Energy 197 (2020) 117200

Contents lists available at ScienceDirect

Energy

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

A hybrid approach based on autoregressive integrated moving average and least-square support vector machine for long-term forecasting of net electricity consumption

Fazil Kaytez

General Directorate of Energy Affairs, Department of Energy Policies and Technologies, 06520, Ankara, Turkey

ARTICLE INFO

Article history:
Received 27 July 2019
Received in revised form
13 February 2020
Accepted 19 February 2020
Available online 21 February 2020

Keywords:
Energy modeling
Least square support vector machines
Autoregressive integrated moving average
Multiple regression
Prediction
Electricity consumption

ABSTRACT

Electricity consumption is on the rise in developing countries. Most of the research studies in energy demand forecasting aim to provide that sufficient electricity is produced to meet future needs. A reliable forecasting model is necessary for accurate investment planning of electricity generation and distribution. The main goal of this study is to develop effective and realistic solutions for electricity consumption forecasting in Turkey. This paper proposes a hybrid model based on least-square support vector machine and an autoregressive integrated moving average. This hybrid approach's forecast results are compared with multiple linear regression approach, a single autoregressive integrated moving average model, official forecasts and similar studies in literature. Also, it is applied to forecast the future net electricity consumption for Turkey until 2022. The study results indicate that the proposed model can generate more realistic and reliable forecasts. It can also be stated that it responds better to some unexpected reactions in the time series.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Electricity consumption has become a topic of high importance for many countries. The increasing demand in the market has largely been triggered by countries' increasing population and industrialization. Long term forecasting of electricity consumption is crucial in planning, analysis and management of power systems in order to ensure continuous, secure and economic supply of electricity [1]. Studies focusing on the forecasting of electricity consumption using different techniques are quite substantial for developing countries like Turkey, India, South Africa etc.

There are few important points to establish accurate forecasting models of consumption. At first, it is very significant to accurately identify the indicators that have strong impacts on the consumption of the country and to add these indicators to the forecast model, before performing the forecasts. Each country may have different model inputs according to their own conditions, and the impact and number of these inputs on consumption may affect the forecast performance significantly. Secondly, important point is to choose a suitable modeling methodology. The non-linear relationship between most of the input and output variables and the difficulties in expressing this relationship mathematically is one of the challenges in this area. The criteria to improve the forecast performance becomes prominent most of the time, instead of theoretical criteria in the model selection. Thirdly, an important point is that the methodology used should be able to respond future events, in other words it should be able to produce forecasts.

Turkey has been one of the fastest growing electricity markets in the world with its increasing population, powerful financal structure and speedy urbanization for two decades. Installed capacity of Turkey has risen to around 87.139 MW in the first half of 2018. However, Turkey is an energy importing country and its energy import dependency has been increasing in recent years. Today, electricity generation in Turkey is fundamentally dependent on natural gas and coal power plants. According to Ministry of Energy and Natural Resources (MENR) reports, electricity generation in Turkey is supplied from natural gas by 37%, coal by 33%, hydroenergy by 20%, wind by 6%, geothermal by 2% and other resources by 2% in 2017 [2]. Turkey has the highest rate of growing electricity demand among OECD countries over the last decades, with an average annual growth rate of 6%. Energy supplies in Turkey increased nearly 60% between 2005 and 2016. However, electricity consumption of the country rose even faster nearly 70%

E-mail address: fkaytez@enerji.gov.tr.

https://doi.org/10.1016/j.energy.2020.117200
0360-5442/© 2020 Elsevier Ltd. All rights reserved.

over the same period. Turkey's net electricity consumption (internal consumption, grid losses and energy theft included) is 207,38 TWh for the year 2014 and 217,31 TWh for 2015. Also, the country's gross electricity production (gross generation + import − export) is 265,7 billion kWh for the year 2015 and 278,3 TWh for the 2016 by annual increase of 3,3%. In 2017, the net electricity consumption in Turkey was 249,02 billion kWh and it increased by 7,7% compared to the prior year. And Turkey's electricity consumption in the year 2022 is estimated to be around 336,52 TWh according to MENR's scenarios [2].

The electricity consumption forecasts in Turkey are officially determined by the MENR. The Model for Analysis of Energy Demand (MAED) is used in these official estimates and is a simulation model for evaluating the medium and long-term demand for electricity and other energy demands of Turkey. In this model, a very large set of independent variables (economic, social, technical, etc.) are used as model inputs. And the electrical power generation planning studies, which constitute a significant part in this research, are conducted by Turkish Electricity Transmission Corporation (TETC), and "Wien Automatic System Planning Package" (WASP) software model is preferred for this purpose.

Although the results of the MAED model are comparable to the actual gross consumption values, its error rates are quite high. And, the fact that the MAED is not able to model Turkey's net electricity consumption can be considered as a deficiency. Because, the gross consumption is a statistical information that does not present the actual demand. Whereas, the net consumption is a significant statistical indicator that includes the actual consumption data (industry, residence, lighting etc.) provided from all of the distribution system operators (DSOs). Due to great differences between the MAED model results and the actual values, MENR and TETC generally revise their long-term forecasts once every six months. It is quite important, therefore, to select an appropriate consumption estimation model that will provide precise estimates of future demand of electricity in Turkey. In order to obtain more precision forecasts, Turkey needs to use alternative models and methodologies.

To predict exactly future electricity consumption and demand, several studies have offered different models that use econometric, artificial intelligence (AI) and hybrid approaches in the literature. Regression analysis, Auto-Regressive Iterative Moving Avarage (ARIMA), Seasonal Autoregressive Iterative Moving Avarage (SARIMA) and grey theory have been presented as econometric solutions. For example, Bianco et al. [3] forecasted Italy's electricity consumption by linear regression models and researched the effect of economic and demographic variables which are the historical consumption, population, GDP (Gross Domestic Product) per capita and GDP on the annual electricity consumption in Italy to improve a long-term consumption forecasting model. Erdogdu [4] used a ARIMA model in electricity demand analysis, Tunc et al. [5] predicted the electricity consumption demand with regression analysis. Pappas et al. [6] proposed a ARIMA model for Greece electricity consumption and the model results compared with some analytical time-series models. Ediger and Akar [7] used ARIMA and SARIMA methods for electricity demand forecasting. Also, there are several studies that projected the energy consumption by traditional methods. Egelioglu et al. [8] studied the influence of economic variables on the annual electricity consumption. Crompton et al. [9] applied the Bayesian vector autoregressive methodology to forecast China's energy consumption. Grey theory was utilized by some researchers to forecast electricity consumption. Nai-ming et al. [10] used a discreate grey and markov model to forecast the total amount of China's energy consumption. Hamzacebi et al. [11] forecasted Turkey's electric demand for the 2013−2025 period bu using an optimized grey modeling (1,1) approach. Akay et al. [12] proposed a grey prediction with a rolling mechanism to predict Turkey's total electricity consumption. In some studies, the AI methods have been also submitted for modeling of electricity consumption and demand. Jasinski [13] provided an innovative approach to the modeling of Poland's electricity consumption using artificial neural networks (ANN). Dozic et al. [14] carried out an analysis of European Union energy system using AI approaches. Azadeh et al. [15] presented an AI approach for optimum estimation of renewable energy consumption and compared the approach's results with some fuzzy regression models. Some researchers have proposed different models to improve the prediction performance using neural networks and neuro-fuzzy approach. Beccali et al. [16] presented a forecasting model based on an Elman ANN method for the short-time prediction. In recent years, the Support Vector Machine (SVM) and Least Square Support Vector Machine (LSSVM) have been used in various energy research, such as modeling, forecasting and power engineering. Esen et al. [17] reported a modeling study of a solar air heater system efficiency by using LSSVM technique. Sulaiman et al. [18] presented a new method for power tracing in a power system using a hybrid artificial bee colony and LSSVM algorithms. Ekici [19] developed a LSSVM based intelligent model to predict the next day's solar insolation. Bessedik et al. [20] carried out the prediction of flashover voltage of insulators using LSSVM and particle swarm optimization. Kavaklioglu [21] used support vector regression model for forecasting Turkey electricity consumption. Furthermore, hybrid approaches have been presented to suggest some models for catching nonlinear characteristics of electricity demand and consumption. Barak and Sadegh [22] used a ensemble ARIMA-ANFIS hybrid algorithm for forecasting energy consumption. Tutun et al. [23] developed a new forecasting framework in net electricity consumption using LASSO (Least Absolute Shrinkage and Selection Operator). Gurbuz et al. [24] used an artificial bee colony model to optimize time series in forecasting of electricity consumption. Some researchers focused on comparing models to each other in order to decide the best method for estimating the electricity consumption. Geoffrey and Kelvin [25] compared the accuracy in predicting electricity energy consumption in Hong Kong among three different approaches: regression analysis, decision trees and neural networks. Bilgili et al. [26] compared the precision in forecasting electricity demand in residential and industrial sectors of Turkey using regression and ANN models. Kaytez et al. [27] presented a LSSVM model for electricity consumption of Turkey and also compared the LSSVM forecast results among other two different approaches: ANN and regression analysis. Ayvaz and Kusakci [28] used a Nonhomogeneous Discrete Grey Model (NDGM) to forecast Turkey's electricity consumption from 2014 to 2030. Also, several studies were carried out on Turkey's electricity consumption forecast using diverse approaches. A summary of the methods and researchers are given in Table 1. As a result, the researchers need to suggest better approaches in order to obtain accurate results in forecasting studies. However, new approaches are always needed with more accurate scenarios and extensive data sets.

