Energy Economics 80 (2019) 937–949

Contents lists available at ScienceDirect

Energy Economics

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

Using machine learning tools for forecasting natural gas consumption in the province of Istanbul☆

Omer Faruk Beyca [a], Beyzanur Cayir Ervural [d], Ekrem Tatoglu [b,*], Pinar Gokcin Ozuyar [c], Selim Zaim [d]

[a] Istanbul Technical University, Faculty of Management, Macka, Istanbul 34722, Turkey
[b] Ibn Haldun University, School of Business, Basaksehir, Istanbul 34494, Turkey
[c] Ozyegin University, Faculty of Business, Cekmekoy, Istanbul 34794, Turkey
[d] Istanbul Sehir University, College of Engineering and Natural Sciences, Dragos, Kartal, Istanbul 34865, Turkey
[d] Konya Food and Agriculture University, Faculty of Engineering and Architecture, Meram, Konya, 42080, Turkey

ARTICLE INFO

Article history:
Received 23 April 2018
Received in revised form 10 February 2019
Accepted 9 March 2019
Available online 13 March 2019

JEL classification:
C45
C53
E17
Q47

Keywords:
Natural gas forecasting
Machine learning
Artificial neural network
Support vector regression
Emerging countries
Istanbul

ABSTRACT

Commensurate with unprecedented increases in energy demand, a well-constructed forecasting model is vital to managing energy policies effectively by providing energy diversity and energy requirements that adapt to the dynamic structure of the country. In this study, we employ three alternative popular machine learning tools for rigorous projection of natural gas consumption in the province of Istanbul, Turkey's largest natural gas-consuming mega-city. These tools include multiple linear regression (MLR), an artificial neural network approach (ANN) and support vector regression (SVR). The results indicate that the SVR is much superior to ANN technique, providing more reliable and accurate results in terms of lower prediction errors for time series forecasting of natural gas consumption. This study could well serve a useful benchmarking study for many emerging countries due to the data structure, consumption frequency, and consumption behavior of consumers in various time-periods.

1. Introduction

Natural gas has always been one of the key energy sources in the world and has numerous advantages compared to other fossil fuels in terms of its cleanliness, affordability, ease of storage, and transportability (Liang et al., 2012). Due to limited natural gas reserves, accurate energy planning studies and optimal energy projections are crucial for long-term utilization and sustainability of energy (Tolba and Biswas, 2013). The worldwide natural gas consumption has grown considerably in recent years because of the rapidly growing population, industrialization, and urbanization (International Energy Agency, 2017). Global natural gas consumption is envisaged to rise from 120 trillion cubic feet (cf) in 2012 to 203 trillion cf. in 2040 (EIA, 2015). As an energy source, natural gas has by far the largest share constituting 23.8% of the world's primary energy consumption.

Turkey is acknowledged as one of the big emerging countries in terms of economic growth featuring the world's 17th leading economy with a gross domestic product (GDP) of USD 820 billion (IEA, 2017; World Bank, 2016). In terms of energy demand, it is also one of the most rapidly growing countries in the world along with China (Tatoglu et al., 2015). Based on official statistics, Turkey imported 99.2% of natural gas from various countries in 2015. Natural gas comprised 37.8% of total electricity production in 2015 (Republic of Turkey Ministry of Foreign Affairs, 2017).

Previous research has revealed that in mega-cities like Istanbul, high level of urbanization indicates a close interaction between GDP and energy consumption (Zhang et al., 2013). With over 15 million residents, Istanbul has turned into a global mega-city, featuring eighth within the top ten metro-regions in terms of population size (BP Report, 2017). Commensurate with worldwide developments, this mega-city's

☆ We would like to express our sincere thanks to the journal editor Professor Richard S.J. Tol and two anonymous reviewers for their invaluable comments and feedback on an earlier version of this study.We are also grateful to Dr. Mahmut Akşit from IGDAŞ (Istanbul Gas Distribution Industry and Trade Inc.) for providing the natural gas consumption data of the province of Istanbul.

* Corresponding author at: Ibn Haldun University, School of Business, Department of Management, Basaksehir, Istanbul 34494, Turkey.
E-mail addresses: beyca@itu.edu.tr (O.F. Beyca), beyzanur.ervural@gidatarim.edu.tr (B.C. Ervural), ekrem.tatoglu@ihu.edu.tr (E. Tatoglu), pinar.ozuyar@ozyegin.edu.tr (P.G. Ozuyar), selimzaim@sehir.edu.tr (S. Zaim).

https://doi.org/10.1016/j.eneco.2019.03.006
0140-9883/

economy provides a new vision to be able to seize the competitive opportunities of the region. Istanbul is also the engine of economic growth in Turkey, and becomes the key industrial, logistics and financial center, accounting for nearly 40% of the country's GDP and harboring the sizable amount of foreign investments (Energy and Renewables - Invest in Turkey, 2018). Istanbul is also the largest province comprising nearly 18% of share in the total natural gas demand in Turkey, making it the greatest city of natural gas marketing in Turkey.

Under these circumstances, the subject of energy forecasting has received growing attention from both scholars and practitioners to maintain energy security and to keep uninterrupted energy flow for the future. The demand and supply of energy is one of the key subjects that is always at the forefront of the agenda. For the greater metro-region of Istanbul, the city's rising GDP growth in conjunction with its superb location contributes to a massive growth in energy consumption. Thus, policy makers have to seriously consider threats that may be posed by institutional voids and market imperfections.

The extant literature emphasizes that forecasting research on primary energy demand nationwide and also at the provincial level in Istanbul are essential vehicles of providing useful information on the managerial decision-making and policy development process (Melikoglu, 2013). However, it is imperative to assist local policy makers with advanced aiding techniques for making superior decisions to assess several energy-related alternatives for fulfilling growing energy demand (Dong et al., 2013; Popoola et al., 2018). The planning for the consumption of natural gas is one of such tools. The natural gas consumption has grown in the metro-region of Istanbul as well as Turkey over the past two decades and has turned out to be the highest growing primary energy source in Turkey (Hacisalihoglu, 2008). Local policy makers in Istanbul, as in other mega-cities, should formulate creative strategies to manage this increasing energy demand and also effectively deal with the uninterrupted and dynamic demand for inexpensive supply of natural gas. Seemingly, rigorous estimation of natural gas consumption and also of other energy sources are vital for devising effective policies for a competitive mega-city, as most of the purchasing agreements rely on managerial decisions (de Almeida et al., 2004; Wiser et al., 2004). To avoid natural gas crises, well-designed projections and predetermined investment decisions should be implemented properly at both country and regional levels.

In the light of these explanations, efficient utilization of energy resources help energy authorities accurately predict energy need in the future. The studies on predictions of natural gas consumption were particularly performed at various levels: (a) global, national, and provincial; (b) industrial and residential; and (c) commercial and residential (Soldo, 2012; Potočnik et al., 2014). In this study, we investigate consumption of natural gas at the provincial level that combines all stakeholders from commercial, industrial, and residential sectors in Istanbul.

This study offers three main contributions to the extant research on the forecasting of natural gas consumption. First, it provides a simple and accurate prediction model for natural gas consumption. The proposed models can be adapted to different types of natural gas consumption areas including residential, industrial, and nonresidential natural gas consumption in addition to forecasting of electricity/energy consumption. Secondly, the present study reveals that drivers of natural gas consumption behavior change relying on the use of machine learning tools. Finally, forecasting natural gas consumption of Istanbul, which is Turkey's largest natural gas-consuming mega-city, could well serve a useful benchmarking study for many emerging countries due to the data structure, consumption frequency, and consumption behavior of consumers in various time-periods.

The rest of this study is organized as follows. The next section presents brief information on the natural gas sector in Turkey. Section 3 provides a review of earlier research on forecasting of natural gas demand and overview of the methods adopted as well as the descriptions of the proposed procedure. Application of the model along with the findings is presented in Section 4. The conclusion and the managerial implications are set out in the final section.

2. Natural gas sector in Turkey

Natural gas in Turkey was first explored in Hamitabat and Kumrular by the Turkish Petroleum Corporation (TPAO) in 1970, and was first utilized at the Pınarhisar Cement Plant in 1976. The demand for natural gas has increased exponentially since the late 1980s. To fulfill the growing energy demand for natural gas, Turkey signed its first agreement with the former Soviet Union in 1984 for the delivery of natural gas, indicating a consideration of growing environmental concerns and also fulfilling the policy of diversifying energy resources. Since then, additional purchase agreements/deals have been conducted rapidly to respond natural gas requirements.

Based on the report by the Energy Information Administration of the U.S., Turkey ranked 9th in the world with the import of 48.2 billion cubic meters of natural gas in 2015 (International Energy Agency, 2017). Turkey entirely depends on foreign natural gas resources (99.2%) because of the scarcity of domestic natural gas reserves. Fig. 1 shows Turkey's natural gas imports over the period of 2005–2015.

The Turkish Petroleum Pipeline Corporation (BOTAŞ) is Turkey's state-held gas trade and transmission company that has exclusive rights on import, transmission, distribution, and sales activities throughout the country. Of the total imports of 48.2 bcm in 2015, BOTAŞ imports natural gas mostly from the Russian Federation (55.1%), followed by Iran (16.16%), Azerbaijan (12.74%), Algeria (8.09%), Nigeria (2.6%), and others based on long-term contracts, as shown in Fig. 2.

Fig. 3 displays the sectoral distribution of natural gas consumption in 2015 (Energy Governance in Turkey, 2015). Natural gas is mainly utilized in electricity generation (39.61%), industry (29.1%) and residential consumption (22.92%), government offices and commercial sites (6.02%), and others (2.36%).

In terms of the distribution of natural gas consumption by cities, Istanbul accounts for the highest share of natural gas distribution (18%) in Turkey as of 2015 followed by Izmir (9%), Kocaeli (8%), Ankara (8%), Sakarya (8%), and Bursa (7%), as seen in Fig. 4. It is readily apparent from Fig. 4 that highly industrialized and populated cities are those with the largest levels of natural gas consumption.

