Estimation of Turkey's Natural Gas Consumption by Machine Learning Techniques

Osman Emin ERDEM* (iD) , Saadettin Erhan KESEN (iD)

Konya Technical University, Faculty of Engineering and Natural Sciences, Industrial Engineering Department Selçuklu, Konya, Turkey

Highlights
• Real data for energy demand of Turkey are obtained.
• Five machine learning techniques are utilized to predict the future energy demand of Turkey.
• Relative performances of the techniques are compared.
• R statistical programming language is used to code algorithms.

Abstract

Technological advancements coupled with growing world population require the increasing need of energy. Natural gas is one of the most important usable energy resources. Turkey is with high external dependency on energy as it has its own limited natural and underground energy resources. Thus, in order to effectively and productively use of natural gas purchased from foreign countries and to make reliable and robust energy policies for the years ahead, it is crucial to make a reasonable and plausible prediction for natural gas consumption of Turkey. In this paper, we estimate the natural gas consumption using machine learning techniques on the basis of real monthly data representing natural gas consumption of Turkey between the years 2010 and 2018. The performances of machine learning techniques involving Artificial Neural Networks, Random Forest Tree, Regression, Time Series and Multiple Seasonality Time Series are compared in predicting the natural gas consumption of Turkey. Experimental results show that among the five techniques, artificial neural networks produce the best estimation, having the lowest mean square errors, followed by regression method. Time series shows the worst performance among all the techniques.

1. INTRODUCTION

Recent developments in industry and increasing world population have resulted in more energy need than ever before. In this context, Turkey has to make sustainable energy policies to effectively manage and meet the energy demand in the long run since it has scarce natural underground resources and growing population and economies.

Energy resources are mainly divided into two groups as primary and secondary energy resources. Primary energy resources are also classified into two categories such as renewable and non-renewable energy resources. While oil (petroleum), natural gas, coal are considered as non-renewable energy resources, hydraulic energy, wind power, solar energy are described as renewable energy resources. Energy resources generated utilizing primary energy resources are known as secondary energy resources.

As a result of development plan currently implemented in Turkey, becoming an important actor in global economy is important. In line with this target, it is significant for Turkey to increase its production capacity in the industry branches where it dominates in the global market, to find new industry branches and in particular to meet its increasing energy demand. With respect to this aim, effective management of limited energy resources is becoming a critical issue.
While non-renewable energy resources used in Turkey can be classified as coal, oil, natural gas renewable energy resources include hydraulic, wind, solar and bio-mass energy.

---

*Corresponding author, e-mail: sekesen@ktun.edu.tr

Primary energy demand in the world is increasing day by day and it is expected that this demand will increase by 40% until 2030. The region where the world's primary energy demand is expected to meet and which also covers Turkey approximately constitutes 70% of total oil and gas reserves of the world. Turkey's geopolitical importance is to increase with its presence in the region because neighboring countries retain three fourth of the total oil and gas reserves. It forms a direct corridor with the market in Caspian, Central Asian, Middle Eastern countries and Europe. This is because Turkey continues to collaborate on many projects.

Turkey has very limited natural gas deposits. Its natural gas deposits are located in Hamitabat (Kırklareli) and Mardin. When considering natural gas supply and demand balance in Turkey, no disruptions in meeting annual gas demand do not seem. But, due to the high demand in the winter months, the disruptions experienced in the natural gas countries that are imported can cause supply-demand imbalance. In 2017 natural gas consumption in Turkey was realized as 53.5 billion cubic meters [1].

Energy plays a critical role in the development of countries around the world. Energy is the main input of development for the countries. After the oil crisis that began in 1970s, international energy companies started to search for new energy sources. With this search, the interest in natural gas has increased and the importance of natural gas in international markets has increased gradually [2].

Natural gas demand in international markets is growing up due to its low-price as compared to its alternatives, low operating and maintenance costs, no storage costs, no environmental pollution, no transportation problems. Natural gas is preferred for industrial, residential and electricity production purposes. As a result of these advantages, the share of natural gas in the world energy resources is increasing and is subject to extremely critical trade flows in the global economy [3].

In order to manage energy policies well, developing countries need to determine what their future energy consumption will be. Accordingly, prospective consumption can be determined using forecasting models. Machine learning techniques can produce more successful results as the number of observations increases. Although machine learning techniques have extensive use and enormous potential, it is important to understand the limit of the technique. Apart from solid parameters, it has very little flexible estimation. Care must be taken as to which algorithm can fully define the actual data before implementation. Therefore, algorithm selection is very important [4].

The prediction model aims at educating a prediction model known as supervised learning about what it wants to learn and how it aims to learn. The educational process of the supervised learning model clearly describes what it needs to learn in the educational process and how it aims to learn.

Machine learning is a sub-branch of artificial intelligence which is used effectively in many fields today. Machine learning is based on the principle of automatic learning and development. Machine learning searches for some patterns in the data with various algorithms and methods, and then learns by looking at the labels corresponding to these patterns, and then allows the development of systems that can make inferences by using their experience (when faced with a situation similar to what they have learned).