The main motivation of this study is to propose a hybrid ARIMA and LSSVM model for long term electricity consumption forecasting. In this study, the optimal model for Turkey is investigated by using different combinations of certain energy and socioeconomic indicators. Also, this study compares the accuracy in predicting electricity consumption in Turkey among two diverse approaches: the proposed model and MAED model (official estimates). The analysis shows that the proposed model can produce more meaningful results and better capture the unanticipated effects in the time series. Therefore, this study contributes to the area of energy consumption forecasting.

The rest of the paper is organized as follows. The methods used

Table 1
Some studies on electricity consumption for Turkey.

Authors	Consumption type used	Method used	Independent variables	Data used	Estimated years
Kavaklioglu [21]	Electricity consumption	SVR	Years, population, Gross national product, import, export	1975–2006	2007–2016
Sozen et al. [29]	Net electricity consumption	ANN	Installed capacity, Gross generation, years, population	1953–2000	Model is presented
Kavaklioglu et al. [30]	Total electricity consumption	ANN	GNP, population, import, export	1975–2006	2007–2027
Sozen and Arcaklıoglu [31]	Net electricity consumption	ANN	Installed capacity, gross generation, import, export, net consumption, population	1975–2003	2004–2020
Kankal et al. [32]	Net electricity consumption	ANN	Import, export, population, GDP	1980–2007	2008–2014
Hamzacebi [33]	Sectoral consumption	ANN	Industrial sector, residence, Transportation, agriculture	1970–2004	2003–2020
Hamzacebi and Es [34]	Annual electricity consumption	Optimized Grey Model	Data of total electricity consumption	1945–2010	2011–2025
Sozen and Arcaklioglu [35]	Total net electricity consumption	ANN	GNP, GDP, population, installed capacity, gross generation, import, export	1968–2005	Model is presented

in the proposed hybrid approach are explained briefly in Section 2; It is described in detail how the proposed approaches is used to forecast the net electricity consumption in Section 3; In Section 4, It is reported the experimental results and compared the proposed model with the single ARIMA and multiple linear regression (MLR) models, as well as analogous works in literature and the official forecasting model (MAED) using different types of performance criteria. Also, Turkey's future electricity consumption results up to 2022 are obtained with the proposed model; Finally, conclusions from this paper are presented in Section 5.

2. Theoretical background

Regression models are statistical methods used to analyze the relationship between a single dependent variable and a set of independent variable and traditionally used in electricity consumption forecasting. Numerous theoretical relations can also be expressed with regression analysis. However, a limitation on regression models is that they are just current in the series which data have been extracted. For this reason, to obtain a universal experimental relationship, it is essential to have a lot of various series of data. Generally, multiple regression processes will estimate a linear equation of the form. An example of a regression equation is given below [36–38]:

$$y^{'} = a_1 \times x_1 + a_2 \times x_2 + a_3 \times x_3 + \ldots\ldots + a_n \times x_n \quad (1)$$

$$y = y^{'} + e \quad (2)$$

where, $y^{'}$ is the regression model output, y represents the real output, a_1 to a_n express the regression coefficients, x_1 to x_n express the independent variables (or influence parameters) and e is the error term.

ARIMA model is a model that can be fitted to time series in order to better understand or predict future values in the series. Also, it is a simple and popular linear approach. ARIMA model consists of two components: Auto regressive (AR) component and Moving average (MA) component. The AR component is written as [39]:

$$y_t = u_t + a_1 y_{t-1} + \ldots + a_p y_{t-p} + c \quad (3)$$

where y_t is the actual value, p is the order of the AR, u_t is the white noise(error), the c is a constant and $a_1, \ldots, a_p$ are the AR parameters. Likewise, The MA component can be expressed as:

$$y_t = u_t + m_1 u_{t-1} + \ldots + m_q u_{t-q} + \mu \quad (4)$$

where u_t, u_{t-1}, …,u_{t-q} are error terms, μ is the expectation of actual value (y_t), $m_1, \ldots, m_q$ are the MA parameters and q is the order of MA. Integrating these models with the same training data will create the ARIMA(p,q) model [39]:

$$y_t = u_t + a_1 y_{t-1} + \ldots + a_p y_{t-p} + m_1 u_{t-1} + \ldots + m_q u_{t-q} + c \quad (5)$$

where p is the autoregressive term and q is the moving average term.

The fundamental assumption of this model is that time series data includes statistical stationarity, meaning that measured statistical properties such as autocorrelation, variance and mean remain invariable over time [40]. However, if the training data is not stationary, the ARIMA model needs differenced data to transform it to stationarity. This is displayed as ARIMA (p,d,q) where d is the degree of differencing [41].

The SVM that was firstly proposed by Vapnik [42] is a important machine learning techniques for solving problems in nonlinear classification and function estimation. This technique maps the data into a higher dimensional space (future space) and constructs an optimal separating hyperplane in this space. The quality of SVM solution does not depend directly on the dimensionality of the input space [42–45]. SVM is quite powerful on non-linear regression forecasting problems. LSSVM was introduced by Suykens et al. [46]. In fact, LSSVM is a reformulation of standard SVM. The performance of LSSVM is better than conventional SVM in terms of fast convergence, simple calculation and high precision. On the other hand, it is very difficult to select suitable LSSVM parameters in many applications. The selection of LSSVM parameters affects the regression precision of LSSVM [46,47].

The LSSVM simplifies the SVM procedure. The LSSVM uses a organized least squares function with equality constraints, leading to a linear system which meets the Karush-Kuhn-Tucker (KKT) conditions for obtaining an optimal solution. Therefore, the regression problem can be solved by a linear equation system rather than quadratic programming, as in SVM. LSSVM is generally used for the control theory, classification and regression problems. This section briefly introduces LSSVM for regression. The Least Squares Support Vector Regression (LSSVR) method is to approximate an unknown function by using given a sample of a training data series. $\{x_i, y_i\}_{i=1}^{l}$. The linear SVM regression algorithm tries to find the function

$$y = f(x) = w^T \phi(x) + b \tag{6}$$

where the $x \in \Re^n, y \in \Re$ and $\phi(.) : \Re^n \rightarrow \Re^{nh}$ is the mapping to the high dimensional feature space. LSSVM method considers the regression problem as the following optimization problem.

$$\min J_1(w, b, e) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^{l} e_i^2 \tag{7}$$

subject to constrains to $y_i = w^T \phi(x_i) + b + e_i, \quad i = 1,\ 2, \ldots l$ (8)

where e_i is the prediction error term for data point i, γ is a regularization constant and $\phi(x_i)$denotes an infinite dimensional feature map. Introduce the lagrangian function as

$$L_1(w, b, e; \alpha) = J_1(w, b, e) + \sum_{i=1}^{l} \alpha_i \left(y_i - w^T \phi(x_i) - b - e_i \right) \tag{9}$$

where $\alpha = (\alpha_1; \alpha_2; \ldots; \alpha_l) \in \Re$ are the Langrange multipliers, which can be positive or negative in LSSVM formulation. The conditions for optimality are

$$\begin{aligned} &\frac{\partial L_1}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{l} \alpha_i \phi(x_i), \\ &\frac{\partial L_1}{\partial b} = 0 \rightarrow \sum_{i=1}^{l} \alpha_i = 0, \\ &\frac{\partial L_1}{\partial e_i} = 0 \rightarrow e_i = \frac{1}{\gamma} \alpha_i, \\ &\frac{\partial L_1}{\partial \alpha_i} = 0 \rightarrow y_i = w^T \phi(x_i) + b + e_i, \end{aligned} \tag{10}$$

For $i = 1,\ 2,\ 3 \ldots l k(x_i, x_j)$specifies a kernel function whose value equals the inner product of x_i and $x_j$vectors in the feature space $\phi(x_i)$and $\phi(x_j)$.

$$k(x_i, x_j) = \phi(x_i)^T \phi(x_j) \tag{11}$$

The basic features of a kernel function are derived from Mercer's theorem. Applicable kernel functions must fulfill Mercer's conditions [48,49]. The linear and the polynomial functions are used as kernel functions in many applications.

$$k(x_i, x_j) = \exp\left(- \parallel x_i - x_j \parallel^2 \Big/ \sigma^2 \right) \ (RBF) \tag{12}$$

The mercer's condition is valid for all σvalues in the radial basis function case and positive t values in the polynomial function case. Hereby, it can be showed $K = (k_{ij})_{lxl},\ k_{ij} = k(x_i,\ x_j)$and Vas V = diag(1/γ,1/γ, …,1/γ), thus LSSVM regression model can be described as:

$$\begin{bmatrix} A & 1 \\ 1^T & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ b \end{bmatrix} = \begin{bmatrix} y \\ 0 \end{bmatrix}, \tag{13}$$

where $A = K + V$. Hereby, the regression model in Eq. (8) is found by solving Eq. (16). By using the lagrangian method, the LSSVM model can be expressed as

$$f(x) = \sum_{i=1}^{l} \alpha_i K(x, x_i) + b \tag{14}$$

where α_i and $b$are the solutions to the linear system. LSSVM regression model in Eq. (16) contains two parameters: regularization parameter (C) and the kernel function parameter (σ) [46,50].