Turkey is geographically located between south-north and east-west energy corridors, and this unique location is envisaged to make Turkey become a powerful energy hub in view of the energy security in the region. Some of the ongoing important gas pipeline projects are as follows: Turkey-Greece Interconnector (ITG), Baku-Tbilisi-Erzurum Natural Gas Pipeline (BTE), Blue Stream Natural Gas Pipeline, Western Route (Russia-Turkey Natural Gas Pipeline), and lastly Iran-Turkey Natural Gas Pipeline. The details of these projects can be found on the official website of Turkey's Ministry of Foreign Affairs (Republic of Turkey Ministry of Foreign Affairs, 2017). Additionally, the Trans-Anatolian Natural Gas Pipeline Project (TANAP) and the Turkey-

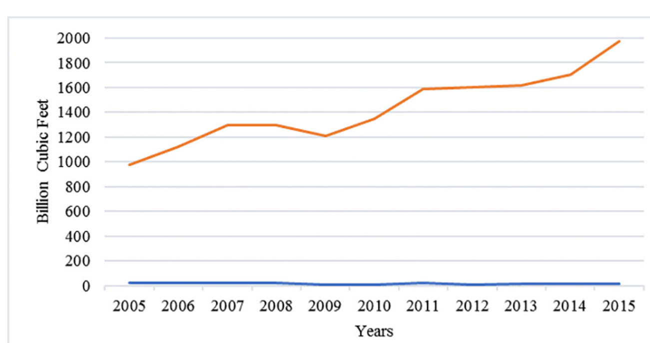

Fig. 1. Turkey's natural gas imports (International Energy Agency, 2017).

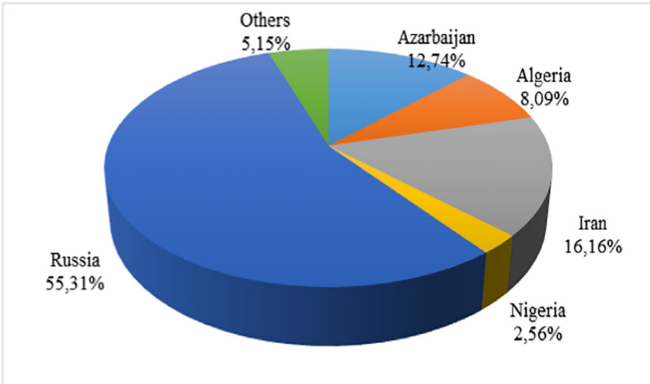

Fig. 2. Imported natural gas by country (2015).

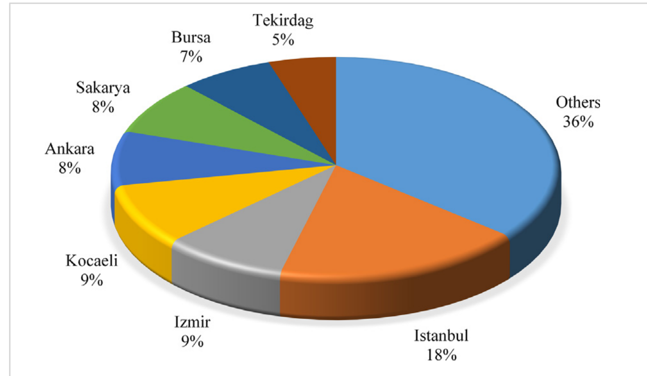

Fig. 4. Natural gas consumption by cities (2015).

Bulgaria Interconnector (ITB) are other remarkable ongoing projects aimed to transport natural gas from the Middle East and Caspian Basin to Europe through new optional paths.

3. Background literature

There are numerous techniques utilized in the previous literature for energy demand forecasting and also particularly for natural gas consumption (Dilaver and Hunt, 2011; Soldo, 2012; Taşpinar et al., 2013; Rodger, 2014; Yu and Xu, 2014; Khan, 2015; Szoplik, 2015; Baldacci et al., 2016; Ervural et al., 2016; Wang et al., 2016; Zeng and Li, 2016; Panapakidis and Dagoumas, 2017), ranging from artificial neural network (ANN), autoregressive integrated moving average (ARIMA), time series techniques to support vector machines (SVM) and fuzzy logic (FL) methods. Table 1 provides a summary of existing studies on forecasting of natural gas consumption.

Several methods have also been applied to specifically forecast Turkey's natural gas demand at both national and provincial levels. Relying on a broader perspective, Karadede et al. (2017) proposed a breeder hybrid algorithm which consists of nonlinear regression-based breeder genetic algorithm and simulated annealing for natural gas demand forecasting in Turkey. Ozdemir et al. (2016) developed a hybrid GA-simulated annealing algorithm based on linear regression to forecast the natural gas demand for Turkey. Boran (2015) applied a grey prediction with a rolling mechanism method to estimate natural gas consumption in Turkey. Taşpinar (2015) assessed residential natural gas consumption to model air pollution in Turkey using multi-parameter time series modeling methods. Melikoglu (2013) developed a logistic model based on GDP, purchasing power parity per capita, demographics, and population change factors to estimate natural gas demand for Turkey during the period of 2013–2030. Taşpinar et al. (2013) applied a multilayer ANN method with time series approach in order to predict short-term natural gas consumption in Turkey. In doing this, some meteorological data such as temperature and moisture are also included to algorithm to increased prediction power. Olgun et al. (2012) implemented the SVM and ANN model to forecast natural gas demand for Turkey. Kaynar et al. (2011) applied ANN, ANFIS, and ARIMA models to estimate natural gas consumption for Turkey.

At the provincial level, Akpinar and Yumusak (2016) attempted to estimate natural gas consumption for the province of Sakarya in Turkey by utilizing time series decomposition. Ervural et al. (2016) developed a forecasting method integrating genetic algorithm (GA) and ARMA methods to estimate natural gas consumption in Istanbul. In a much recent study, Özmen et al. (2018) predicted natural gas consumption of the capital city of Ankara in Turkey utilizing the algorithms of multivariate adaptive regression splines and conic multivariate adaptive regression splines.

In this study, three alternative approaches are employed for rigorous estimation of natural gas consumption within the province of Istanbul, namely, MLR, SVR and ANN techniques (Hacisalihoglu, 2008; Erdogdu, 2010; Demirel et al., 2012; Taşpinar et al., 2013). The key thrust for choosing these prediction methods is that they are promising and powerful machine learning tools due to their easy adaptation abilities to dynamic data structures (Sun et al., 2017). For instance, the SVR method is seemingly a new and capable forecasting technique and has been effectively applied to deal with time series problems (Kavaklioglu, 2011). As for the SVR method, four kernel functions including linear kernel, polynomial kernel, multilayer perception kernel and radial basis function are selected relying on three performance measures (viz., fitting accuracy, prediction accuracy and overall accuracy) (Vapnik, 1998).

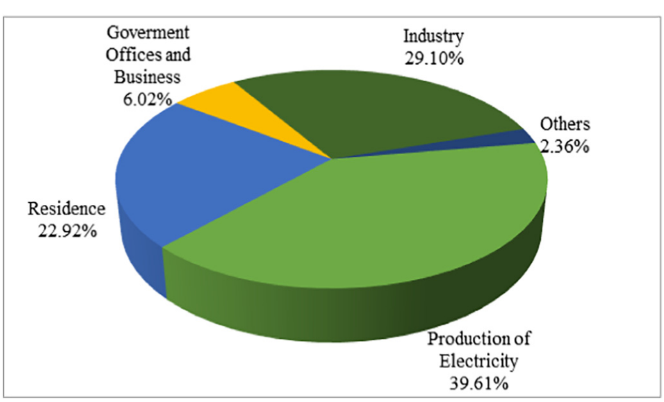

Fig. 3. Natural gas consumption by sector (2015).

3.1. An overview of the methods employed

In this section, three forecasting models are applied: traditional MLR, ANN, and SVR models. SVR was chosen due to its growing popularity and solution ability in the complex. In addition, a traditional MLR model was employed to compare and demonstrate predictive performances of the ANN and SVR methods. In the ensuing sections, a detailed description of each of these techniques is given.

3.1.1. Regression model

Regression analysis is one of the most commonly used traditional forecasting/prediction techniques to identify the causality between the dependent and independent (explanatory) variables. The association between dependent variable and predictor variables is formulated as a linear model in Eq. (1) (Chatterjee and Hadi, 2000).

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \varepsilon \quad (1)$$

Table 1
A summary of existing studies on forecasting of natural gas consumption.

