



Received: June 23, 2020
Accepted: November 25, 2020
Published Online: December 31, 2020

AJ ID: 2020.08.02.OR.03
DOI: 10.17093/alphanumeric.756651
Research Article

The Estimation of Turkey's Energy Demand Through Artificial Neural Networks and Support Vector Regression Methods

Mehmet Kayakuş, Ph.D. * 

Assist. Prof., Department of Management Information Systems, Faculty of Social Sciences and Humanities, Akdeniz University, Antalya, Turkey.
mehmetkayakus@akdeniz.edu.tr

* Akdeniz Üniversitesi Manavgat Sosyal ve Beşeri Bilimler Fakültesi 07600 Manavgat, Antalya, Turkey

ABSTRACT

Energy demand and consumption are very important for the development and progress of countries. Energy demand is increasing rapidly day by day, especially in developing countries. Energy policies should be determined correctly to sustain the industry sector and make the right investments. Forecasting energy demand in the near and long term is important for the strategy that countries will follow. In this study, by using the monthly electricity energy data realized in Turkey between January 2016 and March 2020 and other data affecting this, a model to estimate electrical energy consumption was developed. In this model, artificial neural networks (ANN) and support vector regression (SVR) were used as methods. This study used 15 independent variables as the input value, and Turkey's energy consumption value as the dependent variable was estimated. Correlation, coefficient of determination, MAE, MSE, RMSE, and MAPE statistical methods were used to measure success and error rate, and both models were found to have acceptable error values and success estimation rates. According to the results, it was concluded that the ANN method was more successful than the SVR method.

Keywords:

Energy Consumption, Artificial Neural Networks, Support Vector Regression



1. Introduction

Countries have to consume more energy to continue their progress and to reach the status of developed countries. It shows that energy demand will be important in the coming years as it is today. The need for energy is increasing day by day for the development and progress of countries.(Bayramođlu, Pabuçcu, & Boz, 2017).

Forecasting energy demand is very important in terms of making sustainable and correct investments. Accurate planning will both provide economic benefits and reduce dependence on other countries.

There are many different forecasting techniques in the literature for energy demand forecasting. For example, energy demands were estimated using time series (Özden & Öztürk, 2018), gray prediction (Akay & Atak, 2007), regression model (Karaca & Karacan, 2016; Yüzük, 2019), ANFIS and ARMA models (Bayramođlu et al., 2017; Demirel, Kakilli, & Tektaş, 2010), genetic algorithm (Yiđit, 2011), fuzzy logic, hybrid models (Çınar, 2007), support vector regression and artificial neural network forecasting techniques.(Es, Kalender, & Hamzaçebi, 2014).

One of the forecasting tools frequently used in the energy demand forecasting field is ANN. Sözen et al. used ANN in order to make a prediction for Turkey's net energy consumption (Sözen & Arcaklioglu, 2007; Sözen, Arcaklioglu, & Özkaymak, 2005), Pao for Taiwan's electricity energy consumption (Pao, 2006), Kavaklioglu et al. for Turkey's electricity consumption (Kavaklioglu, Ceylan, Ozturk, & Canyurt, 2009), Geem and Roper for the energy demand of South Korea (Geem & Roper, 2009), Ekonomou for Greece's total energy demand (Ekonomou, 2010), Limanond et al. for Thailand's transportation energy demand (Limanond, Jomnonkwao, & Srikaew, 2011), Kankal et al. for Turkey's overall energy consumption (Kankal, Akpınar, Kömürçü, & Özşahin, 2011), Geem for South Korea's transportation energy demand (Geem, 2011), Bilgili et al. for the demand of residential and industrial sectors in Turkey (Bilgili, Sahin, Yasar, & Simsek, 2012), Kocadayı et al. for TR81 region energy consumption (Kocadayı, ErKaymaz, & Uzun, 2017), Özden at el. for the electricity consumption of an industrial area (Özden & Öztürk, 2018), Yüzük for energy consumption with 17 independent variables in his thesis. (Yüzük, 2019) (Es et al., 2014).

One of the methods frequently used on energy demand and consumption estimation is the support vector regression method. Kaytez et al. used SVR in order to predict Turkey's electricity consumption (Kavaklioglu, 2011), Wang et al. for China's hydropower consumption (Wang, Yu, Tang, & Wang, 2011), Jung et al. for building energy consumption (Jung, Kim, & Heo, 2015), Yang et al. to analyse the data affecting electricity consumption.

In this study, in order to estimate the electricity consumption in Turkey, a model consisting of 15 independent variables was designed. Artificial neural networks and support vector regression methods were used to estimate this model. In this study, the data such as hydraulic, imported coal, hard coal, lignite, natural gas, sun, wind, geothermal, biomass, asphaltite, fuel oil, electricity imports, Turkey's temperature average, Turkey's monthly population, the number of days worked, the number of vacation days between 2016-2020 were used as the independent variable.