2.1. The proposed hybrid approach for the forecasting of electricity consumption

A common approach in analyzing time series is to allocate the series into three components including trend, seasonality, and one non-systematic component called residual. This approach provides a structural way of thinking about forecasting problem in a time-series, both in terms of the complexity of the modeling and especially in terms of how best to capture these components in a particular model.

Some factors such as economic crises and economic growth significantly affect the trend in electricity consumption. Therefore, there is a need to explore the nonlinear component of electricity consumption time series for the upgrade of forecasting accuracy. Xie [51] proposed a hybrid methodology combining SARIMA and least square support vector regression (LSSVR) for container throughput forecasting and used this hybrid approach to adapt the time series of container throughput. This research uses a similar approach to fit the nonlinear part of the long term electricity consumption series.

In this section, the time series of electricity consumption is separated into two parts with ARIMA model: trend component and irregular (residual) component. The proposed hybrid approach is described in Fig. 1 and the following steps:

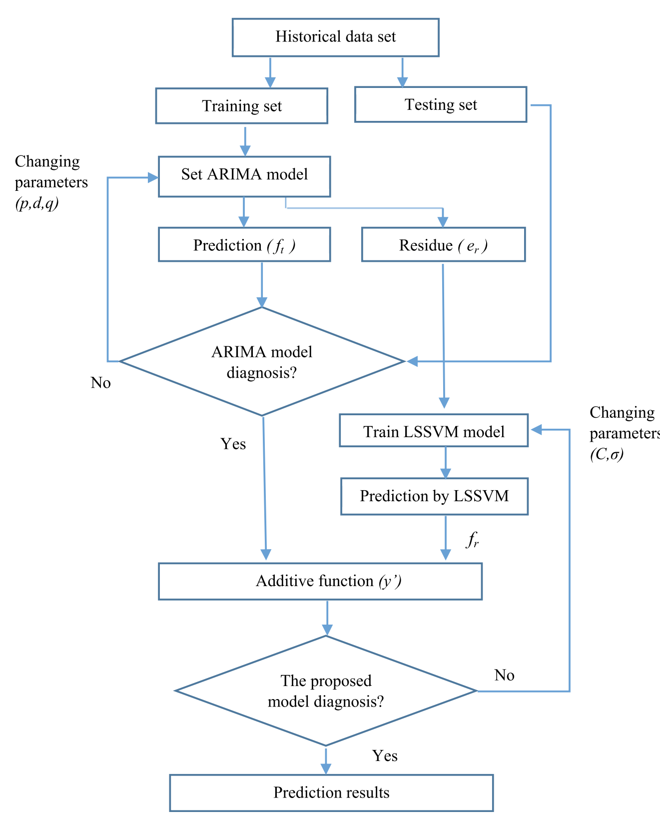

Fig. 1. The proposed approach for long-term electricity consumption forecasting.

Step 1: ARIMA approach is utilized to fit the trend component of the series and to create a series of forecasts (f_t).

Step 2: A series of irregular components (e_r) are attained by subtracting the forecasts of the trend component from the observed values (y_t)

$$e_r = y_t - f_t \tag{15}$$

Step 3: LSSVM is utilized to fit the irregular component (e_r) and to generate a series of forecasts (f_r).

Step 4: And the final forecasts ($y^{'}$) of net electricity consumption are calculated, as is seen in Eq. (16).

$$y^{'} = f_t + f_r \tag{16}$$

3. Case studies

This section describes in detail how the proposed approach is used to forecast the net electricity consumption.

3.1. Data collection

In the multiple linear regression analysis, three different models are created for Turkey's electricity consumption model. These models are selected by considering the models' correlations with electricity consumption. Indicators such as Gross Electricity Generation (GEG), Population (P), Installed Capacity (IC), Import (I), Export (E) and Total Subscribership (TS) are utilized in different combinations to find the optimal model. The economic indicator Gross Domestic Product (GDP) has not been preferred in any of the models. Because the causalities between electricity consumption and the economic growth are still controversial. Although GDP is a very effective parameter in electricity consumption in some of the countries, changes in GDP does not cause any significant impact in some other countries. For instance, some studies carried out in Turkey indicate that the increase or decrease in electricity consumption affects the GDP [52,53]. Basically, it may be said that economic crisis are more effective on Turkey's electricity consumption trends rather than GDP. However, it is really difficult to predict the economic crisis.

In order to create the electricity consumption models in regard to the available information for Turkey, six categories of data were collected from different official sources. Historical population data from 1970 to 2017 were gathered from Turkish Statistical Institute (TURKSTAT) [54] The electricity import and export amounts were provided by the TETC [55] for the same time period. Additionally, installed capacity and gross electricity generation data were also provided from TETC [56,57]. Between 1970 and 2017, Turkey's net electricity consumption data and the number of subscribers were obtained from the annual reports of Turkish Electricity Distribution Corporation (TEDC) [58] which is a public institution. The data used in these models are divided into three groups. In all prediction models, %85 of the input data are used for training, %11 of them are used for testing and %4 of them data are used for validation purposes.

3.2. Performance criteria

In order to evaluate the performance of the proposed model in work, various statistical tests are carried out. These tests are the mean squared error (MSE), the root mean squared error (RMSE), the mean absolute percentage error (MAPE). In notation, the mathematical formulas used to compute MSE, RMSE and MAPE are as follows:where M is the number of forecasting periods; d_t is the actual value at period t; and z_t is the forecasting value at period t (Table 2).

Moreover, when evaluating which the ARIMA model best fits the data, Akaike Information Criterion (*AIC*), Bayesian Information Criterion (*BIC*, or Schwarz criterion) and Hannan-Quinn Criterion (*HQC*) values are used to compare them. The AIC is described as

$$AIC = 2k + \left[\ln\left(\frac{RSS}{n_{obs}}\right)\right] \tag{17}$$

where *RSS* is the residual sum of squares, k is the number of independent parameters, and n_{obs} is the total number of observations. The *BIC* has a larger penalty term than the *AIC*. The *BIC* is described as

$$BIC = n_{obs}.\ln\left(\sigma^2\right) + k.\ln(n_{obs}) \tag{18}$$

where σ^2 is the error variance, n_{obs} and k parameters are the same as for *AIC*. Furthermore, The *HQC* is one of the indicators used for ARIMA model parameter selection.

$$HQC = n_{obs}.\left[\ln\left(\frac{RSS}{n_{obs}}\right)\right] + k.\ln[In(n_{obs})] \tag{19}$$

where all parameters are the same as for *AIC* and *BIC*.

3.3. Experiment

The analyzes conducted within the scope of this study mainly consist of three stage: (*i*) the selection of the appropriate multiple linear regression model. (*ii*) the implementation of the proposed hybrid model, and (*iii*) in order to verify the effectiveness of the proposed hybrid model, evaluation of the model's predictive ability by performance metrics and comparison of the results with the official estimates and similar studies in the literature. Fig. 2 shows the general flow of this study.

(i) *Stage I:* As known, determining a correlation and the degree of this correlation between numerous variables lays at the heart of majority of the problems encountered in scientific researches. Regression is one of the most practiced techniques in studying correlation between variables. Multiple linear regression analysis is used for modeling the net electricity consumption of Turkey in the first stage of the analysis. The three models taking different demographic, socio-economic and energy indicators into consideration are as follows:

Model 1: $y = b_1x_2 + c_1x_3 + f_1$

Model 2: $y = a_1x_1 + b_2x_2 + c_2x_3 + d_1x_4 + e_1x_5 + g_1x_6 + f_2$

Table 2
Performance metrics and output variables.