Author(s)	Goal	Method	Findings
Sánchez-Úbeda and Berzosa (2007)	Forecasting industrial end-use natural gas consumption	Decomposition approach	A novel prediction model is presented that provides forecasting in a medium-term horizon with a very high resolution.
Elragal (2009)	Forecasting of daily consumption of natural gas	ANN using PSO	The proposed method provides more accurate result than single ANN.
Wadud et al. (2011)	Forecasting of natural gas demand in Bangladesh.	Dynamic econometric model	Total gas demand model is constructed as a function of per GDP, price, and population variables.
Iskin et al. (2012)	Exploring renewable energy pricing	Analytic network process	Integrates social, technical, environmental and economic factors in order to determine renewable energy pricing.
Karimi and Dastranj (2014)	Prediction of natural gas consumption in Iran	ANN with GA	The proposed ANN-GA model is a powerful tool that can be consistently and accurately used to forecast daily natural gas consumption.
Yu and Xu (2014)	Short term gas load forecasting in Shanghai	A combined GA-BP neural network model	The combined GA–BP model improved by modified additional momentum factor is superior to others.
Askari et al. (2015)	Forecasting of natural gas consumption	A clustering-based forecasting algorithm for fuzzy time series	The proposed is faster and more accurate than the conventional algorithms.
Azadeh et al. (2015)	Prediction of long term gas consumption in Iran	Adaptive neuro fuzzy inference system	The proposed method is found better in terms of applicability and superiority than conventional approaches.
Fagiani et al. (2015)	Forecasting of domestic water and natural gas demand using database (AMPds)	GP and combination with extended Kalman filter	A strong relationship is found between natural gas and water consumptions.
Khan (2015)	Natural gas demand forecasting in Pakistan	Econometric model	The study mainly examines the impacts of price, income and cross price elasticities on natural gas demand for residential, commercial, industrial transport, and power sector.
Izadyar et al. (2015)	Forecasting of monthly residential natural gas consumption in Iran.	Extreme learning machine (ELM), ANNs and genetic programming (GP)	ELM predictions are superior then ANN and GP.
Szoplik (2015)	Forecasting natural gas demand of Szczecin in Poland.	ANNs	The model considers calendar and weather factors, which have an impact on gas consumption.
Cardoso and Cruz (2016)	Prediction of natural gas consumption in Brazil	ARIMA, ANNs and a hybrid methodology	The results obtained with the ANN and hybrid methods provide better results than the ARIMA method.
Liu et al. (2016)	Primary energy consumption in Spain	Grey models combined with input-output models	The proposed model has higher forecasting accuracy than Grey models and other combination methods.
Naji et al. (2016)	Estimation of energy consumption in buildings	SVR and ANFIS	ANFIS predictions are better than SVR results in terms of RMSE.
Shaikh and Ji (2016)	Medium- to long-term natural gas consumption forecasting in China	Logistic model	To improve estimation accuracy, Levenberg-Marquardt algorithm is adopted to get the parameters of the logistic model.
Shakouri and Kazemi (2016)	Residential and commercial energy demand forecasting in Iran	ARMAX model	Regardless of de-subsidization, the energy demand grows by an average rate of around 3% annually.
Alcaraz and Villalvazo (2017)	Estimation of the natural gas shortage	Econometric analysis based on panel data	The natural gas shortage reduced Mexican GDP annual growth rate by 0.28 percentage points in the second quarter of 2013.
Panapakidis and Dagoumas (2017)	Day-ahead natural gas demand forecasting of Greece	Wavelet transform, GA, ANFIS and feed-forward neural network (FFNN)	ANFIS with a FFNN model provides high accuracy.
Scarpa and Bianco (2017)	Long-term natural gas consumption in Italian residential sector	Regression algorithm and Kalman filter method.	The explanatory variables, namely heating degree days, price of natural gas and gross domestic product (GDP) per capita, were used to estimate residential natural gas consumption. Standard regression technique was validated by using a Kalman filter.
Chen et al. (2018)	Forecasting day-ahead high-resolution natural-gas demand in Germany	A functional autoregressive model with exogenous variables (FARX).	FARX model provides stronger performance compared to alternative approaches.
Ding (2018)	Forecasting China's natural-gas demand.	A novel self-adapting intelligent grey model.	The experimental results indicate that the performance of the new model is better than the competing models.
Fan et al. (2018)	Forecasting the demand of natural gas in China from 2011 to 2017.	A combined model based on the grey model and the self-adapting intelligent grey model with a genetic algorithm.	The proposed GM-S-SIGM-GA model is superior to other single forecasting models in terms of MAPE.
Liu et al. (2018)	Investigation of the consumption of natural gas in Chinese households.	The generalized least square method	The price and income affect both the volume of natural gas and the average household consumption.
Wang et al. (2018)	A high accuracy natural gas consumption model in China.	A hybrid model based on the PSO-wavelet neural network (WNN).	The PSO-WNN model well forecasts gas consumption as reflected by a MAPE value of 2.32% for prediction, and outperforms others.
Hribar et al. (2019)	Forecasting the residential natural gas demand in the city of Ljubljana, Slovenia.	Machine-learning models are considered, such as linear regression, kernel machine and artificial neural network.	The recurrent neural network and linear regression are the most accurate ones.

In this formulation $\beta_0, \ldots, \beta_p$ are the regression coefficients to be estimated according to observations. To avoid multi-collinearity problems, correlations between the predictors should be controlled (the correlation coefficient of the explanatory variables should not exceed 0.7) (Anderson et al., 2014). The last term in the formulation, ε, denotes the random error and is referred as the residual for checking the overall significance of the model and each regression coefficient (Braun et al., 2014). Error term is independently and normally distributed, with a mean of zero and a constant variance of σ^2 (Montgomery et al., 2012).

3.1.2. Support vector machine

SVM emerged as a robust and accurate data mining technique in pattern recognition, classification and regression problems, and introduced by Vapnik (1998). In recent years, it has become a popular data mining method due to its success in solving problems. Its conceptual underpinning stems from statistical learning theory. SVM is a supervised learning technique that produces input-output mapping functions from a set of training data. SVM was developed to overcome some deficiencies of ANN, such as network construction issues, overfitting problems, and determining a number of data points for model training (Kavousi-Fard et al., 2014).

The SVM technique just described is valid for binary classification problems, however, the SVM model also includes general prediction problems, consisting of a SVM version for regression that is the SVR method. The purpose of the SVR is to find a function that deviates at most insensitive loss function (ε) from the actual outputs.

When reducing the error, the over-fitting risk decreases by simultaneously increasing the function's flatness. The regression model is created through a series of high-dimensional functions. The formulations between Eqs. (2) and (10) explain SVR methodology from a mathematical perspective:

$$f(x) = w * \varphi(x) + b \tag{2}$$

where $\varphi(x)$ denotes the kernel transformation function for the inputs, and w and b are parameters. The coefficients are calculated by minimizing the regularized risk function that is given below:

$$R(f) = C\,\frac{1}{n}\sum_{i=1}^{n} L_\varepsilon(y_i, f(x_i)) + \frac{1}{2}\,\|w\|^2 \tag{3}$$

where the parameter ε is the tolerance value.

$$L_\varepsilon(y, f(x)) = \begin{cases} 0, & |y-f(x)|<\varepsilon \\ |y-f(x)|-\varepsilon, & |y-f(x)|\geq\varepsilon \end{cases} \tag{4}$$

In the above equation, $L_\varepsilon(y,f(x))$ is labeled a ε insensitive loss function. The loss equals zero if the estimated value is within the ε tube. Analogous to the soft margin hyperplane, one needs to introduce two slack variables, ξ_i and ξ_i^* representing the positive and negative deviations, respectively, out of the ε zone. Eq. (2) is reformulated in the ensuing constraint form,

$$\min\ \frac{1}{2}\,\| w \|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{5}$$

subject to:

$$\begin{aligned} &[w * \varphi(x_i) + b] - y_i \leq \varepsilon + \xi_i^* \\ &y_i - [w * \varphi(x_i) + b] \leq \varepsilon + \xi_i \\ &\xi_i, \xi_i^* \geq 0 \end{aligned} \tag{6}$$

The following Lagrangian form solves this constraint optimization problem:

$$\begin{aligned} min\ \frac{1}{2}\,\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) - \sum_{i=1}^{n}\beta_i[\, w * \varphi(x_i) + b - y_i + \varepsilon + \xi_i] \\ - \sum_{i=1}^{n}\beta_i^*[y_i - w * \varphi(x_i) - b + \varepsilon + \xi_i^*] - \sum_{i=1}^{n}\alpha_i\xi_i + \alpha_i^*\,\xi_i^* \end{aligned} \tag{7}$$

The dual Lagrangian form is given below:

$$\begin{aligned} \max \sum_{i=1}^{n} y_i(\beta_i - \beta_i^*) - \varepsilon\sum_{i=1}^{n}(\beta_i - \beta_i^*) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_i(\beta_i - \beta_i^*) \\ \times \left(\beta_j - \beta_j^*\right) K(x_i, x_j) \end{aligned}$$

subject to

$$\begin{aligned} &\sum_{i=1}^{n}(\beta_i - \beta_i^*) = 0 \\ &0 \leq \beta_i \leq C \\ &0 \leq \beta_i^* \leq C \\ &i = 1, 2, 3, \ldots, n \end{aligned} \tag{8}$$

Consistent with the Karush-Kuhn-Tucker theorem, regression model is expressed by:

$$f(x) = (\beta_i - \beta_i^*)\,K(x_i, x_j) + b \tag{9}$$

$K(x_i,x_j)$ is a kernel function whose value equals the inner product of two vectors, x_i and x_j, in the feature space $\varphi(x_i)$ and $\varphi(x_j)$. The most popular kernel functions are radial basis kernel, polynomial kernel, sigmoidal kernel and linear kernel.

There is no standard to apply the proper kernel type for identified data patterns. In this study, polynomial kernels were used to examine natural gas consumption. Polynomial kernels demonstrate successful performance since they can determine the dimension of the feature space by their degree, and the computational simplicity provides them an advantage for algorithm speed. Moreover, polynomial kernels have been widely utilized in diverse applications (Ikeda, 2004). The polynomial kernel with an order of d and constant of a_1 and a_2 can be expressed in Eq. (10) as:

$$K(x_i, x_j) = (a_1 x_i x_j + a_2)^d \tag{10}$$

The performance of the SVR hinges on a good setting of the hyper-parameters (C, ε) and the kernel parameters (d). The identification of all three parameters significantly influences the estimation accuracy of SVR tools. More specifically, C provides the tradeoff between the training error and the model robustness. If C takes too large a value, then the empirical risk of the objective function will be minimized. The number of support vectors identifies with the ε parameter, and determine the width of the loss function in SVR. Larger values of ε are obtained if a smaller number of support vectors are conducted to function (Hong, 2011, 2009; Kavousi-Fard et al., 2014).

3.1.3. Artificial neural network models

ANNs are inspired by the neurological functions of the human brain and are formulated on the human cognitive system. Researchers in various fields showed a great interest in ANN because of its ability to find solutions under multi-dimensional, nonlinear and complex data sets. ANN can easily cope with nonlinear models and under incomplete data structure, and provide successful results. ANNs include a large number of computational elements (neurons) interacting across weighted connections. ANN shows some characteristics such as the working principle of the human brain; after learning complex knowledge patterns, it generalizes this knowledge to apply in various situations.

ANNs can be classified into various versions such as supervised, unsupervised learning techniques and feed-forward and feedback recall architectures. In classification and prediction problems, a back propagation ANN is one of the most preferred ANN tools (Pankratz, 2009).

The basic feature of ANN is that high flexibility capabilities in large diversity functional relationships from input to output. Due to the working principle, ANN shows good performance even under vague, missing and unclear data sets. Moreover, there is no need a priori hypothesis and a specific functional structure between input and output. For this reason, in case of lack of information or assumptions, ANN is widely used as a practical alternative (Haykin, 1999).

The performance of ANN depends on a well-designed network architecture based on different combinations of parameters such as number of neuron, layers, iteration and determining the connection weights, learning algorithm and transfer function. The illustrative structure of an ANN model comprises one input layer, one output layer and one hidden layer as given in Fig. 5.