Machine learning algorithms are generally classified in three ways: (i) supervised learning, (ii) unsupervised learning, and (iii) reinforcement learning. Algorithms in the supervised learning class use labeled data to make predictions based on what they have learnt. That is, the data to be used in education and the classes (categories/labels) of the data are known in advance. With this information, the system learns and interprets the new incoming data with this information. In this study, in order to estimate Turkey's natural gas energy consumption for upcoming years, the performances of five machine learning techniques including Artificial Neural Networks, Random Forest Tree, Regression, Auto Regressive and Moving Average Time Series, and Multiple Seasonality Time Series are evaluated under the criterion of mean square errors using R statistical software.

The remaining of the paper can be summarized as follows. İn Section 2, we review the relevant studies regarding machine learning techniques for estimation. Section 3 is devoted to explaining the machine learning techniques that we use to estimate Turkey's natural gas consumption. Section 4 demonstrates experimental results. Concluding remarks are stated in Section 5.

2. LITERATURE SURVEY

Since there are many studies on the use of forecasting techniques in the literature, the scope of the literature research has been narrowed as the use of machine learning techniques in energy estimation.

[5] analyzed data involving Gross Domestic Product, building surface area, the number of vehicles, population, import and export values between years 1970 and 2010 and estimated Turkey's net energy demand for years between 2011-2015. In the study, multiple regression technique and artificial neural networks were used for estimation and it was found that the estimation performance of artificial neural networks was better than multiple regression.

[6] presented the results of the predictions made by simple regression, which is one of the machine learning techniques, and multiple regression are presented. The effects of the factors that are thought to have an effect on electricity consumption were also examined.

[7] investigated long term electrical energy consumption by using artificial neural networks. They compared the results with regression techniques and box-jenkins models. As a result of the study, it has been seen that artificial neural networks give more successful performance in electrical energy consumption.

[8] estimated the generation of electricity from wind energy using artificial neural network method. In the wind speed data test stage, it was shown that the estimations made by the model formed from the regression curves using the output power of different types of wind turbines during the training phase were consistent.

[9] estimated electricity production of a hydroelectric plant for 12 months fed from Turkey's Isparta Aksu district and Yenice stream. Monthly average relative humidity and monthly average flow rate, monthly average rainfall data as input using feed forward neural network model 12-month electricity production forecast study was performed. It is shown that the model successfully represents monthly electricity consumption.

[10] clustered the data with the approach developed in order to estimate the amount of energy production for the wind power plants and the electricity energy prices to be given to the market and estimated the interval by applying local percentage slice regression. The model was tested in wind power monitoring and forecasting center (RITM) and it was found that successful results were obtained. In the next step, a method for estimating electricity prices has been developed.

[11] tried to make predictions about natural gas demand in terms of supply of energy in Konya for future periods. For this purpose, as a result of time series analysis, it was concluded that the natural gas demand of the households will fluctuate and have seasonality in months. In the analysis conducted for the industrial sector, it is shown that consumption is expected to enter an upward trend over the years.

[12] examined 48-hour wind forecast for Turkey and studied the consistency of these estimates. Hourly data obtained from automatic observation stations in Balıkesir-Balya and İstanbul-Çatalca meteorology radar fields were used as observation data. In the study, MM5 medium scale estimation model developed by NCAR (National Center for Atmospheric Research) is used. It is stated that the correlation of the predictions and observations is high.

[13] created a fuzzy logic model for Turkey's short-term electricity demand forecasts and compared with the models obtained by regression analysis and time series. They performed the performance evaluation of the models by comparing the mean absolute relative error rates of these methods. As a result of the study, the estimation made by time series analysis was found to be superior to the other two models.

Eren [14], using Turkey's annual natural gas consumption and population data, gray model (1,1) acquired natural gas consumption and population estimates for the 2017-2030 period with the model.

[15] examined the effect of economic and population factors on annual electricity consumption in Italy in a long-term prediction model. Using the data from 1970 to 2007, they made a prediction study using a regression model.

[16] analyzed the results based on historical data using the time series model for the US natural gas production forecast.

[17] used multiple regression, ARMAX and artificial neural network models to estimate Istanbul's natural gas consumption. As a result of this study, back propagation artificial neural network model was found to be superior to other models.

[18] based on NARX artificial neural networks, tried to dynamically predict the daily movements of natural gas prices in the German market. For this purpose, some of the traders interviewed with the characteristics of the impact of the features determined and investigated which features can be effective for a sensitive analysis.

[19] calculated the multiple regression method and the artificial neural network method using a computer program called GasDay to estimate the daily demand of natural gas consumption. They concluded that the computer program used was good in estimating results for the relevant data set.

[20] presented a breeder hybrid algorithm relied upon nonlinear regression for forecasting of natural gas demand of Turkey between years 2001 and 2014. They stated that their study distinguish from the literature as their algorithm continuously evolves to the best solution.