Performance metric	Calculation
MSE	$\frac{1}{M}\sum_{t=1}^{M}\left(d_t - z_t\right)^2$
RMSE	$\left\{\frac{1}{M}\sum_{t=1}^{M}(d_t - z_t)^2\right\}^{\frac{1}{2}}$
MAPE	$\frac{100}{M}\sum_{t=1}^{M}\left\|\frac{d_t - z_t}{d_t}\right\|$

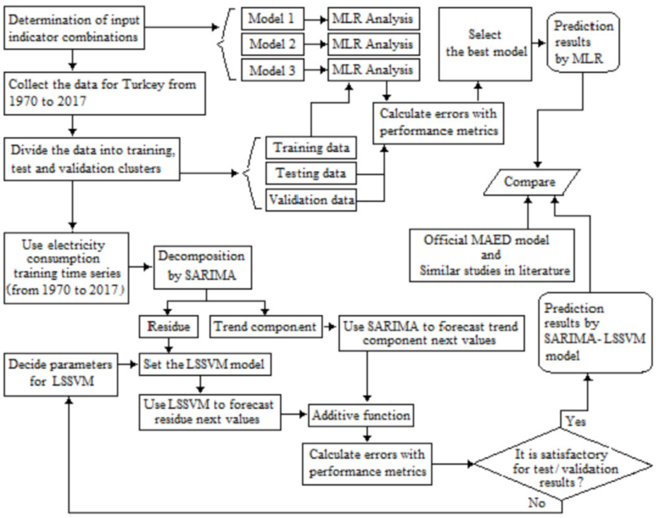

Fig. 2. The flow chart of the experiment.

Model 3: $y = a_2x_1 + b_3x_2 + c_3x_3 + d_2x_4 + f_3$

where y presents the estimated net electricity consumption; a_1-a_2, b_1-b_3, c_1-c_3, d_1-d_2, f_1-f_3, e_1 and g_1 values are the regression coefficients. The x values express the six independent variables used as the predictors of y (Table 3). The results of the linear regression models are presented in Table 4 and Table 5. The coefficients of Adjusted R^2 are obtained as 0.988, 0.994 and 0.996 for Model 1, 2 and 3, respectively.

As it can be seen in Table 5, the training errors of the Model 1 and Model 2 are rather low, but the testing errors in Model 1 and Model 2 are high. On the other hand, Model 3 has higher training error but has a lower testing error rate with 2.54 of RMSE. In addition, the higher Adjusted R^2 coefficient of the Model 3 indicates that the Model 3 is more appropriate than the other models. Therefore, the Model 3 regression equation is taken into consideration in the comparison of the results.

(ii) *Stage II:* Except for some periods of economic recession, the country's net electricity consumption show a linear distribution year by year. A considerable fluctuation and seasonality are not observed in the historical consumption data that examined. Firstly, The ARIMA model is fitted to a stationary time series and the yearly electricity consumption data need regular differencing to become stationary. Autocorrelation function (ACF) and a partial autocorrelation function (PACF) are used to identify the order of the ARIMA model (Fig. 3). The model's parameters are estimated by a maximum likelihood function. The goodness-of-fit of the model is tried on the model residuals. When evaluating which temporary model best fits the data, *AIC* and *BIC* are used to select the best fitted model. The aim is to choose the model orders that result in minimum values of *AIC* and *BIC*. In this study, the ARIMA model is applied via MATLAB 2010b programming language that can be used to create econometric models. Statistical measurements of the most successful ARIMA models are summarized in Table 6. It can be derived that the best-fitted model to the data of the time series is the ARIMA (1,1,2), as it has the lower value of *AIC*, *BIC*, and *HQC*. Also, R^2 (simple and adjusted) statistical criteria, which are used to assess the fit of regression models, has a higher value than others. The ARIMA (1,1,2) model indicates one lag of the dependent variable, the variable being used is of first-difference stationary and two lags of the error term.

Table 3
List of all indicators used in MLR analysis.

Variable	Description
x_1	Installed capacity
x_2	Gross electricity generation
x_3	Population
x_4	Total subscribership
x_5	Export
x_6	Import
y	Net energy consumption

The residual error components in the estimates obtained with the ARIMA model are separated for LSSVM analysis. In the LSSVM analysis, data normalization is performed before the training process begins and each input space is normalized into the [-1, +1] range using Eq. (20).

$$x_n = 2\frac{x - x_{\min}}{x_{\max} - x_{\min}} - 1 \tag{20}$$

where X_n and X represent the normalized and original data, $X_{\min}$ and $X_{\max}$ represent minimum and maximum values in the original data respectively. For the LSSVM model, MATLAB 2010b platform is preferred for implementing and analysis of the algorithm. In this study, Radial Basis Function (RBF) is used as the kernel function. RBF is especially chosen due to its robust and sensitive performance factors. The RBF parameters selected in the LSSVM analysis also should reflect the distribution of input values of the training data. Otherwise, it may cause serious variations in the optimal regression line. On the other hand, there is no standard method for determining the parameters, C ve σ^2, for the LSSVM model. Comprehensive grid search and a cross-validation method are traditionally employed to obtain optimal parameter values. The grid-search technique is applied to find out the optimal parameters values which include regularization parameter C and the kernel function parameter (σ^2) which is the bandwidth in the common case of the RBF kernel. To capture the optimal values of these parameters, a grid-search technique is used in this study. The MATLAB2010b platform provides a function "*tunelssvm*" which can be used for estimation of optimal parameters. Thus, this study use the LSSVM training algorithm for residual components as follows:

Step 1. Solve LSSVM problem for residual components (Eq. (9) or 14).

Step 2. Set the solution LSSVM model as $w^{(0)}$ and $b^{(0)}$.

Step 3. Generate optimal C ve σ^2values. *(C, σ^2) = tunelssvm(x,y, 'gridsearch', 'crossvalidatelssvm', {10,'mse'})*

Step 4. Training *model = trainlssvm(x,y, C,σ^2)*

Step 5. Simulation and residual predictions (f_r).

Step 6. Additive function calculation ($f_t + f_r$) (Eq. (16)).

Step 7 Performance measurement

Step 8. Store if the performance is better than the previous one, and go to Step 3.

Step 9. Stop processing the steps if the number of epochs reaches 1100.

Step 10. Send the solution with best performance to the *Step 11.*

Step 11. Set new $w = w^{(k)}$ and $b = b^{(k)}$ and output the regressor (Eq. (7)).

This algorithm is applied to seek and acquire the two optimal parameter sets when the MSE is at its minimum. The searching process is operated with real code in the MATLAB platform. As can be shown in Fig. 4, the better results are obtained while $\sigma^2 < 0.5$ and $C < 300$. After grid search, the optimum parameters are found as

Table 4
Regression coefficients and Adjusted R2 values for linear models.

	a	b	c	d	e	g	f	Adjusted R^2
Model 1	–	1,78,801	0.027,973	–	–	–	−49567	0.988
Model 2	0.00522	0.004977	1,57,688	0.99,865	0.27,655	0.04896	−29233	0.994
Model 3	0,41,919	0,5900	−0,58,430	1,3557	–	–	17,3255	0.996

Table 5
Statistical values for different regression models.

Model	Statistical Values				
	Adjusted R^2	MSE-Training	MSE-Test	RMSE-Training	RMSE-Test
Model 1	0.988	1,96	8,29	1,4	2,88
Model 2	0.994	1,19	16,32	1,09	4,04
Model 3	0.996	6,05	6,45	2,46	2,54

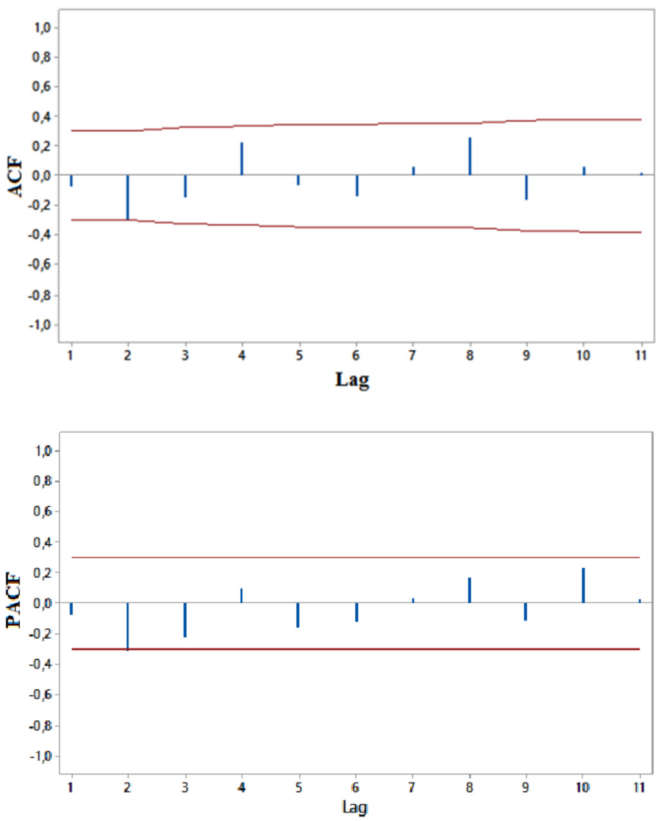

Fig. 3. ACF and PACF plots.

Table 6
Statistical criteria and measures for the ARIMA model.