The sample data within ANN methodology is split into two major subsets that are labeled as test and training sets. Whereas test set is applied to measure the model's performance in the testing process, during the training process the ANN learns the link between output and input layer (Hornik, 1993).

In order to get the optimal network architecture, various combinations are assessed. One of the factors among these combinations is identifying a proper transfer function. We have employed a hyperbolic tangent sigmoid transfer function as stated in Eq. (11):

$$f(s) = \frac{1-e^{-s}}{1+e^{-s}} \tag{11}$$

where s is the weighted input summation of the hidden layer, and $f(s)$ is the output of the hidden layer.

The most widely utilized learning algorithm is the backpropagation algorithm (Wang and Ramsay, 1998). The central idea of the backpropagation algorithm is to minimize the sum of square errors by backward propagating through the ANN. To enhance the learning process of the backpropagation algorithm, two parameters of the algorithm that include the learning rate and the momentum should be modified based on convergence rates.

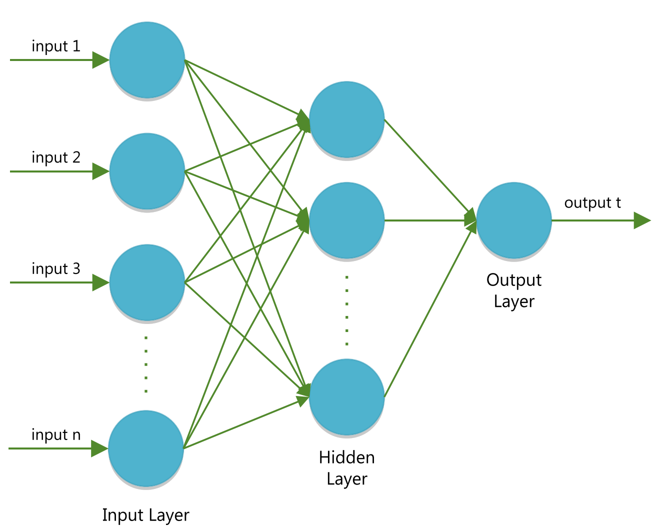

Fig. 5. ANN architecture.

The Levenberg-Marquardt algorithm, widely used to compute the weights of ANN in backpropagation algorithm, is derived from Newton's method (Hagan and Menhaj, 1994). The Newton's method for minimizing a function $V(x)$ with respect to the vector x is given in Eq. (12) as:

$$\Delta(x) = -\left[\nabla^2 V(x)\right]^{-1} \nabla V(x) \tag{12}$$

where $\nabla^2 V(x)$ denotes the Hessian matrix and $\nabla V(x)$is the gradient vector.

For more of a mathematical background of the method, readers can get more information from (Hagan and Menhaj, 1994). To summarize, the Levenberg-Marquardt algorithm presents a reconciliation between the speed of Gauss-Newton and the guaranteed

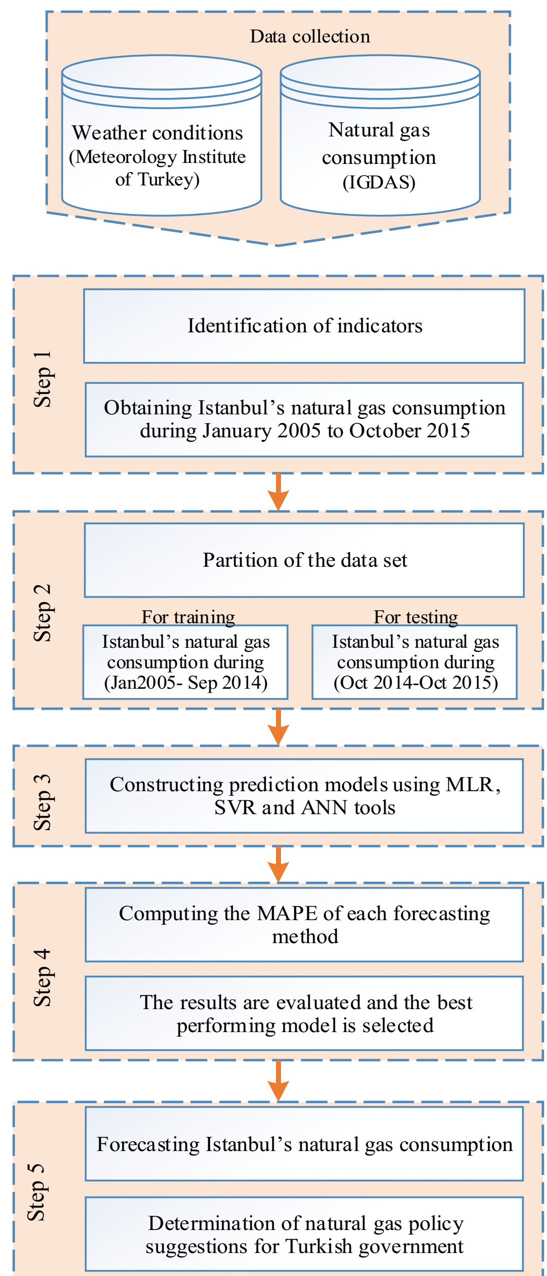

Fig. 6. Steps of the methodology.

convergence of steepest descent (Saini and Soni, 2002; Catalão et al., 2007).

3.2. The proposed procedure

The proposed forecasting procedure is presented in Fig. 6. There are three main phases in this method: (a) to determine the most suitable technique for estimating the consumption of natural gas in Istanbul; (b) to choose best features which will provide the most information about the expected forecasting risk; and (c) to offer recommendations for decision makers.

Step 1: The inputs of the proposed forecasting model are gas consumption of past months, temperature forecast, and various time features such as season and month of the year. This first step includes data collection, its analysis and extraction of its features.
Step 2: The second step involves application of a simple seasonal exponential smoothing technique to forecast future temperature data that will be used as input in the regression model. Data is split into two data sets, namely, training and testing data sets. Training data set is used to calculate the model parameters, while the latter is utilized to measure the model's performance.
Step 3: In the third step, various machine learning techniques including SVR, ANN and MLR are used to construct the forecasting model for natural gas demand in Istanbul.
Step 4: In the fourth step, performance of the forecasting models is compared through mean absolute percentage error (MAPE), and the parameters are tuned to achieve most accurate forecasting results.
Step 5: In the last step, natural gas consumption of Istanbul is predicted for the next year (for 12 months); then, some policy suggestions for natural gas are presented for energy policy makers/managers.

4. Application of the model

In our proposed methodology, firstly, we examine the historical natural gas consumption values of Istanbul, and then we utilize several machine learning techniques for forecasting purpose. After applying machine learning models, we measure their performance by comparing the estimated values with real consumption values.

4.1. Assessing the data

Monthly values of natural gas consumption for Istanbul over the period of 2004–2015 were obtained from Istanbul Gas Distribution Company (IGDAS), Turkey's largest natural gas distributor, and this was plotted in Fig. 7.

Based on a thorough review of pertinent literature (e.g. Hamzaçebi, 2007; Soldo, 2012; Melikoglu, 2013; Alcaraz and Villalvazo, 2017; Scarpa and Bianco, 2017; Liu et al., 2018) and also discussions with experts and academics with relevant knowledge and expertise, we initially identified a set of predictor variables which include seasonal index, temperature, price of natural gas, oil price, population of the province of Istanbul, exchange rate and GDP of the province of Istanbul. These variables are widely used in anticipating energy consumption behavior at the provincial level.

Seasonality of the natural gas consumption data can easily be detected by analyzing the time series plot. Natural gas demand decreases in hotter months while demand increases in colder months. Although there was no significant evidence of this trend, variance of the time series changes throughout time. Variability in winter months is greater than variability in summer months. This characteristic of the time series makes it challenging to model with traditional time series method.

Besides historical data, some other exogenous variable was needed. We checked the relationship between monthly average temperature and monthly consumption of natural gas data that were obtained from the Meteorology Institute of Turkey (national weather forecasting institute of Turkey). In Fig. 8, scatter plot of monthly average temperature and natural gas consumption was presented.

Fig. 8 shows that there exists a significant negative association between temperature and consumption values. The relationship appears linear up to 20 °C, and then temperature does not affect natural gas demand significantly. Pearson correlation coefficient calculated at 0.91 supports that natural gas consumption is affected by temperature and temperature should be used as exogenous variable in the forecasting model.

Seasonality effect was analyzed by calculating a variable called seasonality index. Seasonality index was computed by taking average of natural gas consumption values on a monthly basis over the period of 2004–2014. Fig. 9 presents the scatter plot between seasonal index and consumption, and supports a linear relationship. Pearson correlation coefficient between seasonality index and consumption values is 0.93.

As we focus on estimating energy consumption at the provincial level, price of natural gas (Turkish Lira/m^3), and population of the province of Istanbul are included in the model. We also incorporated two important macro-economic variables, namely GDP of the province of Istanbul (USD) and exchange rate (Turkish Lira/USD).

Lastly, we selected historical consumption data as an input variable. Since we intended to make 12 months forecast, we preferred to use lag 12 as input variable for our forecast model. Relationship between lag 12

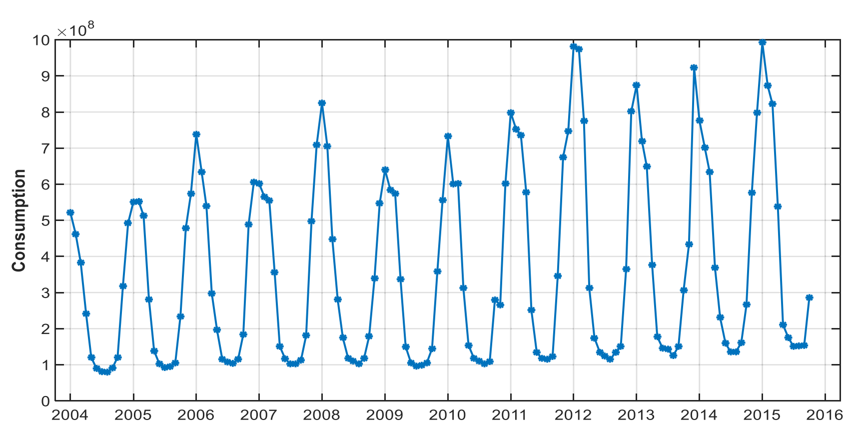

Fig. 7. Time plot of monthly natural gas consumption of Istanbul.