[21] developed a heuristic algorithm composing of genetic algorithm and simulated annealing algorithm to forecast the future energy demand of Turkey. Under two different scenarios, they evaluated effectiveness of the developed algorithm against linear regression. Results show the superiority of the algorithm they developed.

[22] used support vector machines and artificial neural networks in order to estimate the natural gas demand of Turkey. Results show that support vector machines have less error rate as compared to artificial neural network.

This study differs from the literature in some aspects. Five different machine learning models were used to estimate Turkey's monthly consumption of natural gas by using R statistical package and outputs are measured by mean square error model.

3. MACHINE LEARNING TECHNIQUES

In this section, the general working principles of the five different machine learning techniques used in the study will be discussed.

3.1. Artificial Neural Networks

Artificial neural networks provide an understanding of the relationship between input signals and output signals using sensory inputs using a model of how we understand the response from the brain. It uses artificial neural neurons or knuckles to solve problems that the brain has learned in an artificial neural network, such as parallel structures that the brain uses connecting cells called neurons.

A single neural cell model can explain the artificial neural network for a fairly simple biological model. In Figure 1, the input signals describe the relationship between the dendrites (numeric variable x) and the output signal (numeric variable y) by the functional work network. The biological neuron assigns a weight value to each dendrite signal according to its importance. The input signals are collected by the body cells and the signals go through the activation function (the activation function is shown as f).
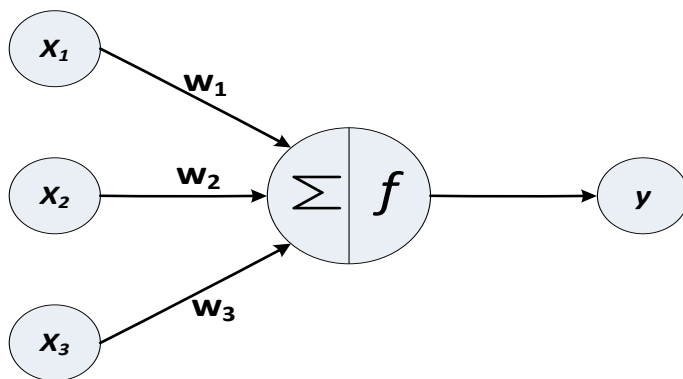

Figure 1. *A simple neuron*

In general in the artificial neural network n input dendrite returns the answer with the following formula. These w weights result from each n input value (the value shown as x_i) as the sum of the large or small contributing values of the input signals. It is used with the total activation function $f(x)$ and the result function $y(x)$ [23], as is shown in Equation (1)

$$y(x) = f(\sum_{i=1}^{n} w_i \cdot x_i). \tag{1}$$

Activation function is the mechanism that transmits the information received from the artificial neural networks to the network. Just as the artificial neural network is modeled after the biological version, the activation function is modeled after the natural design.

In the biological state, the activation function determines and processes all signals that cross the threshold limits. This is known as the threshold activation function in artificial neural networks. The most commonly used activation function in ANN is sigmoid function, but some alternative uses are available. For example, some activation functions such as Linear, Saturated Linear, Hyperbolic Tangent, and Gaussian are also available [23].

In network topology, the learning ability of the artificial network is the root of the topology, or the structure and components of neuron bonds. Numerous artificial architectural forms can be created. They can be distinguished by three key characters.

- The number of layers.
- Whether network feedback is allowed or not.
- The number of nodes in each network layer.

Topology can be learned by the network in determining the complexity of tasks [23]. The unprocessed signals reach the node activation function and are converted therefrom. The input and output nodes are organized in a group known as a layer. The data that reach the input nodes are processed by the weight values in the network and generates the answer. Single-layer network is used in simple-featured classification, especially for direct separable classification, but more complex networks require more layers to learn more tasks. Layers are added to create more complex networks. In multi-layer networks, one or more layers are added to process incoming signals before they can reach the output node. Having more than one hidden layer in the artificial neural network is called as deep artificial network or as deep learning [23].

In feedback networks, the basic idea is to change the input weights to minimize the differences between the network output and the actual output until an acceptable error is achieved. Updating the weights means learning the network.

3.2. Regression

Regression deals with the relationship between dependent and one or more independent variables. How the independent variables affect the dependent variable is determined by a coefficient. This coefficient is called the regression coefficient of the variable. The regression coefficient indicates the degree of dependence. The causal relationship between dependent and independent variables allows the use of this method.

In the regression analysis, the aim for the term $Y_i = \beta_0 + \beta_1 X_i$ is to produce the coefficients β_0 and β_1 that will produce the results closest to the Y value. Using independent X_i values and dependent true Y_i values in the data set, β_0 and β_1 coefficient values are determined by least squares method to minimize the deviation values of the data.

3.3. Random Forest Tree

The random forest method works by focusing on the collection of decision trees. This method was supported by Leo Breiman and Adele Cutler and based on the principle that random property selection in decision tree models was packed by increasing diversity [23].