ARIMA Model	R^2	Adjusted R^2	p-value	*AIC*	*BIC*	*HQC*
ARIMA (1,1,2)	0,852	0,85	0,00	−3,3705	−3,2501	−3,3256
ARIMA (1,1,0)	0,842	0,84	0,00	−3,3529	−3,2325	−3,3080
ARIMA (1,1,1)	0,839	0821	0,00	−3,3314	−3,1708	−3,2715

$\sigma^2 = 0{,}4351$ and $C = 241{,}2335$ when the MSE value is equal to 0.88 during the process of training. These optimal parameter sets are employed to establish the LSSVM model. Subsequently, it is necessary to combine the ARIMA(1,1,2) model predictions with residual values that estimated with the LSSVM model for the final results of the proposed model. Fig. 5 represents the convergence curve during the training process of the ARIMA-LSSVM hybrid approach. The training performance curve which is epochs against MSE indicates the best training performance values is at 0,174 at epoch 252 In the following section, the analyzes in *Stage III* phase of the study are presented.

4. Experimental results and observations

In this section, net electricity consumption in Turkey is forecasted by using ARIMA-LSSVM hybrid approach. The proposed approach is discussed for the forecasting of net electricity consumption based on other different methods.

4.1. Comparison of the model results

In Table 7, the MAPE (%), MSE and RMSE results are shown for all the models implemented in this study. The MAPE value in the training phase obtained by the ARIMA-LSSVM hybrid approach is 0,876%, which is lower than that of the MLR and single ARIMA models. In testing level, the ARIMA-LSSVM hybrid approach outperforms the other models in net electricity consumption forecasting and has a lower MAPE value of 1002% in contrast to MAPEs of 4,58% and 2,36% for the MLR and single ARIMA models, respectively. It can be said that the these error levels in the training and testing process might be acceptable levels of significance. Fig. 6 shows the forecasting performance of the three models for some test years. And it can clearly be seen that ARIMA-LSSVM hybrid approach showed the best overall performance for net electricity comsumption forecasting. And Fig. 7 compares the performance of the three models between 2007 and 2017.

4.2. Validation of the forecast results of the proposed hybrid approach

The country's actual consumption values and the official predictions (MAED model) are obtained from the TETC [56,57] and TURKSTAT [54] official reports. The comparison between the results of proposed model and the MAED model for the validation data, 2010 and 2011, is given in Table 8. While measuring the MAPE values, the MAED model outputs are evaluated with actual gross consumption values. Because, MAED model outputs are quite different from actual net consumption values and they are closer to actual gross consumption values. As the proposed model is designed to estimate the actual net consumption value, the proposed approach results are compared with the actual net consumptions.

MAED model produces equal results for the high and low demand cases in the year 2010. In fact, this result is unexpected for two different input series. Logically, the actual value is expected to be at any point in the middle of the high and low forecast series. Moreover, the forecasts of the MAED model in two different series for the years 2010 and 2011 are quite different from the actual gross consumption values and its error rates are higher than acceptable levels. The validation results show that the proposed hybrid approach is quite effective in determination of long-term net electricity consumption prediction and has good generalization

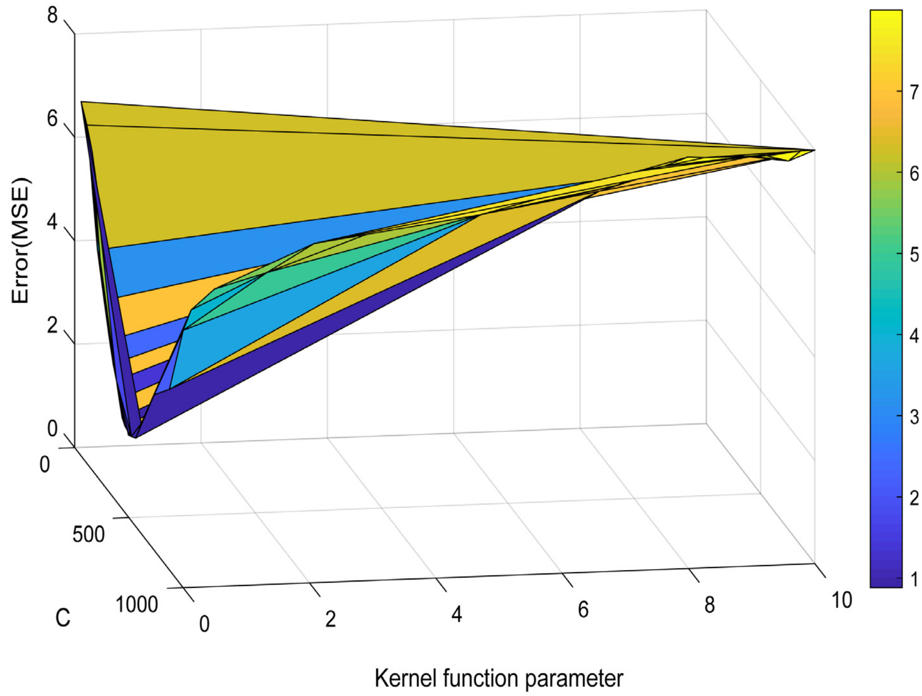

Fig. 4. Grid search for LSSVM.

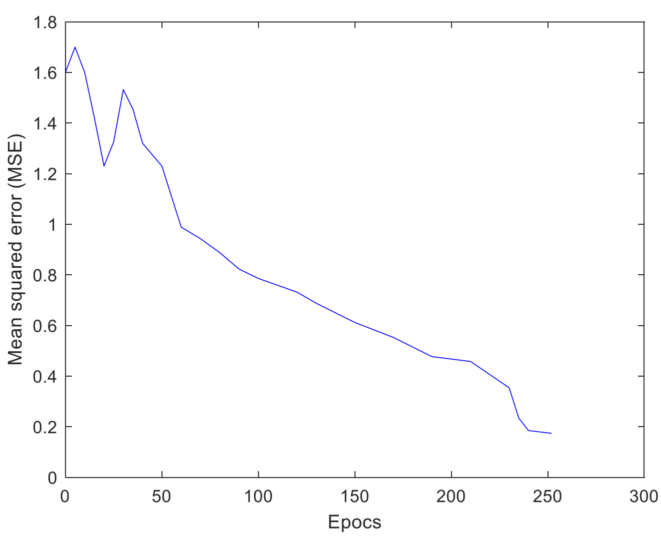

Fig. 5. The convergence curve in the training process for ARIMA-LSSVM model.

ability for two different series (Table 9). Also, better results are obtained than the studies in literature carried out by Kavaklioglu et al. [21], Kavaklioglu [30] and Hamzacebi and Es [34] for estimated net electricity consumption. When the conducted studies are analyzed in Table 9, more consistent values are obtained by the proposed model for the years 2010–2011. A very minor error value occurred between the actual values and estimated values and because while the actual values are 172,05 and 186,10 TWh for 2010–2011, the proposed model has predicted values of 170,38 and 183,41 TWh, respectively. Also, error values are determined at 0,971% and 1445% MAPE error ratio, which is lower than in previous studies using different methods, as shown in Table 9.

4.3. Comparison of the proposed hybrid approach results with official estimates

MENR and TETC use three types of estimation sets in their long-term forecasts. These are the high demand, low demand and base demand series. These series are prepared in three different scenarios. The factors such as financial crisis or economic recession and the lower-than-expected generation capacities of the new power plants to be added to the grid are taken into account in the low demand series. And the cases such as the absence of the economic crisis, full-capacity new power plants connected to the grid are taken into account in the high demand series. TETC creates a base demand series by analyzing high and low demand conditions jointly and determines all demand series and electric generating capacity additions in a period of 10 years [57]. This study employ

Table 7
The performance values of the models.

Model	Prosedure					
	Training			Test		
	MAPE (%)	MSE	RMSE	MAPE (%)	MSE	RMSE
MLR (Model 3)	4,49	8,45	2,91	4,58	10,96	3310
Single ARIMA (1,1,2)	2,60	5,47	2,34	2,36	4,86	2203
ARIMA-LSSVM hybrid approach	0,876	0174	0,418	1002	0,818	0904

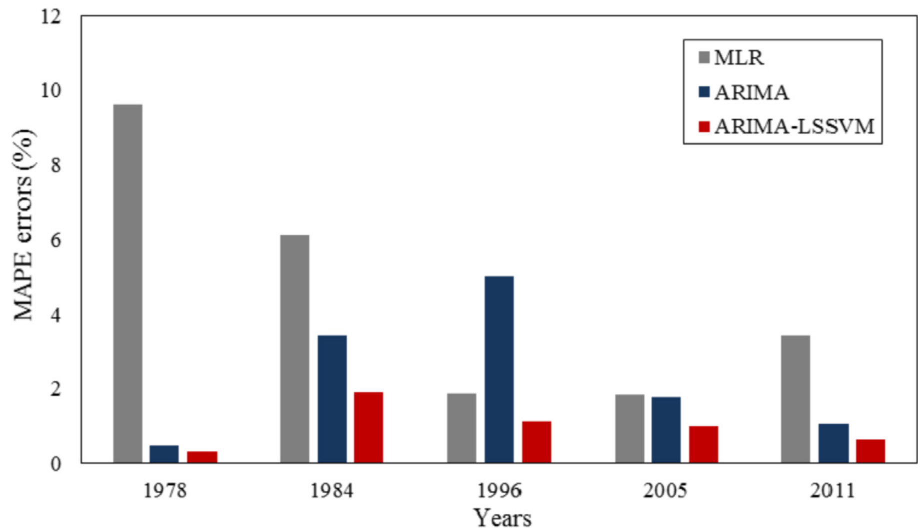

Fig. 6. Comparison of forecasting performance of different models for five testing years.

