy 87, 3584–3590.
De Felice, M., Alessandri, A., Ruti, P.M., 2013. Electricity demand forecasting over Italy: potential benefits using numerical weather prediction models. Electr. Power Syst. Res. 104, 71–79.
Ekonomou, L., 2010. Greek long-term energy consumption prediction using artificial neural networks. Energy 35, 512–517.
Erdogdu, E., 2007. Electricity demand analysis using cointegration and ARIMA modelling: a case study of Turkey. Energy Policy 35, 1129–1146.
García-Ascanio, C., Maté, C., 2010. Electric power demand forecasting using interval time series: a comparison between VAR and iMLP. Energy Policy 38, 715–725.
Geem, Z.W., Roper, W.E., 2009. Energy demand estimation of South Korea using artificial neural network. Energy Policy 37, 4049–4054.
Ghanbari, A., Kazemi, S.M.R., Mehmanpazir, F., Nakhostin, M.M., 2013. A Cooperative Ant Colony Optimization-Genetic Algorithm approach for construction of energy demand forecasting knowledge-based expert systems. Knowl.-Based Syst. 39, 194–206.
Günay, M.E., Yildirim, R., 2011. Neural network analysis of selective CO oxidation over copper-based catalysts for knowledge extraction from published data in the literature. Ind. Eng. Chem. Res. 50, 12488–12500.
Hamzaçebi, C., 2007. Forecasting of Turkey's net electricity energy consumption on sectoral bases. Energy Policy 35, 2009–2016.
Hamzacebi, C., Es, H.A., 2014. Forecasting the annual electricity consumption of Turkey using an optimized grey model. Energy 70, 165–171.
IMF, 2013. International Monetary Fund, World Economic Outlook Database, ⟨http://www.imf.org/external/pubs/ft/weo/2013/01/weodata/index.aspx⟩; (accessed 13.08.15).
Inglesi, R., 2010. Aggregate electricity demand in South Africa: conditional forecasts to 2030. Appl. Energy 87, 197–204.
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. Appl. Energy 88, 1927–1939.
Kavaklioglu, K., 2011. Modeling and prediction of Turkey's electricity consumption using support vector regression. Appl. Energy 88, 368–375.
Kavaklioglu, K., Ceylan, H., Ozturk, H.K., Canyurt, O.E., 2009. Modeling and prediction of Turkey's electricity consumption using Artificial Neural Networks. Energy Convers. Manag. 50, 2719–2727.
Kennedy, P., 2008. A Guide to Econometrics, 6 ed. Blackwell Publishing, Malden.
Kheirkhah, A., Azadeh, A., Saberi, M., Azaron, A., Shakouri, H., 2013. Improved estimation of electricity demand function by using of artificial neural network, principal component analysis and data envelopment analysis. Comput. Ind. Eng. 64, 425–441.
Kialashaki, A., Reisel, J.R., 2014. Development and validation of artificial neural network models of the energy demand in the industrial sector of the United States. Energy 76, 749–760.
Kıran, M.S., Özceylan, E., Gündüz, M., Paksoy, T., 2012. A novel hybrid approach based on Particle Swarm Optimization and Ant Colony Algorithm to forecast energy demand of Turkey. Energy Convers. Manag. 53, 75–83.
Kucukali, S., Baris, K., 2010. Turkey's short-term gross annual electricity demand forecast by fuzzy logic approach. Energy Policy 38, 2438–2445.
Larose, D.T., 2006. Data Mining Methods and Models. Wiley, New Jersey.
MGM, 2014. Turkish State Meteorological Services, Official Statistics, ⟨http://www.mgm.gov.tr/veridegerlendirme/il-ve-ilceler-istatistik.aspx?k=C⟩; (accessed 13.08.015).
Nawaz, S., Iqbal, N., Anwar, S., 2014. Modelling electricity demand using the STAR (Smooth Transition Auto-Regressive) model in Pakistan. Energy 78, 535–542.
O'brien, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. Qual. Quant. 41, 673–690.
OECD, 2015. Organization for Economic Cooperation and Development ⟨http://stats.oecd.org⟩; (accessed 13.08.15).
Pan, X., Brian, L., Zhang, C.A., 2013. A Comparison of Neural Network Backpropagation Algorithms for Electricity Load Forecasting. In: Proceedings of IEEE International Workshop on Intelligent Energy Systems, pp. 22–27.
Pao, H.T., 2009. Forecasting energy consumption in Taiwan using hybrid nonlinear models. Energy 34, 1438–1446.
TEIAS, 2007. Turkish Electricity Energy 10-year Capacity Projection Report (in Turkish) ⟨http://www.teias.gov.tr/YayinRapor/APK/projeksiyon/index.htm⟩; (accessed 13.08.15).
TEIAS, 2012. Turkish Electricity Energy 10-year Capacity Projection Report (in Turkish) ⟨http://www.teias.gov.tr/YayinRapor/APK/projeksiyon/index.htm⟩; (accessed 13.08.15).
TEIAS, 2013. Turkish Electricity Transmission Company, Electricity Generation & Transmission Statics of Turkey ⟨http://www.teias.gov.tr/TürkiyeElektrikİstatistikleri/istatistik2013/istatistik2013.htm⟩; (accessed 13.08.15).

TurkStat, 2013. Turkish Statistical Institute ⟨http://www.turkstat.gov.tr⟩; (accessed 13.08.15).
Walpole, R.E., Myers, R.H., Myers, S.L., Ye, K., 2012. Probability & Statistics for Engineers & Scientists, 9 ed. Pearson, Boston.
Wilamowski, B.M., Chen, Y., 1999. Efficient algorithm for training neural networks with one hidden layer. In: Proceedings of International Joint Conference on Neural Networks 3, pp. 1725–1728.
WWID, 2014. World Wide Inflation Data, Historic inflation Turkey ⟨http://www.inflation.eu/inflation-rates/turkey/historic-inflation/cpi-inflation-turkey.aspx⟩; (accessed 13.08.15).
Zahedi, G., Azizi, S., Bahadori, A., Elkamel, A., Wan Alwi, S.R., 2013. Electricity demand estimation using an adaptive neuro-fuzzy network: a case study from the Ontario province – Canada. Energy 49, 323–328.