The random forest method is a collection of decision trees based on the collection of random variables of each tree. For the P-dimensional random vector, $X = (X_1, \ldots, Xp)^T$ represents the real value or prediction value, while Y represents the real value response. $P_{XY}(X, Y)$ represents the default port. The objective is to find the $f(X)$ prediction function to predict the Y value. The prediction function $L(Y, f(X))$ is determined by an error function and is defined to reduce the expected value of the error values.

$E_{XY}L(Y, f(X))$ is the expected value index for the distribution of X and Y values.

$L(Y, f(X))$ measures how close $f(X)$ is to Y, giving penalties based on how far from Y to $f(X)$. Conventionally, L is the sum of squares of the error function for regression and is calculated by $L(Y, f(X)) = (Y - f(X))^2$, and takes the value of 0 and 1 for the classification.

The sum of the terms $h_1(x), \ldots h_j(x)$, so-called basic learning, is f and it is the forecast function of sum of basic learning. This regression is the average values [24], as shown in Equation (2).

$$f(x) = \frac{1}{j}\sum_{j=1}^{j} h_j(x). \tag{2}$$

When using the random forest tree method, the most important factor in the problem is whether to make a classification estimate or a numerical estimate. As a result of the researches in the literature and the classification problem m-try that determines the number of variables used in each tree, the related problem

is selected as the square root of the number of variables of interest in the data set, the problem is selected as the ratio of the number of variables of interest in the data to 3 in the case of numerical estimation.

3.4. Time Series

A series of historical data varying over a period form a time series. The reason for the differentiation of these values is examined by time series analysis and a model representing the process and behavior is prepared. With the help of this model, future demand is estimated. To establish the associated mathematical model, the following five factors that make up the behavior of a time series must be known [25].

Average: The data takes place around an average value.

Trend (Trend): Data tends to increase and decrease continuously. Finding a continuous line represents the first step in setting up the time series model.

Seasonal Change: It represents seasonal increases or decreases of data. The concept of seasons means that data have similar tendency in time process. The process of interest over time may increase or decrease in specific months.

Change of Period: changes in the longer process except seasonal variation of data. It can also be considered as a period change showing demand fluctuations spreading over long time.

Random Events: These are irregular and not easy to understand.

A time series model is a function of the numerical values that the dependent variable has taken in the past. Many time series data also include this process.

This can be expressed by an equation given in Equation (3)

$$x_t = a + a_1 x_{(t-1)} + a_2 x_{(t-2)} + a_3 x_{(t-3)} + \cdots + \varepsilon. \tag{3}$$

Here, a represents the constant term, while coefficients such as $a_1, \ldots$ represent the relationship of the delayed numerical values to the current numerical value. In addition, the term error ε defines random shocks. It is generally referred to as $AR(p)$.

If the delayed error term of the series also affects the current error term, the moving average process must be defined. In a moving average process, the predicted value of the variable relates to the predicted value of the error terms.

A moving average process ,denoted as $MA(q)$, can be expressed in Equation (4)

$$x_t = e_t - a_1 e_{(t-1)} - \cdots \quad ; \quad t = 1,2,\ldots,n. \tag{4}$$

Most time series include both AR and MA processes. In addition, I denotes the trend contained by the series. $ARIMA(p, d, q)$ tells us that the time series includes the $AR(p)$ process while expressing the situation caused by the delay of p-degree itself. In the other notation, q expresses the relationship of error terms with historical values, while expressing the random process, $MA(q)$ is one of the methods of flattening error terms. In addition to these representations, the presence of a non-stationary state or encountering a stationary time series is indicated by d and is clearly indicated at which level the time series is considered. Relating term can be expressed as in Equation (5) [26]

$$Y_t = m + a_1 y_{(t-1)}, \ldots, a_p y_{(t-p)} + u_t - b_1 u_{(t-1)} - \cdots - b_q u_{(t-q)}. \tag{5}$$

3.5. Multiple Seasonality Time Series

Time series data usually assume a single seasonal effect, but in some cases two or more seasons may be affected. The TBATS model can be used for this purpose. The model is named by the combination of the initials of the words Trigonometric, Box-Cox transform, ARMA errors, Trend, and Seasonal components.

Being trigonometric allows the model to work with non-integer values. The Box-Cox transform model is a transformer that converts non-normal dependent variables into normal dependent variables. ARMA errors refer to the degree parameters mentioned in the time series. The trend is used to determine periods of increase or decrease in the time series.

The TBATS model is used to identify increases in non-stationary periods. Parameters of the model Trigonometric Box, Box-Cox parameter φ, ARMA parameters p, q, while defining seasonal non-integer frequencies m_t, k_t parameters are used [27]. Equation (6) shows the TBATS model

$$TBATS(\omega, \varphi, p, q, \{m_1, k_1\}, \{m_2, k_2\}, \dots, \{m_t, k_t\}). \tag{6}$$

4.   EXPERIMENTAL RESUTS

Turkey's natural gas energy consumption was estimated by machine learning techniques with real data obtained from www.jodidata.org. It is aimed to compare the real data with the prediction data and select the best model according to the prediction performance.