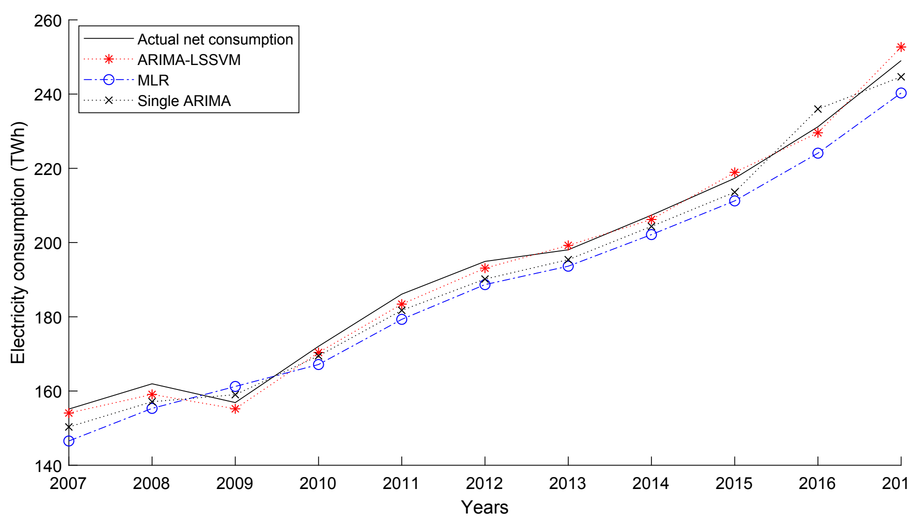

Fig. 7. Comparison of the model results with the actual values.

Table 8
Accuracy and validation of results.

	Consumptions		Official (MAED) predictions				The proposed hybrid approach	
Years	Gross realization (TWh)	Net Realization (TWh)	High demand (TWh)	MAPE (%)	Low demand (TWh)	MAPE (%)	Prediction values	MAPE (%)
2010	210,4	172,05	202,73	3,65	202,73	3,65	170,38	0,97
2011	229,3	186,10	215,91	5,84	213,88	6,72	183,41	1,45

Table 9
Comparison with similar studies in the literature.

Actual net consumption values (TWh)		Kavakliogluet al. [30].		Kavaklioglu [21]		Hamzacebi and Es [34]			This study	
Years	Values	ANN approach	MAPE (%)	SVM approach	MAPE (%)	Optimized grey model (OGM) approach		MAPE (%)	Forecasting values	MAPE (%)
						Direct OGM	Iterative OGM			
2010	172,05	182,68	6178	170,23	1058	–	–	–	170,38	0,971
2011	186,10	189,32	1730	175,05	5938	174,20,174,20		6394	183,41	1445

the official base forecasting series as a reference for the comparison of the analysis results.

In this section of the study, the proposed model results and the official estimates are compared with the gross and net realization values of electricity consumption of Turkey for the years 2000–2017. It can be clearly seen in Fig. 8 that the prediction curve of the proposed approach is closer to the actual net consumption values. However, it can also be said that the forecasts made by the MAED model have relatively high values than the actual values and have unacceptable error rates.

4.4. Future predictions of electricity consumption

In order to perform the consumption forecast with the proposed model, the values of the model inputs is firstly identified for the coming years. The installed capacity and gross power generation data used in the model are obtained from the reports published by TETC [56,57]. TURKSTAT projections [54] on Turkey's population in the coming years are utilized for population data. And, extrapolation is used to obtain the future estimates of the linearly increasing number of subscribers, since TEDC and the private electricity distribution companies have no any study on the estimated number of subscribers in the future.

Based on the consumption model by the proposed approach, future predictions of consumption are evaluated from year 2019 to year 2022, over the 4-year period. The results are given in Fig. 9 graphically and Table 10 numerically. Input variables such as installed power capacity or gross electricity generation are likely to be affected or changed by many economic and political decisions in the coming years. On the other hand, Turkey's long-term forecasts in the official report are given for the years of 2019–2022. Therefore, the future results with the proposed model are limited for this

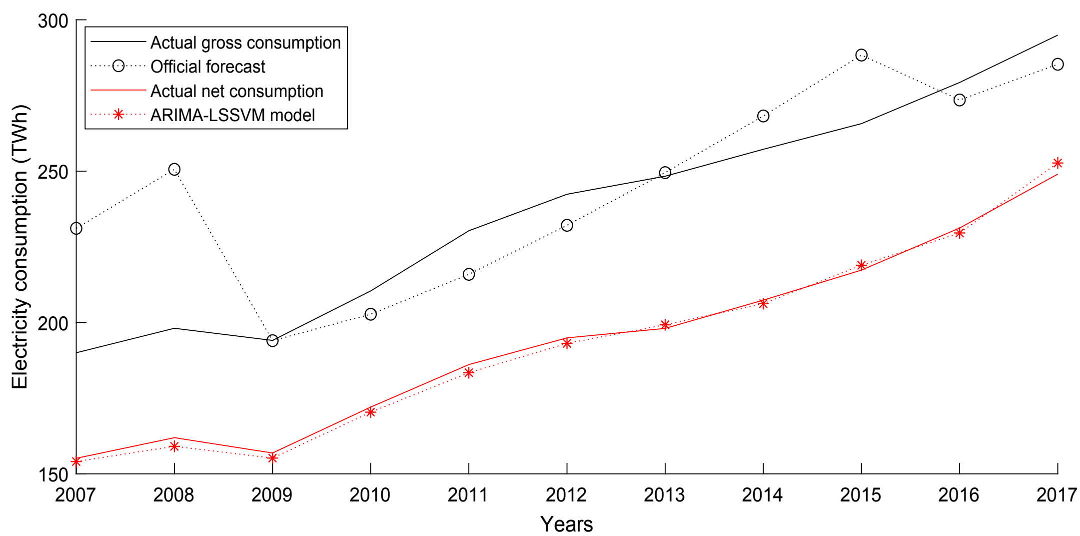

Fig. 8. Comparison of results of the MAED model (official forecasts) and the proposed ARIMA-LSSVM approach between 2007 and 2017.

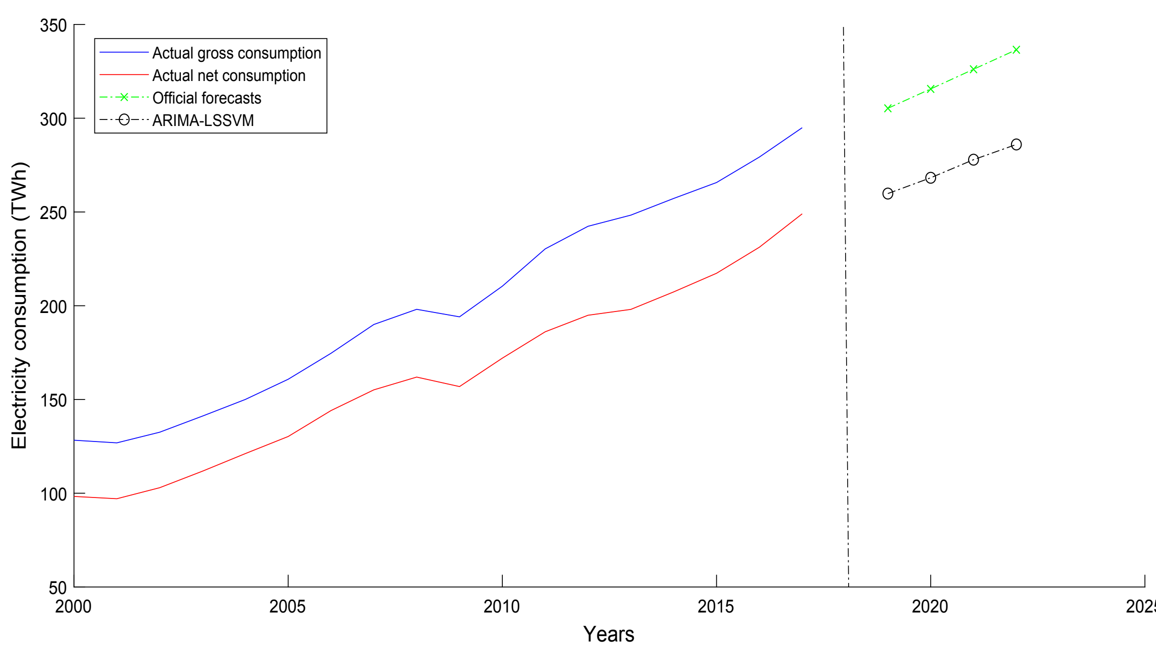

Fig. 9. Comparison of forecast results of the proposed approach and MAED model for the years 2019–2022.