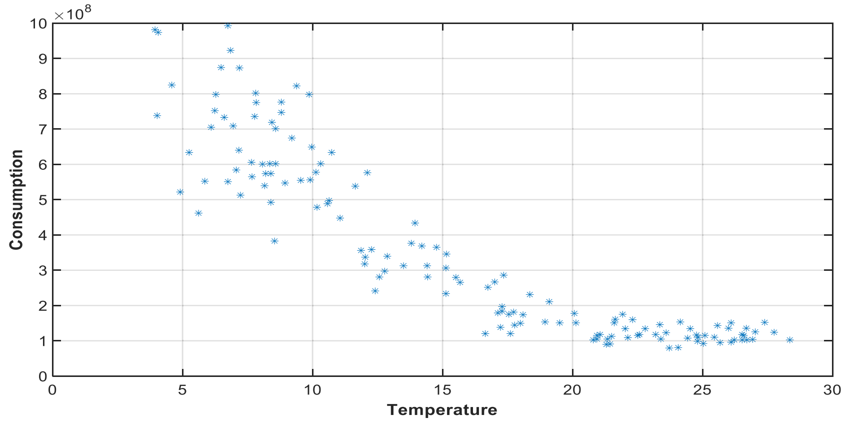

Fig. 8. Scatter plot of consumption values vs. temperature values.

and consumption data is presented in Fig. 10. Correlation coefficient value is 0.92, which indicates a strong positive linear relationship between gas consumption data and lag 12.

In addition, we checked the autocorrelation values from lag 1 to lag 15 (Table 2 and Fig. 11). Autocorrelation results indicated that lag 12 had the highest autocorrelation value (0.83). Analysis of time series of the natural gas consumption data suggests that temperature, seasonality index and lag 12 should be used as input variables in the forecasting model.

We applied stepwise regression model to identify which of the predictor variables affect highly the level of natural gas consumption. Stepwise regression results reveal that GDP of the province of Istanbul and exchange rate have no statistically significant effects on natural gas consumption and are excluded from further analyses. Thus, as

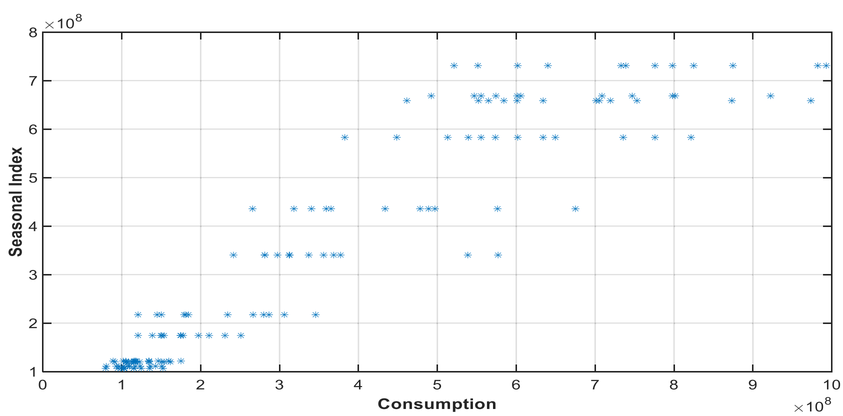

Fig. 9. The scatter plot of seasonal index vs. consumption.

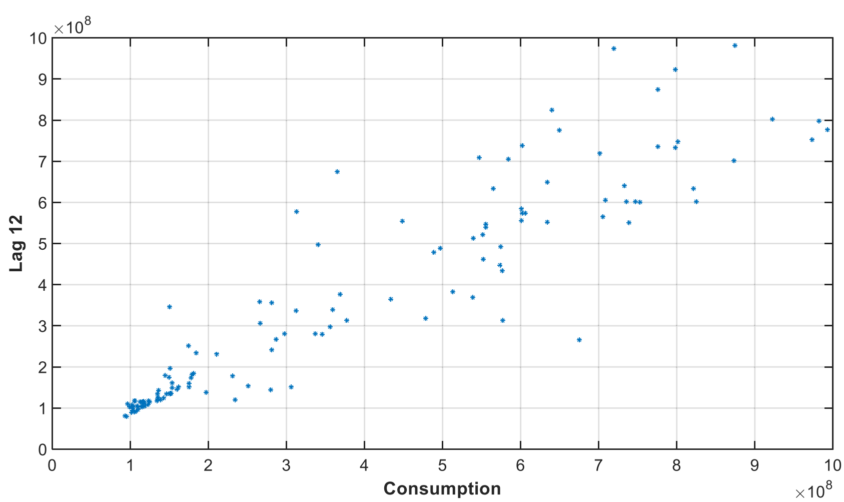

Fig. 10. The scatter plot of lag 12 vs. consumption.

Table 2
Autocorrelation values for consumption.

Lags	1	2	3	4	5	6	7	8	9	10	11	**12**	13	14	15
Correlation	0.82	0.46	0.02	−0.39	−0.66	−0.75	−0.66	−0.41	−0.03	0.39	0.71	0.83	0.71	0.39	−0.01

shown in Table 3, we focus on utilizing the following variables with high level of predictive power on natural gas consumption: seasonal index, temperature, price of natural gas, population of the province of Istanbul and natural gas consumption with lag 12.

4.2. Partitioning the data set

The proposed model is used for monthly natural gas consumption data from January 2005 to October 2015. As noted earlier, the model employs seasonal index, temperature, price of natural gas, population of the province of Istanbul and natural gas consumption with lag 12 as input variables while monthly consumption of natural gas is used as an output variable. Future temperature data used in the regression model as input needed to be forecasted from historical temperature data.

The data is split into training and test data. The data for the period 1/1/2005–30/10/2014 (129 months) were used in training the proposed models. The dataset between 1/11/2014 and 30/10/2015 was used to assess the performance of the proposed models. Future temperature data was predicted using a simple seasonal exponential smoothing method as shown in Fig. 12.

The *R* square value was found to be 0.98, and the MAPE was 4.5%, indicating that future temperature predictions were accurate enough to be used in the regression model. Price of natural gas on monthly basis is obtained from IGDAŞ (Istanbul Gas Distribution Industry and Trade Inc.) and population values of the province of Istanbul (monthly basis) are gathered from Turkish Statistical Institute (TUIK).

4.3. Analyzing the data set

Prior to implementation of a forecasting technique, the existence of stationarity should be checked first. When working with ordinary least square (OLS) or ARIMA estimation on variables that are non-stationary, statistically significant relationship between variables is observed erroneously. This fallacy is termed “spurious regression”. To avoid this fallacy, we test the natural gas consumption and temperature data for stationarity (Stock and Watson, 2015).

If a time series data is not stationary, usually, it is said to follow a “unit root” process. Unit root processes are a generalization of random walk processes with serially correlated errors. Both monthly natural gas consumption and temperature data are initially tested for stationarity using the Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests. The test results indicate that both series do not have unit root attesting the existence of stationarity.

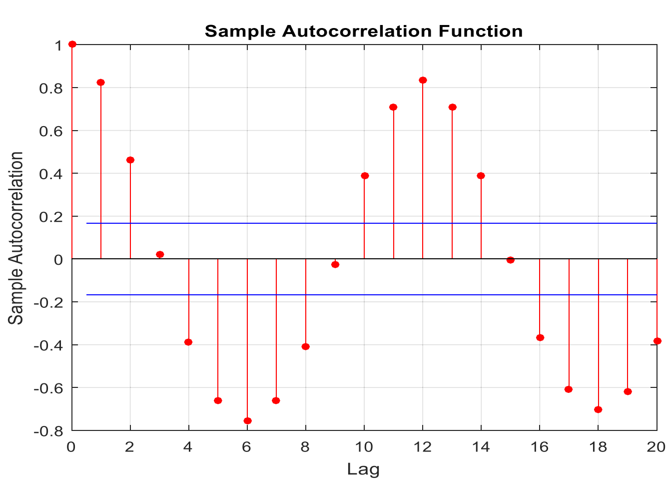

Fig. 11. Autocorrelation graph of consumption values.

In this study, we selected two powerful machine learning techniques which are ANN and SVR and one traditional technique as MLR to identify the best prediction method to project the monthly consumption of natural gas in Istanbul. R software package was used to assess each method.

The MLR equation model is shown in Eq. (13):

$$\begin{aligned}\text{Natural gas consumption} &\\ = &-0.294\ \text{Temperature} + 0.511\ \text{Seasonal index}\\ &+ 0.184\ \text{Natural gas consumption with lag 12}\\ &+ 0.328\ \text{Population of the province of Istanbul}\\ &-0.183\ \text{Price of natural gas consumption}\end{aligned} \tag{13}$$

The regression model, which is given above, has an excellent fit with the dataset ($F = 393.7$, $p < 0.01$) and explains 94% of the variation in the natural gas consumption. It should also be noted that our regression model does not violate the assumptions of homoscedasticity, independence, multi-collinearity and normality of errors. The regression coefficients of all variables are found significant ($p < 0.01$). As it is evident from Eq. (13), seasonal index is noted to be the leading criterion ($\beta = 0.511$) that affects monthly natural gas consumption in Istanbul followed by population of the province of Istanbul ($\beta = 0.328$), temperature ($\beta = -0.294$), natural gas consumption with lag 12 ($\beta = 0.184$) and price of natural gas consumption ($\beta = -0.183$), respectively.

When developing the SVM models, the main objective is to minimize the ε insensitive errors on training set. Therefore, selection of corresponding kernel functions and insensitive coefficients is crucially important. Linear, sigmoid, polynomial and radial basis functions are the most commonly used kernels. In this study, the SVR technique with polynomial cubic kernel function was used since SVR model with polynomial cubic kernel function outperformed the other SVR models. To further improve the prediction performance of SVR with polynomial cubic kernel function, we produced the best values of regularization parameter ($C = 1$) and insensitive coefficient ($\varepsilon = 0.001$) to minimize the error in the test dataset.