Using the regression method, the Gross Inland Deliveries Observed dependent variable and independent variables of "Import" and "Of Which: Electricity.and.Heat.Generation" were used to create a regression model. The scatter diagram of the variables is shown in Figure 2.

Independent variable of "Import" stands for amount imported or when it has crossed the physical boundaries of the country, customs clearance has taken place or not. The independent variable "of which: Electricity and Heat Generation" covers the deliveries of natural gas for the generation of electricity and heat in power plants. Both main-activity and auto producer plants are included. The simple linear equation that represents the relationship between dependent and two independent variables is found as Gross Inland Deliveries = $-1.16 \times 10^3 + 1.326 \times$ Import $- 1,18 \times 10^{-1} \times$ of which: Electricity and Heat Generation.
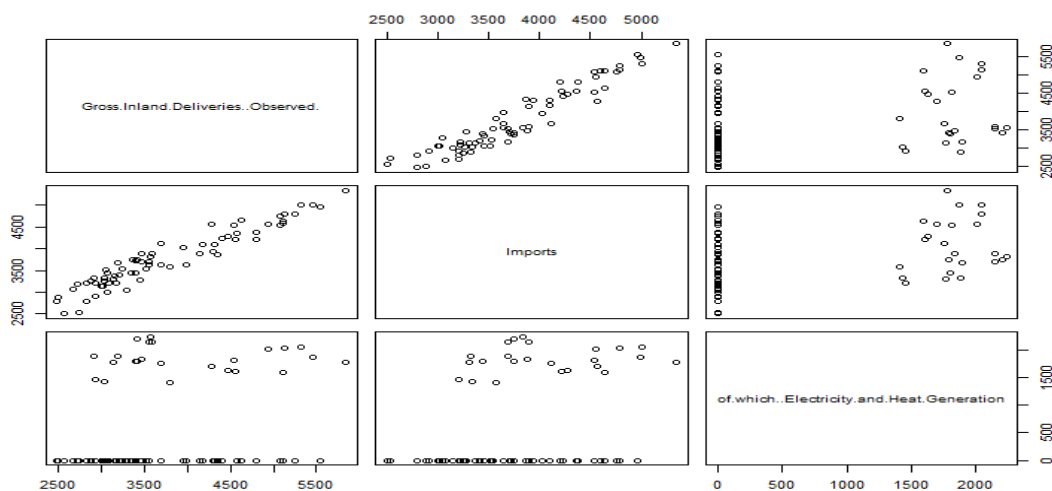

Figure 2. *Scatter diagram of the corresponding variables*

Multiple R-squared value of the regression model was found to be 0.9201, meaning that the independent variables that we give to the model can explain the value we want to estimate by 92%.

The mean squares error method will be used to measure estimation errors. It is the average of the squares of the difference between the observed and estimated values, as shown in Equation (7)

$$MSE = \frac{\sum_{i=1}^{n}(O_i - E_i)^2}{n} = \frac{\sum_{i=1}^{n}(Error_i)^2}{n}. \tag{7}$$

Estimation of 36 months was made by using regression model. Actual values were compared with the estimation results. The average squares error was calculated as 102969.

In the artificial neural network model, Import and Electricity.and.Heat.Generation variables are used as input data and Gross Inland Deliveries Observed is estimated as output data. 80% of the data was used as the educational data of the model and remaining 20% was used as the test data. Accordingly, by using the data between 2010 and 2015, the model was educated and the data for the 36 months were estimated.

In which layers the model will work best, the number of layers and the number of nodes have been changed to improve the model. Developed models are decided as the most appropriate model with the least error result. Among the models with the lowest error rate in multiple model trials, the error rate of the 4-node model in Figure 3 was lower than the error rate of the 3-node model in the first layer in the second layer in Figure 4. For this reason, the most suitable model among artificial neural networks is the single-layer 4-node model.
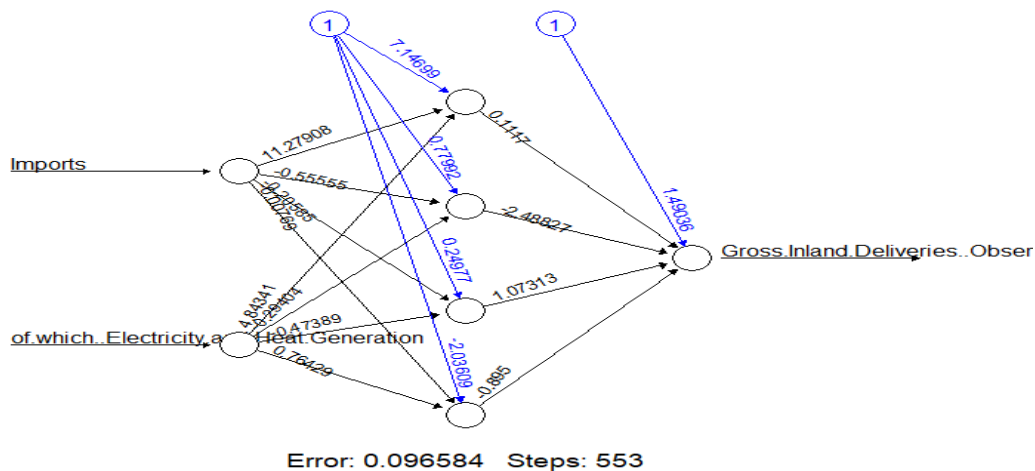