Table 10
Future forecast results of the models and other inputs until 2022.

Years	Installed capacity (MW)	Gross electricity generation (GWh)	Population (million)	Total subscribership (million)	ARIMA-LSSVM model		Official forecasts (MAED model)	
					Total Consumption (TWh)	Per-Capita consumption (kWh)	Total Consumption (TWh)	Per-Capita consumption (kWh)
2019	95.167	424.470	82,89	45,07	259,82	3.135	305,29	3.683
2020	100.413	441.331	83,9	46,45	268,24	3.197	315,62	3.762
2021	104.644	462.885	84,91	47,83	277,89	3.272	326,11	3.841
2022	105.300	464.672	85,91	49,21	286,01	3.329	336,52	3.917

time period.

The values of the input data for the prediction model between 2019 and 2020 have similar upward trends. Turkey population growth has almost a linear development with time. From a population of 821,897 million in the year 2019, it is predicted to reach 85,91 million in 2022 which is a large population to plan for. In Turkey as a developing country, the total numbers of electrical subscriber indicate similar dramatic increases with the country's population. Also, the installed power and electricity production of the country will increase with Turkey's official power investment program from year to year. It is estimated that the gross electricity generation and the installed capacity of Turkey will reach to 464.672 GWh and 105.300 MW by the year 2022, respectively.

In 2017, the largest electricity consumption consuming countries are Ireland (54,4 MWh/per capita), Norway (23,7 MWh/per capita) and Bahrain (18,7 MWh/per capita). Also, some countries such as Ethiopia, Niger and Nigeria have the lowest consumption with 0.1 MWh/capita consumption. In Turkey as a developed country, the per-capita electricity consumption was 3081 kW-hours in 2017 compared to 2847 kW-hours in 2014. By 2022, the per capita electricity consumption in Turkey is estimated to reach 3917 kWh in official reports [57]. In contrast, the proposed model estimates that per capita consumption will be 3329 kWh.

The results show that the net electricity consumption of Turkey will sit on a steady upward trend until year 2022. The proposed approach predicts that the consumption will reach 286,01 TWh in the year 2022. In the last decade, while Turkey's annual electricity consumption growth is around 4–5%, ARIMA-LSSVM model shows this growth rate will be reduced by at least 0.5–1.0% in.

5. Conclusions

Electric grids should be able to handle the management of renewable sources, follow reliability issues and improving the electricity generation. This structure of the grid which is called the smart grid has mutual effects in the energy market. Furthermore, the members of the energy market request stability in market parameters to make the highest feasible gain, and this gain is almost related to the precision forecast of electric consumption and production cost. The main goal of consumption prediction is providing management and planning for future electric energy consumption or electric load. Also, this issue is very important for power market operation, power market design, power systems planning, power systems control and security of procurement. Namely, by precise forecasting, 1% decrease of the performance criteria is effective in the power system to get the range of 3–5% which can reduce the generation costs about 0.1%–0.3% [59]. Also, the higher forecasts than the real values may lead to misinformation and consequently unnecessary investment decisions for the energy investors. Electricity generation, transmission and distribution facilities require an investment of billions of dollars as a whole. The lower forecasts are another important difficulty in forecast studies. Basically, this may slow down the new energy investments and cause to potential power outages in the long term.

The main contribution of this study is that the ARIMA-LSSVM hybrid model is first used for electricity consumption forecasting and the major conclusion of this research is that the long-term electricity consumption can be modeled as a function of the proposed independent variables using the proposed approach. The proposed model has some advantages such as working with a small number of parameters, fast and simple programming and high accuracy, but the fact remains that the processes in determining the input variables and the control parameters in LSSVM model can directly affect the forecast performance. In this study, the ARIMA model is applied to capture linear patterns hidden in the electricity consumption data, while the LSSVM model is utilized to capture nonlinear patterns existing on data, resulting in a hybrid approach which can increase forecasting accuracy. In the light of the results presented in the previous sections, the testing and validation processes point out that the model is able to conduct long-term electricity consumption forecasts. The final results for Turkey demonstrate that the consumption forecasts obtained by the ARIMA-LSSVM hybrid model are close to the actual values with 1002% MAPE error ratio on the test set, whilst the official MAED model forecasts more than 11% error ratio for some years. It is clearly seen that the proposed approach is more successful than the MAED model. A striking example; the official reports of the year 2008 [55,56] estimated the gross electricity demand in Turkey would be averaged to 216,992 GWh in 2009. However, the gross electricity consumption in Turkey was 194,080 GWh in 2009. There was an unacceptable deviation of 11.81% in the official estimates. 22,912 GWh of energy, which is a forecast error, approximately equals to the annual total production of Turkey's public hydroelectric power plants that have higher installed capacities such as Ataturk (2405 MW), Karakaya (1800 MW) and Keban (1330 MW). Therefore, advanced intelligent forecast tools are definitely needed in Turkey. The proposed approach can be used by the Turkey's related organizations to forecast future electric consumption values so as to make successful future planning. This model can also be used for other developing countries.

Acknowledgement

The author would like to thank anonymous referees and the Editors for their valuable and constructive comments on the paper.

References

[1] Bianco V, Manca O, Nardini S, Minea AA. Analysis and forecasting of nonresidential electricity consumption in Romania. Appl Energy 2010;87:3584–90. https://doi.org/10.1016/j.apenergy.2010.05.018.
[2] Ministry of energy and natural resources (MENR) Turkey. Electric; 2018. http://www.enerji.gov.tr/en-US/Pages/Electricity. [Accessed 21 October 2018].
[3] Bianco V, Manca O, Nardini S. Electricity consumption forecasting in Italy using linear regression models. Energy 2009;34:1413–21. https://doi.org/10.1016/j.energy.2009.06.034.
[4] Erdogdu E. Electricity demand analysis using cointegration and ARIMA modeling: a case study of Turkey. Energy Pol 2007;35(2):1129–46. https://