The architecture the ANN model comprises one input layer, one hidden layer and one output layer. In the hidden layer, sigmoid function is utilized as activation function while linear function is employed in the

Table 3
Stepwise regression results.

Variables	Standardized regression weights	Std. error	t value
Seasonal index	5.11E−01	1.03E−01	4.980***
Temperature	−2.94E−01	6.44E−02	−4.559***
Price of natural gas	−1.83E−01	5.63E−02	−3.251**
Population of the province of Istanbul	3.28E−01	5.33E−02	6.160***
Natural gas consumption with lag 12	1.85E−01	7.97E−02	2.319*
F-statistics	393.7***		
Adjusted R square	0.94		
N	142		

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

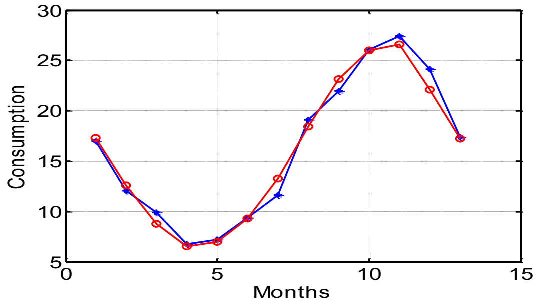

Fig. 12. The plot of real temperature values and predictions of testing data.

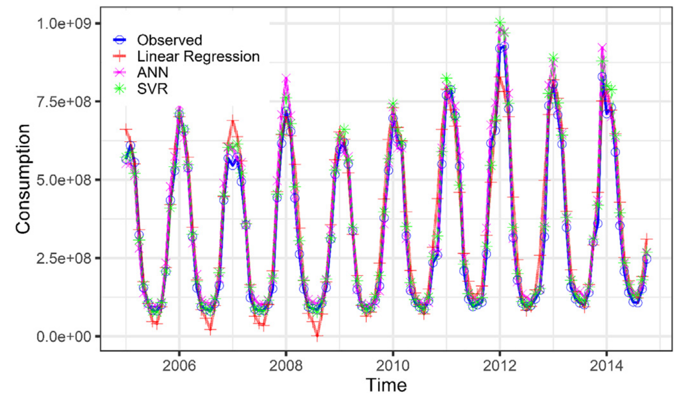

Fig. 13. The plot of real consumption values and predictions of training data.

output layer. Levenberg-Marquardt optimization algorithm is used to calculate the weights of neurons.

By changing the number of nodes (1 through 5) during the testing of the model, it was decided to use three nodes at the hidden layer where the minimum mean square error (MSE) value occurred at this point for the training data set. The MSE values were at minimum for both the training and testing data sets when three nodes were used in the hidden layer. If there were more than three nodes at the hidden layer, it would inevitably lead to an overtraining problem in the ANN model for small data sets. As our data set is relatively small, we limited the number of nodes to 3.

4.4. Comparing the results

The results of analysis obtained using MLR, ANN and SVR models were compared to evaluate the forecasting performance of the each model for the training and test data. Performances of all three models for training data set were measured using 7-fold cross validation technique. To assess the estimation performance of the model, different performance criteria such as MAPE and MSE were utilized. In this study, both MAPE and MSE were used as performance criteria and were calculated in Eqs. (14)–(15) as follows:

$$MAPE = \left[\frac{1}{n} \left(\sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right| \right) * 100 \right] \tag{14}$$

$$MSE = \left[\frac{1}{n} \left(\sum_{t=1}^{n} (A_t - F_t)^2 \right) \right] \tag{15}$$

where, n is the number of forecasted points in time, A_t is the actual value and F_t is the forecast value. The values of MAPE and MSE for both training and test data are provided in Table 4.

Comparison results of prediction values of each method and actual values on training data set with respect to time are presented in Fig. 13. As shown in Fig. 13, ANN, MLR, and SVR methods were able to track the actual trend in monthly natural gas consumption within the province of Istanbul. However, SVR model has the lowest MAPE result as compared with ANN and MLR. Fig. 14 demonstrates the percentage errors of each month in testing data sets for ANN, MLR, and SVR models.

Fig. 15 depicts the comparison of observed values and corresponding forecasts of each model in testing data set. Though three methods performed well and have <5% of error for estimating real consumption values, SVR did better than the ANN and MLR models. As shown in Fig. 15, between June and October, the SVR model predicted much better than the ANN and traditional MLR models.

Table 4
Forecasting performance based on MAPE and MSE results.

Forecasting method/data set	MLR		SVR		ANN	
	MAPE	MSE	MAPE	MSE	MAPE	MSE
Training data	18.70	61,806,049	8.14	30,728,408	9.89	35,903,335
Test data	14.24	90,897,525	5.53	36,130,270	13.10	68,242,576

Finally, in order to show our forecasts are not based on luck, we compared the MAPE results of all three models with that of a seasonal random walk model. The MAPE result of seasonal random walk model is 31.1%, which indicates a much higher level of forecasting error than all three prediction models.

4.5. Future consumption forecast

In this section, forecasted natural gas consumption for the next year (for 12 months) was calculated for energy policy using support vector regression. The results presented in Fig. 16 indicate that the demand for natural gas will increase in the future.

5. Conclusion and implications

Turkey has had the highest rate of energy and natural gas demand over the past two decades. Almost 98.5% of the natural gas used in Turkey is imported from foreign countries such as Russia, Iran, and Azerbaijan. Energy planning in Turkey and Istanbul cannot be carried out without knowing past, present and future consumption of natural gas. Overestimating the demand for natural gas is likely to increase energy costs for unused resources, while underestimating it may lead to severe crises in the country. Therefore, accurate estimation methods of natural gas demand become crucial to determine the energy policy for the province of Istanbul.

This study essentially focused on one very serious issue in the management of natural gas consumption: making sound decisions for natural gas predictions in the province of Istanbul utilizing efficient machine learning tools such as MLR, ANN, and SVR. Due to the successes in the dynamic data sets, ANN and SVR emerge as powerful artificial

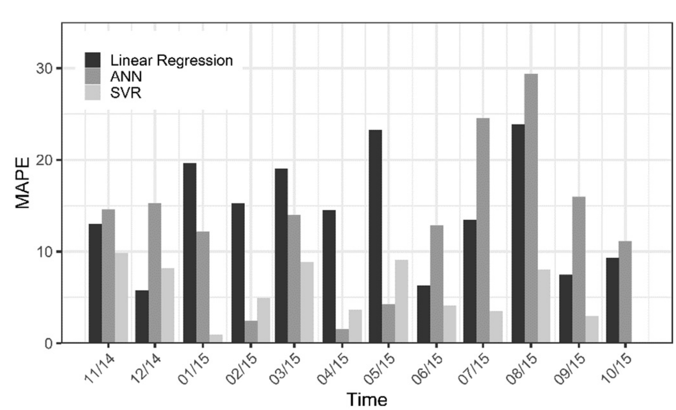

Fig. 14. The percentage errors for ANN, MLR and SVR models.

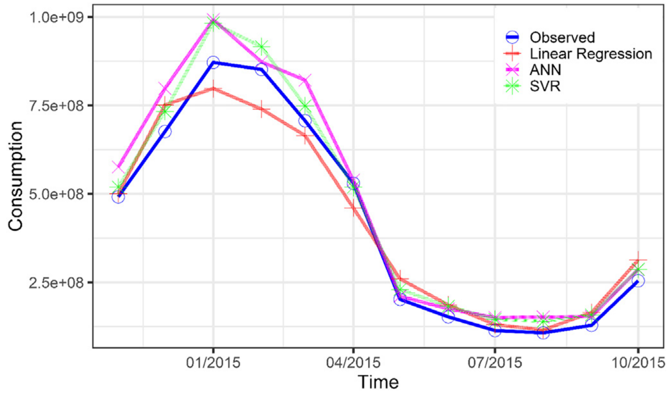

Fig. 15. The plot of August 2011–July 2012 real consumption values and forecasts from ANN, MLR and SVR.

intelligent tools in the forecasting studies while the MLR provides ease of calculation besides timely and accurate results even under big data. These powerful forecasting tools have been compared to each other to analyze their performance. This study revealed that the SVR model with polynomial cubic kernel function outperformed ANNs and the MLR models for monthly estimation of natural gas consumption using the following input variables: seasonal index, temperature, price of natural gas, population of the province of Istanbul and natural gas consumption with lag 12.

The proposed SVR model with polynomial cubic kernel function can be adopted as an important decision-making tool for both public policy makers and regulatory authorities to make rigorous forecasting of natural gas demand for mega-cities like Istanbul. As noted earlier, for the selected input variables that influence natural gas consumption in the province of Istanbul, seasonal index was noted to be the leading criterion followed by population of the province of Istanbul, temperature, natural gas consumption with lag 12 and price of natural gas consumption, respectively. Seasonal variation, population growth and temperature are critical for effective management of natural gas for a large emerging market like Turkey that is entirely reliant on foreign sources of supply where the demand for natural gas is still rising. As put forward by earlier studies on natural gas supply-demand balances, there is no serious problem about meeting the annual natural gas demand. However, during the winter months, high demand, or the fall of temperatures to levels below seasonal norms may lead to the rise of consumption to maximum levels. On the other hand, inaccurate predictions over the same period may cause imbalances between supply and demand. Our model also predicts that population growth leads to a tremendous increase in natural gas consumption. Istanbul is a mega-city attracting many immigrants from different regions of Turkey. According to a recently launched United Nations (UN) report (UN World Organization Prospects, 2018), the current population of the province of Istanbul is expected to reach 16.3 million by 2025 and 17.9 million by 2035. In order to meet increasing demand for natural gas consumption and also to effectively cope with seasonal variations, the Turkish Government should take some constructive measures.

Turkish Government has been suffering a lot from Take or Pay (ToP) penalties, which cost Turkey nearly $6 billion over one year, according to the Turkish Ministry of Energy and Natural Resources. Turkish Government has to pay its key suppliers – Russia, Iran and Azerbaijan – a fixed sum of money regardless of whether it actually imports all the natural gas it agrees to purchase. The requirement of the ToP policy is very significant to ensure reliability in all long-term natural gas contracts, which leads to some speculations for both sides (supplier and buyer). Under the ToP clauses, the customer either takes the product from the supplier or pays a penalty, regardless of whether it takes delivery or not. Uncertainties in demand during certain periods of the year and insufficient underground storage capacity are forcing countries like Turkey that are highly dependent on the imports of natural gas from other countries to cope with this huge financial burden. In addition, due to the volatility of local currency, there are also some serious currency risks, as currency is subject to depreciation over time because of the need for early payment.