Figure 3. *An artificial neural network model with four nodes in the first layer*
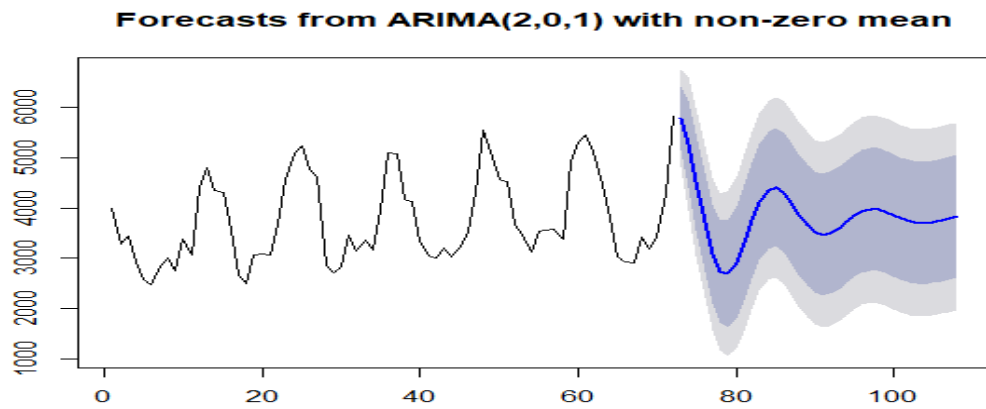
Error: 0.104204   Steps: 58

Figure 4. *Artificial neural network model with 5 nodes in first layer and 3 nodes in the second layer*

The mean squares error of the artificial neural network model was found to be 93381 when predicted values compared with actual values of 36 months.

Since the problem is a numerical estimation in the random forest tree method, according to the information obtained from the research in the literature, there is information to decide how many variables to be used in each tree up to 3 part of the number of variables that are important for the problem. In addition, it is stated that the estimator will decide how many trees to choose according to the size of the data.

80% of the data was used for training of the model and remaining 20% of the data was used for testing. Accordingly, by using the data between 2010 and 2015, the model was educated and the data for the 36 months were estimated. The ntree value in the model specifies how many trees to create, while mtry specifies how many variables to use in each tree.

In the first model made with random forest model (when mtry = 1, ntree = 100), the mean squares of error of the estimate were found to be 377099. In the second model made with random forest model (when mtry = 1, ntree = 1000), the mean squares of error of the estimation were 341566. In the third model made with random forest model (when mtry = 1, ntree = 10000), the mean error squares of the estimation were found to be 340344. In other model studies, increasing the number of trees, ie ntree value, did not improve significantly in the mean error frames.

In the time series model, since the data is monthly based, the frequency value is selected as 12 months and the data set start time is applied starting from 2010. Figure 5 shows the estimated 36 months of the ARIMA model at an estimated 80% and 90% confidence intervals. The mean squares error was calculated as 10508187 in the comparison of the estimates with the ARIMA model and the actual values.

Figure 5. *Estimation values and confidence intervals produced by time series analysis*

The multiple seasonality time series model was created with the idea that there may be more than one seasonal effect. Turkey is a country that moves according to the Hijra calendar, social and cultural sense. In this context, natural gas consumption is expected to decrease during religious holidays. The effect of this social and cultural impact on natural gas consumption is also examined with this model.

The TBATS method was used for the multiple-seasonality effect model and a graph showing the effect of multiple seasonal effects is given in Figure 6. Seasonal parameters of 6 months is selected and the effect of the calendar is calculated as 354.37 days. This is equivalent to 11.64 months, the effect of the gregorian calendar 365.25 days, which will be taken as 12 months. Figure 7 shows the estimated 36 months of the TBATS model at an estimated 80% and 90% confidence intervals. The mean squares error was calculated as 294752 in comparison with the estimations made with the established multi-seasonal effect model and the actual values.
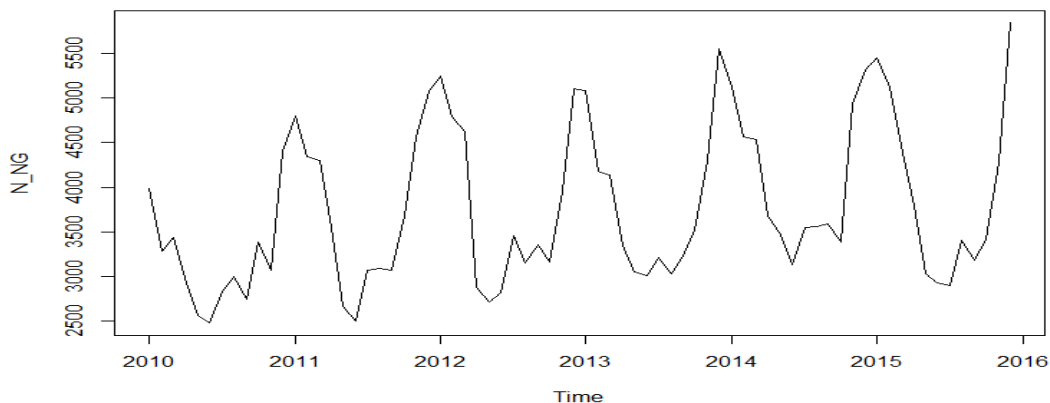