doi.org/10.1016/j.enpol.2006.02.013.
[5] Tunc M, Camdali U, Parmaksizoglu C. Comparison of Turkey's electrical energy consumption and production with some European countries and optimization of future electrical power supply investments in Turkey. Energy Pol 2006;34(1):50–9. https://doi.org/10.1016/j.enpol.2004.04.027.
[6] Pappas SS, Ekonomou L, Karamousantas DC, Chatzarakis G, Katsikas S, Liatsis P. Electricity demand loads modeling using Autoregressive Moving Average (ARMA) models. Energy 2008;33:1353–60. https://doi.org/10.1016/j.energy.2008.05.008.
[7] Ediger VS, Akar S. ARIMA forecasting of primary energy demand by fuel in Turkey. Energy Pol 2007;35(3):1701–8. https://doi.org/10.1016/j.enpol.2006.05.009.
[8] Egelioglu F, Mohamad AA, Guven H. Economic variables and electricity consumption in Northern Cyprus. Energy 2001;26:355–62. https://doi.org/10.1016/S0360-5442(01)00008-1.
[9] Crompton P, Wu Y. Energy consumption in China: past trends and future directions. Energy Econ 2005;27(1):195–208. https://doi.org/10.1016/j.eneco.2004.10.006.
[10] Nai-ming X, Chao-qing Y, Ying-jie Y. Forecasting Chinas energy demand and self-sufficiency rate by grey forecasting model and Markov model. Int J Elec Power 2015;66:1–8. https://doi.org/10.1016/j.ijepes.2014.10.028.
[11] Hamzacebi C, Es HA. Forecasting the annual electricity consumption of Turkey using an optimized grey model. Energy 2014;70:165–71. https://doi.org/10.1016/j.energy.2014.03.105.
[12] Akay D, Atak M. Grey prediction with rolling mechanism for electricity demand forecasting of Turkey. Energy 2007;32(9):1670–5. https://doi.org/10.1016/j.energy.2006.11.014.
[13] Jasinski T. Modeling electricity consumption using nighttime light images and artificial neural networks. Energy 2019;179:831–42. https://doi.org/10.1016/j.energy.2019.04.221.
[14] Đozić UDJ, Urošević BDG. Application of artificial neural networks for testing long-term energy policy targets. Energy 2019;174:488–96. https://doi.org/10.1016/j.energy.2019.02.191.
[15] Azadeh A, Babazadeh R, Asadzadeh SM. Optimum estimation and forecasting of renewable energy consumption by artificial neural networks. Renew Sustain Energy Rev 2013;27:605–12. https://doi.org/10.1016/j.rser.2013.07.007.
[16] Beccali M, Cellura M, Lo Brano V, Marvuglia A. Short-term prediction of household electricity consumption: assessing weather sensitivity in a Mediterranean area. Renew Sustain Energy Rev 2008;12(8):2040–65. https://doi.org/10.1016/j.rser.2007.04.010.
[17] Esen H, Ozgen F, Esen M, Sengur A. Modelling of a new solar air heater through least-squares support vector machines. Expert Syst Appl 2009;36(7):10673–82. https://doi.org/10.1016/j.eswa.2009.02.045.
[18] Sulaiman MH, Mustafa MW, Shareef H, Khalid SNA. An application of artificial bee colony algorithm with least squares support vector machine for real and reactive power tracing in deregulated power system. Int J Elec Power 2012;37(1):67–77. https://doi.org/10.1016/j.ijepes.2011.12.007.
[19] Ekici BB. A least squares support vector machine model for prediction of the next day solar insolation for effective use of PV systems. Measurement 2014;50:255–62. https://doi.org/10.1016/j.measurement.2014.01.010.
[20] Bessedik SA, Hadi H. Prediction of flashover voltage of insulators using least squares support vector machine with particle swarm optimization. Elec Power Syst Res 2013;104:87–92. https://doi.org/10.1016/j.epsr.2013.06.013.
[21] Kavaklioglu K. Modeling and prediction of Turkey's electricity consumption using support vector regression. Appl Energy 2011;88:368–75. https://doi.org/10.1016/j.apenergy.2010.07.021.
[22] Barak S, Sadegh SS. Forecasting energy consumption using ensemble ARIMA-ANFIS hybrid algorithm. Int J Elec Power 2016;82:92–104. https://doi.org/10.1016/j.ijepes.2016.03.012.
[23] Tutun S, Chou C, Caniyilmaz E. A new forecasting framework for volatile behaviour in net electricity consumption: a case study in Turkey. Energy 2015;93:2406–22. https://doi.org/10.1016/j.energy.2015.10.064.
[24] Gurbuz F, Ozturk C, Pardalos P. Prediction of electricity energy consumption of Turkey via artificial bee colony: a case study. Energy Syst 2013;4(3):289–300.
[25] Geoffrey KFT, Kelvin KWY. Predicting electricity energy consumption: a comparision of regression analysis, decision tree and neural Networks. Energy 2007;32:1761–8. https://doi.org/10.1016/j.energy.2006.11.010.
[26] Bilgili M, Sahin B, Yasar A, Simsek E. Electric energy demands of Turkey in residential and industrial sectors. Renew Sustain Energy Rev 2012;16(1):404–14. https://doi.org/10.1016/j.rser.2011.08.005.
[27] Kaytez F, Taplamacioglu MC, Cam E, Hardalac F. Forecasting electricity consumption: a comparison of regression analysis, neural networks and least squares support vector machines. Int J Elec Power 2015;67:431–8. https://doi.org/10.1016/j.ijepes.2014.12.036.
[28] Ayvaz B, Kusakci AO. Electricity consumption forecasting for Turkey with nonhomogeneous discrete grey model. Energy Sources B Energy Econ Plann 2017;12(3):260–7. https://doi.org/10.1080/15567249.2015.1089337.
[29] Sozen A, Akcayol MA, Arcaklioglu E. Forecasting net energy consumption using artificial neural network. Energ Sources Part B 2006;1(2):147–55. https://doi.org/10.1080/009083190881562.
[30] Kavaklioglu K, Ceylan H, Ozturk HK, Canyurt OE. Modeling and prediction of Turkey's electricity consumption using artificial neural Networks. Energy Convers Manag 2009;50:2719–27. https://doi.org/10.1016/j.enconman.2009.06.016.
[31] Sozen A, Arcaklioglu E. Prospects for future projections of the basic energy sources in Turkey. Energy Sources B: Econ Plan Pol 2007;2:183–201.
[32] Kankal M, Akpınar A, Kömürcü MI, Ozsahin TS. Modeling and forecasting of Turkey's energy consumption using socio-economic and demographic variables. Appl Energy 2011;88:1927–39. https://doi.org/10.1016/j.apenergy.2010.12.005.
[33] Hamzacebi C. Forecasting of Turkey's net electricity energy consumption on sectoral bases. Energy Pol 2007;35. https://doi.org/10.1016/j.enpol.2006.03.014. 2009–16.
[34] Hamzacebi C, Es HA. Forecasting the annual electricity consumption of Turkey using an optimized grey model. Energy 2014;70:165–71. https://doi.org/10.1016/j.energy.2014.03.105.
[35] Sozen A, Arcaklioglu E. Prediction of net energy consumption based on economic indicators (GNP and GDP) in Turkey. Energy Pol 2007;35:4981–92. https://doi.org/10.1016/j.enpol.2007.04.029.
[36] Yildiz B, Bilbao JI, Sproul AB. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. Renew Sustain Energy Rev 2017;73:1104–22. https://doi.org/10.1016/j.rser.2017.02.023.
[37] Kallu RR, Keffeler ER, Watters RJ, Agharazi A. Development of a multivariate empirical model for predicting weak rock mass modulus. Int J Min Sci Technol 2015;25(4):545–52. https://doi.org/10.1016/j.ijmst.2015.05.005.
[38] Jia TG, Liu J. Stability of mine ventilation system based on multiple regression analysis. Int J Min Sci Technol 2009;19(4):463–6. https://doi.org/10.1016/S1674-5264(09)60086-7.
[39] Box GEP, Jenkins GM. Time series analysis forecasting and control. 4nd ed. New Jersey: John Wiley&Sons; 1976.
[40] Rosca RE. Stationary and non-stationary time series. USV Ann Econ Public Administr 2011;10:177–86.
[41] Yuan C, Liu S, Fang Z. Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM (1, 1) model. Energy 2016;100:384–90. https://doi.org/10.1016/j.energy.2016.02.001.
[42] Vapnik V. The nature of statistical learning theory. New York: Springer-Verlag; 1995.
[43] Vapnik V. Statistical learning theory. New York: John Wiley; 1998.
[44] Suykens JAK, Vandewalle J. Nonlinear modeling: advanced black-box techniques. Boston: Kluwer Academic Publishers; 1998.
[45] Christianini N, Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods. Cambridge: Cambridge University Press; 2000.
[46] Suykens JAK, Van Gestel T, De Brabanter J, De Moor B, Vandawalle J. Least squares support vector machines. Singapore: World scientific Publishing; 2002.
[47] Pahasa J, Ngamroo I. A heuristic training-based least squares support vector machines for power system stabilization by SMES. Expert Syst Appl 2011;38:13987–93. https://doi.org/10.1016/j.eswa.2011.04.206. 38.
[48] Campbell C. An introduction to kernel methods radial basis function network: design and applications. Berlin: Spronger Verlag; 2000.
[49] Mercer J. Function of positive and negative type and their connection with the theory of integral equations. Philos. Trans. of the Royal Society A. 1909;209:415–46. https://doi.org/10.1098/rsta.1909.0016.
[50] Suykens JAK, Vandewalle J. Least squares support vector machine classifiers. J. Neural Process. Lett. 1999;9(3):293–300.
[51] Xie G, Wang S, Zhao Y, Lai KK. Hybrid approaches based on LSSVR model for container throughput forecasting: a comparative study. Appl Soft Comput 2013;13:2232–41. https://doi.org/10.1016/j.asoc.2013.02.002.
[52] Karagol E, Erbaykal E, Ertugrul HM. Economic growth and electricity consumption in Turkey: a bound test approach. Dogus Univ Journal 2007;8(1):72–80.
[53] Altınay G, Karagol E. Electricity consumption and economic growth: evidence from Turkey. Energy Econ 2005;27:849–56. https://doi.org/10.1016/j.eneco.2005.07.002.
[54] Turkish Statistical Institute (TURKSTAT). Population, annual population growth rate and projections. http://www.tuik.gov.tr/VeriBilgi.do?alt_id=39. [Accessed 3 February 2018].
[55] Turkish Electricity Transmission Co (TETC). Annual development of Turkey's gross electricity generation –imports-exports and demand. http://www.teias.gov.tr/T%C3%BCrkiyeElektrik%C4%B0statistikleri/istatistik2015/index.htm. [Accessed 15 September 2018].
[56] Turkish Electricity Transmission Co (TETC). 10 years of Turkish electricity production capacity projection 2017-2021. http://www.teias.gov.tr/Eng/ApkProjection/Capacity%20Projection%202009-2018.pdf. [Accessed 25 August 2018].
[57] Turkish Electricity Transmission Co (TETC). 10 years of Turkish electricity production capacity projection 2018-2022. https://www.teias.gov.tr/sites/default/files/2018-09/Kapasite_Projeksiyonu_2018_2022.pdf. [Accessed 12 November 2018].
[58] Turkish Electricity Distribution Co (TEDC). Activity reports. http://www.tedas.gov.tr/BilgiBankasi/Sayfalar/IstatistikiBilgiler.aspx. [Accessed 25 January 2018].
[59] Liu Y, Wei W, Noradin G. Electricity load forecasting by an improved forecast engine for building level consumers. Energy 2017;139:18–30. https://doi.org/10.1016/j.energy.2017.07.150.