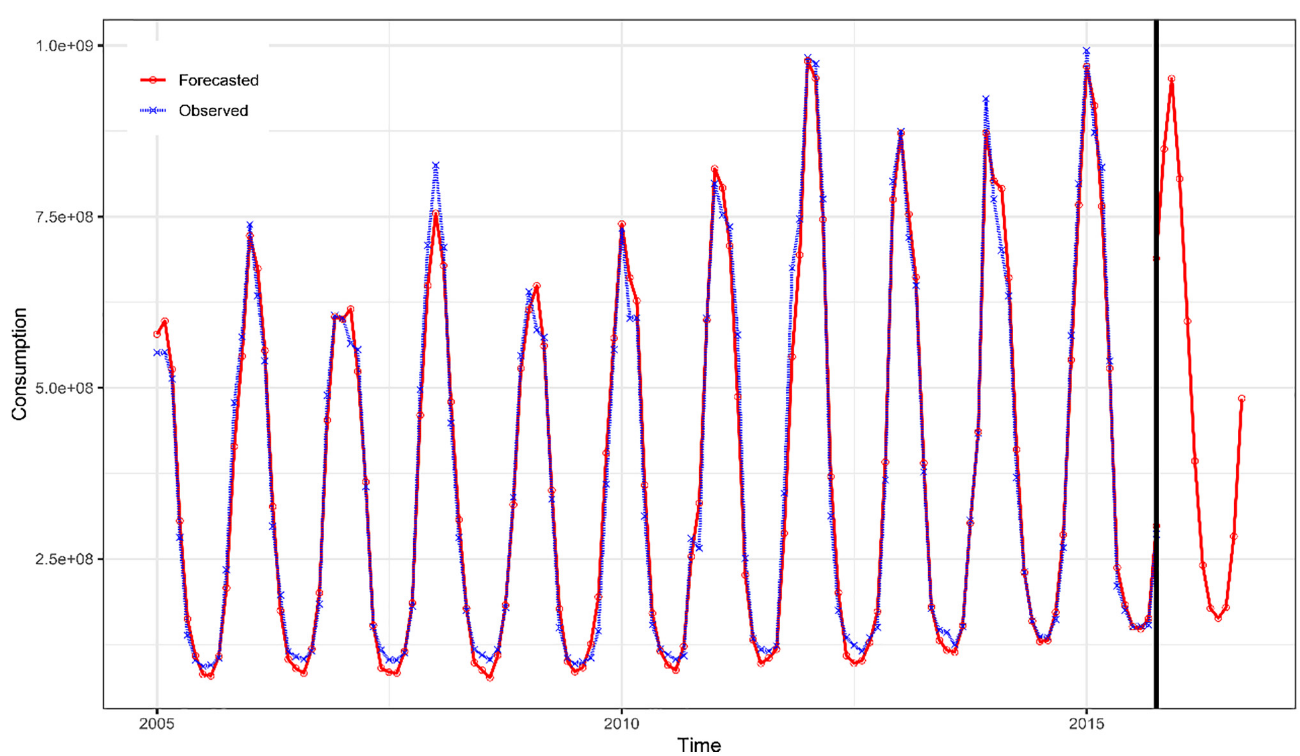

Fig. 16. Forecasted consumption values from 2015 October to 2016 October.

In order to cope with seasonal imbalances of demand and supply, it is imperative for the Turkish Government to increase both storage and withdrawal capacities. In the natural gas sector, gas storages are predominantly used for peak shavings. To avoid the negative effect of ToP, particularly the contracted volume which is not taken and not used because of storage capacity, the Turkish Government has been under great pressure to take some important measures and make some investment decisions. To this end, Turkey rapidly continues investments in natural gas development by closely monitoring novel technological developments and engaging in new natural gas pipeline projects. Within this framework, Silivri Natural Gas Storage Facility, which has a working capacity of 2.84 billion Sm^3, has been taken over by BOTAŞ in order to provide effective seasonal supply-demand balance and supply security. Recently, the Salt Lake Natural Gas Storage Project has been executed immediately in order to effectively eliminate supply interruptions. The working gas capacity was increased from 1 bcm to 5 bcm, and the withdrawal capacity was increased from 40 mcm/day to 80 mcm/day. The target of the project is to increase the capacity gradually by reaching 500 mcm of working gas capacity in 2018, 1.75 bcm in 2020, 3 bcm in 2021, and 5 bcm in 2023.

This study also emphasizes that while consumption of natural gas can be projected in the short run, for long-term planning and unexpected externalities, Turkish Government is in serious need of urgently establishing gas storage facilities to respond to supply or demand elasticity when actual level of demand for natural gas in Istanbul is less than the projected level. It helps to alleviate the adverse impacts of fluctuations in demand and supply of natural gas. From a broader view, Turkey's natural gas expenditure may seriously fall against the negative consequences of contracts that are based on a projected total volume of gas consumption.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eneco.2019.03.006.

References

Akpinar, M., Yumusak, N., 2016. Year ahead demand forecast of city natural gas using seasonal time series methods. Energies 9, 1–17.
Alcaraz, C., Villalvazo, S., 2017. The effect of natural gas shortages on the Mexican economy. Energy Econ. 66, 147–153.
de Almeida, A.T., Lopes, A.C., Carvalho, A., Mariano, J., Jahn, A., Broege, M., 2004. Examining the potential of natural gas demand-side measures to benefit customers, the distribution utility, and the environment: two case studies from Europe. Energy 29, 979–1000.
Anderson, D.R., Sweeney, D.J., Williams, T.A., 2014. Modern Business Statistics with Microsoft Excel. 5 edition. (South-Western College Pub).
Askari, S., Montazerin, N., Zarandi, M.H.F., 2015. A clustering based forecasting algorithm for multivariable fuzzy time series using linear combinations of independent variables. Appl. Soft Comput. 35, 151–160.
Azadeh, A., Zarrin, M., Rahdar Beik, H., Aliheidari Bioki, T., 2015. A neuro-fuzzy algorithm for improved gas consumption forecasting with economic, environmental and IT/IS indicators. J. Pet. Sci. Eng. 133, 716–739.
Baldacci, L., Golfarelli, M., Lombardi, D., Sami, F., 2016. Natural gas consumption forecasting for anomaly detection. Expert Syst. Appl. 62, 190–201.
Boran, F.E., 2015. Forecasting natural gas consumption in Turkey using grey prediction. Energy Sources Part B 10, 208–213.
BP Report, 2017. Statistical Review of World Energy | Energy economics | BP Global [WWW Document]. http://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html, Accessed date: 1 March 2018.
Braun, M.R., Altan, H., Beck, S.B.M., 2014. Using regression analysis to predict the future energy consumption of a supermarket in the UK. Appl. Energy 130, 305–313.
Cardoso, C.V., Cruz, G.L., 2016. Forecasting natural gas consumption using ARIMA models and artificial neural networks. IEEE Lat. Am. Trans. 14, 2233–2238.
Catalão, J.P.S., Mariano, S.J.P.S., Mendes, V.M.F., Ferreira, L.A.F.M., 2007. Short-term electricity prices forecasting in a competitive market: a neural network approach. Electr. Power Syst. Res. 77, 1297–1304.
Chatterjee, S., Hadi, A.S., 2000. Regression Analysis by Example. 5th edition. Wiley.
Chen, Y., Chua, W.S., Koch, T., 2018. Forecasting day-ahead high-resolution natural-gas demand and supply in Germany. Appl. Energy 228, 1091–1110.
Demirel, Ö.F., Zaim, S., Çalişkan, A., Özuyar, P., 2012. Forecasting natural gas consumption in Istanbul using neural networks and multivariate time series methods. Turk. J. Electr. Eng. Comput. Sci. 20, 695–711.
Dilaver, Z., Hunt, L.C., 2011. Industrial electricity demand for Turkey: a structural time series analysis. Energy Econ. 33, 426–436.
Ding, S., 2018. A novel self-adapting intelligent grey model for forecasting China's natural-gas demand. Energy 162, 393–407.
Dong, C., Huang, G.H., Cai, Y.P., Liu, Y., 2013. Robust planning of energy management systems with environmental and constraint-conservative considerations under multiple uncertainties. Energy Conversion and Management, Global Conference on Renewable Energy and Energy Efficiency for Desert Regions 2011 "GCREEDER 2011" 65, pp. 471–486.
EIA, 2015. U.S. natural gas production reaches record high in 2015 - today in energy - U.S. Energy Information Administration (EIA) [WWW Document]. http://www.eia.gov/todayinenergy/detail.php?id=25832, Accessed date: 2 May 2018.
Elragal, H.M., 2009. Improving neural networks prediction accuracy using particle swarm optimization combiner. Int. J. Neural Syst. 19, 387–393.
Energy and Renewables - Invest in Turkey, 2018. Energy and Renewables - Invest in Turkey. URL http://www.invest.gov.tr/en-US/sectors/Pages/Energy.aspx (accessed 02.08.2018).
Energy Governance in Turkey, 2015. https://www.energy-community.org/portal/page/portal/ENC_HOME/DOCS/3894261/25824B882CF017E0E053C92FA8C0EE59.PDF (accessed 03.01.2018).
Erdogdu, E., 2010. Natural gas demand in Turkey. Appl. Energy 87, 211–219.
Ervural, B.C., Beyca, O.F., Zaim, S., 2016. Model estimation of ARMA using genetic algorithms: a case study of forecasting natural gas consumption. Procedia - Social and Behavioral Sciences, 12th International Strategic Management Conference, ISMC 2016, 28-30 October 2016. Antalya, Turkey 235, 537–545.
Fagiani, M., Squartini, S., Gabrielli, L., Spinsante, S., Piazza, F., 2015. Domestic water and natural gas demand forecasting by using heterogeneous data: a preliminary study. In: Bassis, S., Esposito, A., Morabito, F.C. (Eds.), Advances in Neural Networks: Computational and Theoretical Issues. Springer International Publishing, Cham, pp. 185–194.
Fan, G.-F., Wang, A., Hong, W.-C., 2018. Combining grey model and self-adapting intelligent grey model with genetic algorithm and annual share changes in natural gas demand forecasting. Energies 11, 1625.
Hacisalihoglu, B., 2008. Turkey's natural gas policy. Energy Policy 36, 1867–1872.
Hagan, M.T., Menhaj, M.B., 1994. Training feedforward networks with the Marquardt algorithm. IEEE Trans. Neural Netw. 5, 989–993.
Hamzaçebi, C., 2007. Forecasting of Turkey's net electricity energy consumption on sectoral bases. Energy Policy 35 (3), 2009–2016.
Haykin, S.S., 1999. Neural Networks: A Comprehensive Foundation. 2nd ed. Prentice Hall, Upper Saddle River, N.J.
Hong, W.-C., 2009. Hybrid evolutionary algorithms in a SVR-based electric load forecasting model. Int. J. Electr. Power Energy Syst. 31, 409–417.
Hong, W.-C., 2011. Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm. Energy 36, 5568–5578.
Hornik, K., 1993. Some new results on neural network approximation. Neural Netw. 6, 1069–1072.
Hribar, R., Potočnik, P., Šilc, J., Papa, G., 2019. A comparison of models for forecasting the residential natural gas demand of an urban area. Energy 167, 511–522.
IEA, 2017. Turkey - International - Analysis - U.S. Energy Information Administration (EIA) [WWW Document]. https://www.eia.gov/beta/international/analysis.cfm?iso=TUR (accessed 03.01.2018).
Ikeda, K., 2004. Geometry and learning curves of kernel methods with polynomial kernels. Syst. Comp. Jpn. 35, 41–48.
International Energy Agency, 2017. URL http://www.iea.org/ (accessed 23.02.2018).
Iskin, I., Daim, T., Kayakutlu, G., Altuntas, M., 2012. Exploring renewable energy pricing with analytic network process — comparing a developed and a developing economy. Energy Econ. 34, 882–891.
Izadyar, N., Ong, H.C., Shamshirband, S., Ghadamian, H., Tong, C.W., 2015. Intelligent forecasting of residential heating demand for the district heating system based on the monthly overall natural gas consumption. Energ. Buildings 104, 208–214.
Karadede, Y., Ozdemir, G., Aydemir, E., 2017. Breeder hybrid algorithm approach for natural gas demand forecasting model. Energy 141, 1269–1284.
Karimi, H., Dastranj, J., 2014. Artificial neural network-based genetic algorithm to predict natural gas consumption. Energy Syst. 5, 571–581.
Kavaklioglu, K., 2011. Modeling and prediction of Turkey's electricity consumption using support vector regression. Appl. Energy 88, 368–375.
Kavousi-Fard, A., Samet, H., Marzbani, F., 2014. A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting. Expert Syst. Appl. 41, 6047–6056.
Kaynar, O., Yilmaz, I., Demirkoparan, F., 2011. Forecasting of natural gas consumption with neural network and neuro fuzzy system. Energy Education Science and Technology Part A: Energy Science and Research 26, 221–238.
Khan, M.A., 2015. Modelling and forecasting the demand for natural gas in Pakistan. Renew. Sust. Energ. Rev. 49, 1145–1159.
Liang, F.-Y., Ryvak, M., Sayeed, S., Zhao, N., 2012. The role of natural gas as a primary fuel in the near future, including comparisons of acquisition, transmission and waste handling costs of as with competitive alternatives. Chem. Cent. J. 6, S4.
Liu, X., Moreno, B., García, A.S., 2016. A grey neural network and input-output combined forecasting model. Primary energy consumption forecasts in Spanish economic sectors. Energy 115, 1042–1054.
Liu, G., Dong, X., Jiang, Q., Dong, C., Li, J., 2018. Natural gas consumption of urban households in China and corresponding influencing factors. Energy Policy 122, 17–26.
Melikoglu, M., 2013. Vision 2023: forecasting Turkey's natural gas demand between 2013 and 2030. Renew. Sust. Energ. Rev. 22, 393–400.