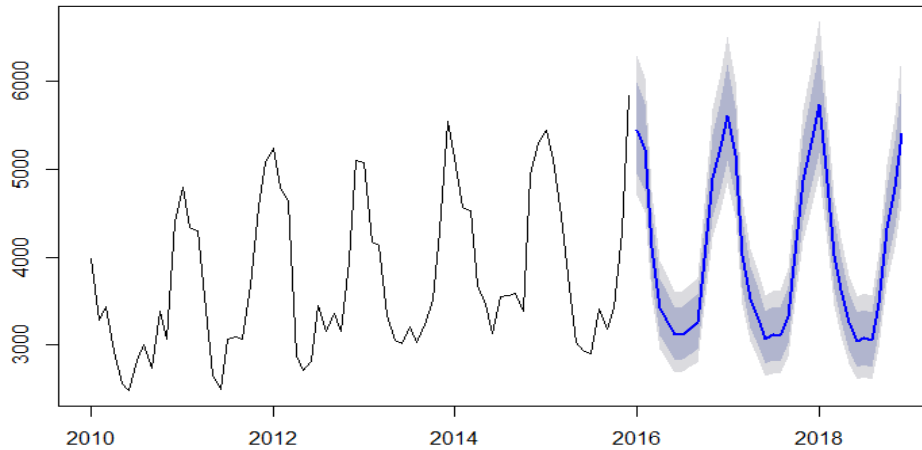***Figure 6.*** *Effects of multiple seasonal effects on consumption*

Figure 7. *Estimation of multi seasonal effect model with confidence intervals*

5. DISCUSSION

Natural Gas is one of the primary energy sources with increasing importance due to being environmentally friendly and cheap in meeting the energy demand in the world. Because of Turkey's booming economy and increasing population, its need for energy is constantly increasing. Despite the increasing energy needs of Turkey, it's very limited natural gas reserves cannot meet their own needs, leading to high external dependency. Because of the fact that a large portion of the world's natural gas reserves is in the Middle East and the Caspian region and Europe's energy needs should be met by these reserves, Turkey has a great geopolitical importance. Despite this situation, it is the fact that large amounts of money is paid for Turkey's natural gas consumption. In order to sustain the growth of our country's economy, which has a high dependency on foreign energy in natural gas energy, it is important to estimate natural gas consumption for the upcoming years.

Machine learning is a set of techniques that have been increasingly used in recent years and have been very successful in forecasting.

In this study, supervised machine learning algorithms involving artificial neural networks, regression, random forest trees, time series and multiple seasonality time series are utilized to estimate Turkey's natural gas consumption. Real consumption data of natural gas on a monthly basis between 2010 and 2018 were used for estimation. The data between the years 2010-2015 was used as training data while the data between 2016-2018 were used as test data in order to see the effectiveness of the algorithms. In determining the effectiveness of the algorithms, the sum of the least squares of error (MSE), which is widely used in the literature, is taken into consideration. Experimental results showed that artificial neural network technique produces the best performance (MSE=93381). The regression method was followed by artificial neural networks (MSE = 102969), while the Multi-Term Time Series method (MSE = 294752) came in third. Random Forest Tree (MSE = 341566) and Time Series (MSE = 105187) were the two worst-performing machine learning methods. It can be concluded that artificial neural networks, as confirmed by the existing literature, has superior performance among the techniques and can be thought of the first technique in order to use in the future for other estimation settings.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

[1] T.C. Enerji ve Tabii Kaynaklar Bakanlığı https://www.enerji.gov.tr/tr-TR/Sayfalar/Dogal-Gaz Publish date April 18, 2019. Data accessed May 15, 2019.

[2] Bayrac, H.N., "The Economic Structure of International Natural Gas Market and Applied Policies", Eskişehir Osmangazi Üniversitesi İİBF Dergisi, 13 (3): 13-36, (2018).

[3] Bayrac, H.N., "The Economic Structure of International Natural Gas Market and Applied Policies", Eskişehir Osmangazi Üniversitesi İİBF Dergisi, 13 (3): 13-36, (2018).

[4] Brett, L., Machine Learning with R Second Edition, Packt Publishing, Birmingham, (2015).

[5] Es, H.A., Hamzaçebi, C., "Forecasting The Net Energy Demand Of Turkey By Artificial Neural Networks", Journal of The Faculty of Engineering and Architecture of Gazi University , 29 (3): 495-504, (2014).

[6] Karaca, C., Karacan, H., "Investigation Of Factors Affecting Demand For Electricity Consumption With Multiple Regression Method", Selcuk University Journal of Engineering, Science and Technology, 4 (3): 183-195, (2016).

[7] Hamzaçebi, C., Kutay, F., "Electrıc Consumption Forecasting of Turkey Using Artifial Neural Networks Up To Year 2010, Journal of The Faculty of Engineering and Architecture of Gazi University, 19 (3): 227-233, (2004).

[8] Şenol, Ü., Musayev, Z., "Estimating Wind Energy Potential by Artificial Neural Networks Method", Bilge International Journal of Science and Technology Research, 1(1): 23-31, (2017).

[9] Makas, Y., Karaatlı, M., "Multi-Period Estimation Of Hydroelectric Power Generation With Artificial Neural Network", The Journal of Faculty of Economics and Administrative Sciences, 21 (3): 757-772, (2016).

[10] Aksoy, S.A., Eryiğit, E., Hashimova, N., İşbilir, M., Avşar, Z.M. and Köksal, G., A "Prediction And Bidding System Design For Wind Power Generation", Endüstri Mühendisliği Dergisi, 24 (3-4): 4-15, (2013).

[11] Çoban, O., Özcan, C.C., "Sektörel Açıdan Enerjinin Artan Önemi: Konya İli İçin Bir Doğalgaz Talep Tahmini Denemesi", SÜ İİBF Sosyal ve Ekonomik Araştırmalar Dergisi, 11 (22): 85-106, (2011).

[12] Dündar, C., Oğuz, K., Dokuyucu, K. and Bacanlı, H.,"Kısa Süreli Rüzgar Enerjisi Tahmini", VI. Yeni ve Yenilenebilir Enerji Kaynakları Sempozyumu, Kayseri, 2-12, (2011).

[13] Yavuzdemir, M., "Türkiye'nin Kısa Dönem Yıllık Brüt Elektrik Enerjisi Talep Tahmini", MSc. Thesis, Ankara Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı, Ankara, (2014).

[14] Eren, T., "Grey Forecasting Method In Natural Gas Consumption Planning And Turkey Implementation", MSc. Thesis, İstanbul Ticaret University, Institute of Science and Technology, İstanbul, (2017).

[15] Bianco, V., Manca, O., Nardini, S., "Electricity consumption forecasting in Italy using linear regression models", Energy, 34 (9): 1413–1421, (2009).

[16] Al-Fattah, S.M., "Time Series Modeling for U.S. Natural Gas Forecasting", International Petroleum Technology Conference, Doha-Qatar, (2005).

[17] Demirel, Ö.F., Zaim, S., Çalışkan, A. and Özuyar P., "Forecasting Natural Gas Consumption in İstanbul Using Neural Networks And Multivariate Time Series Methods", Turkısh Journal Of Electrıcal Engıneerıng Computer Sciences, 20 (5): 695-711, (2012).

[18] Busse, S., Halmholz P. and Weinmann M., "Forecasting Day Ahead Spot Price Movements Of Natural Gas – An Analysis Of Potential Influence Factors On Basis Of A NARX Neural Network", Multikonferenz Wirtschaftsinformatik, 13(1): 1395 - 1406, (2012).

[19] Vıtullo, S.R., Brown, R.H., Corlıss, G.F. and Marx, B.M., "Mathematical Models For Natural Gas Forecasting", Canadian Applied Mathematics Quarterly, 17 (4): 807-827, (2009).

[20] Karadere, Y., Özdemir, G., Aydemir, E., "Breeder hybrid algorithm approach for natural gas demand forecasting model", Energy, 141: 1269-1284, (2017).

[21] Özdemir, G., Aydemir, E., Olgun, M.O., Mulbay, Z., "Forecasting of Turkey natural gas demand using a hybrid algorithm", Energy Sources, Part B: Economics, Planning and Policy, 11(4): 295-302, (2016).

[22] Olgun, M.O., Özdemir, G., Aydemir, E., "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(1): 15-20, (2012).

[23] Brett, L., Machine Learning with R Second Edition, Packt Publishing, Birmingham, (2015).

[24] Cutler, A., Stevens J.R., Cutler D.R., Ensemble Machine Learning: Methods and Applications, Springer, (eds.) Zhang, C., Ma, Y., New York, (2012).

[25] Toomey, J.W., "MRP II Planning for Manufacturing Excellence", Hardbaound, New York, (1996).

[26] Sandy, R., "Statistics for Business and Economics", Mc-Graw Hill Higher Education, C.USA,(1990).

[27] De Livera, A.M., Hyndman, R.J. and Snyder, R.D., "Forecasting time series with complex seasonal patterns using exponential smoothing", Journal of the American Statistical Association, 15(9): 2-39, (2010).