Montgomery, D.C., Peck, E.A., Vining, G.G., 2012. Introduction to Linear Regression Analysis. 5th ed. Wiley series in probability and statisticsWiley, Hoboken, NJ.
Naji, S., Shamshirband, S., Basser, H., Alengaram, U.J., Jumaat, M.Z., Amirmojahedi, M., 2016. Soft computing methodologies for estimation of energy consumption in buildings with different envelope parameters. Energ. Effic. 9, 435–453.
Olgun, M.O., Ozdemir, G., Aydemir, E., 2012. Forecasting of Turkey's natural gas demand using artifical neural networks and support vector machines. Energy Education Science and Technology Part A: Energy Science and Research 30, 15–20.
Ozdemir, G., Aydemir, E., Olgun, M.O., Mulbay, Z., 2016. Forecasting of Turkey natural gas demand using a hybrid algorithm. Energy Sources Part B 11, 295–302.
Özmen, A., Yılmaz, Y., Weber, G.-W., 2018. Natural gas consumption forecast with MARS and CMARS models for residential users. Energy Econ. 70, 357–381.
Panapakidis, I.P., Dagoumas, A.S., 2017. Day-ahead natural gas demand forecasting based on the combination of wavelet transform and ANFIS/genetic algorithm/neural network model. Energy 118, 231–245.
Pankratz, A., 2009. Forecasting with Univariate Box-Jenkins Models: Concepts and Cases. John Wiley & Sons.
Popoola, O.M., Munda, J., Mpanda, A., 2018. Residential lighting load profile modelling: ANFIS approach using weighted and non-weighted data. Energ. Effic. 11, 169–188.
Potočnik, P., Soldo, B., Šimunović, G., Šarić, T., Jeromen, A., Govekar, E., 2014. Comparison of static and adaptive models for short-term residential natural gas forecasting in Croatia. Appl. Energy 129, 94–103.
Republic of Turkey Ministry of Foreign Affairs, 2017. Republic of Turkey Ministry of Foreign Affairs [WWW Document]. http://www.mfa.gov.tr/turkeys-energy-strategy.en.mfa (accessed 23.02.2018).
Rodger, J.A., 2014. A fuzzy nearest neighbor neural network statistical model for predicting demand for natural gas and energy cost savings in public buildings. Expert Syst. Appl. 41, 1813–1829.
Saini, L.M., Soni, M.K., 2002. Artificial neural network based peak load forecasting using Levenberg–Marquardt and quasi-Newton methods. IEE Proc. Gener. Transm. Distrib. 149, 578.
Sánchez-Úbeda, E.F., Berzosa, A., 2007. Modeling and forecasting industrial end-use natural gas consumption. Energy Econ. 29, 710–742.
Scarpa, F., Bianco, V., 2017. Assessing the quality of natural gas consumption forecasting: an application to the Italian residential sector. Energies 10, 1879.
Shaikh, F., Ji, Q., 2016. Forecasting natural gas demand in China: logistic modelling analysis. Int. J. Electr. Power Energy Syst. 77, 25–32.
Shakouri, G.H., Kazemi, A., 2016. Selection of the best ARMAX model for forecasting energy demand: case study of the residential and commercial sectors in Iran. Energ. Effic. 9, 339–352.
Soldo, B., 2012. Forecasting natural gas consumption. Appl. Energy 92, 26–37.
Stock, J.H., Watson, M.M., 2015. Introduction to Econometrics. 3rd. edition. Pearson Education.
Sun, W., Wang, C., Zhang, C., 2017. Factor analysis and forecasting of CO2 emissions in Hebei, using extreme learning machine based on particle swarm optimization. J. Clean. Prod. 162, 1095–1101.
Szoplik, J., 2015. Forecasting of natural gas consumption with artificial neural networks. Energy 85, 208–220.
Taşpinar, F., 2015. Time series models for air pollution modelling considering the shift to natural gas in a Turkish city. Clean: Soil, Air, Water 43, 980–988.
Taşpinar, F., Çelebi, N., Tutkun, N., 2013. Forecasting of daily natural gas consumption on regional basis in Turkey using various computational methods. Energ. Buildings 56, 23–31.
Tatoglu, E., Bayraktar, E., Arda, O.A., 2015. Adoption of corporate environmental policies in Turkey. J. Clean. Prod. 91, 313–326.
Tolba, M.K., Biswas, A.K., 2013. Earth and Us: Population – Resources – Environment – Development. Elsevier.
Vapnik, V.N., 1998. Statistical Learning Theory (Adaptive and Learning Systems for Signal Processing, Communications and Control). Wiley, New York.
Wadud, Z., Dey, H.S., Kabir, M.A., Khan, S.I., 2011. Modeling and forecasting natural gas demand in Bangladesh. Energy Policy 39, 7372–7380.
Wang, A.J., Ramsay, B., 1998. A neural network based estimator for electricity spot-pricing with particular reference to weekend and public holidays. Neurocomputing 23, 47–57.
Wang, J., Mohr, S., Feng, L., Liu, H., Tverberg, G.E., 2016. Analysis of resource potential for China's unconventional gas and forecast for its long-term production growth. Energy Policy 88, 389–401.
Wang, D., Liu, Y., Wu, Z., Fu, H., Shi, Y., Guo, H., 2018. Scenario analysis of natural gas consumption in China based on wavelet neural network optimized by particle swarm optimization algorithm. Energies 11, 825.
Wiser, R., Bachrach, D., Bolinger, M., Golove, W., 2004. Comparing the risk profiles of renewable and natural gas-fired electricity contracts. Renew. Sust. Energ. Rev. 8, 335–363.
World Bank, 2016. Data The World Bank [WWW Document]. http://data.worldbank.org/ (accessed 24.02.2018).
Yu, F., Xu, X., 2014. A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network. Appl. Energy 134, 102–113.
Zeng, B., Li, C., 2016. Forecasting the natural gas demand in China using a self-adapting intelligent grey model. Energy 112, 810–825.
Zhang, X.-H., Zhang, R., Wu, L.-Q., Deng, S.-H., Lin, L.-L., Yu, X.-Y., 2013. The interactions among China's economic growth and its energy consumption and emissions during 1978–2007. Ecol. Indic. 24, 83–95.