2. Methods

2.1. Artificial neural networks (ANNs)

ANNs that mimic the way the human brain works have an important place among artificial intelligence technologies. ANN methodology has many important features such as learning from data, making generalizations, working with an unlimited number of variables (Ataseven, 2013). As seen in Figure 1, the artificial neural network model consists of 3 layers as input layer, hidden layer and output layer.

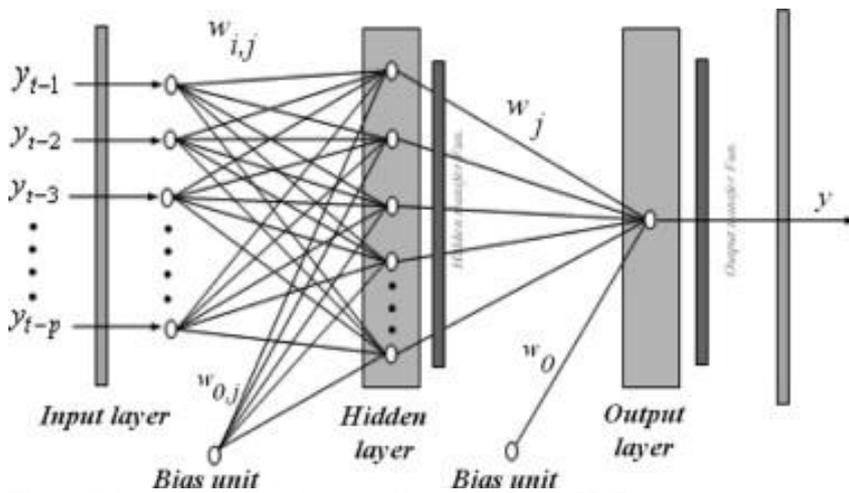


Figure 1. Neural network structure(Khashei & Bijari, 2010)

The input layer is the first layer and ensures that external data is received into the artificial neural network. The input layer consists of parameters that affect the problem, and the number of neurons in the input layer is shaped by the number of parameters. The output layer functions to transmit information outside. The hidden layer is located between the input layer and the output layer. Neurons in the hidden layer have no connections with the external environment. They only receive data from the input layer and send the data to the output layer. There are one or more hidden intermediate layers between the input layer and the output layer from which the data are obtained. Each neuron in a layer is connected to all neurons in a top layer with different numerical weights ($w_{i,j}$). Weights indicate the effect of the neuron in the preceding layer on the neuron in the upper layer. Positive weight values indicate reinforcement and negative values indicate obstacle (Asilkan & IRMAK, 2009; Gallant & Gallant, 1993).

ANN is classified under two architectural structures, "feed-forward" and "feedback", depending on the way the cells attach (Slaughter & Hobson, 2006). Feed-forward networks have a network structure in which data are transferred forward only from input units to output units. The outputs of neurons in one layer are given as inputs to the next layer through weights. There is no connection between neurons in the same layer or to the previous layer. Feedback networks have a network structure in which the data flow can be not only forward but also backwards(Asilkan & IRMAK, 2009). It obtains the information to be used in changing the weights by comparing the output value it generates according to the network input information to the desired value. Training continues until the difference between the entered value and the desired value is less than the predetermined value as the error value. When the error value

falls below the desired value, all weights are fixed, and the training process is ended. (Çuhadar, Güngör, & Göksu, 2009).

Each unit on the network takes the value from the neuron that precedes it, and weighted sums are calculated. The input data are multiplied by weights and advance to the next layer. Weights are chosen randomly at the beginning. Multiplication results are collected in hidden layers, and the result is passed through a transfer function. Signals coming from the input layer move forwardly towards the output layer consecutively.

$$v_i(t) = w_1x_1(t) + w_2x_2(t) + \dots + w_{ij}x_n(t) + \theta_j = \sum_{i=1}^n w_{ij}x_n(t) + \theta_j \quad (1)$$

For $i = 1, \dots, x$, i shows network inputs, w_{ij} shows the link weights between the i input and the j cell. θ_j expresses the threshold (bias) value of the cell. The output of the cell is obtained by adding the result of the addition with the threshold value and passing it through the activation function. The activation functions can take various forms. The type of activation function is indicated by the state of the neuron in the network. The most commonly used activation function in the output layer is a linear function, since using the nonlinear activation function in this layer may cause distortion in the predicted output. Logistics and hyperbolic functions are often used as hidden layer transfer activation functions. (Khashei & Bijari, 2010).

2.2. Support Vector Regression (SVR)

Support Vector Machines (SVMs) are a kind of machine learning tool that can solve classification, regression and innovation detection problems with better generalization compared to other traditional learning algorithms. (Ali & Smith-Miles, 2006). It was first proposed by Vapnik in 1995 for classification and regression type problem solutions (Vapnik, 1995). SVM operates based on structural risk minimization and is also successful in applications with high dimensions but few data. The type of SVM used in classification applications is known as SVC (Support Vector Classification), and the type used in regression applications is known as SVR (Shen, Pei, Fisher, & Lee, 2006). The purpose of SVM is to separate the input data, whose classes are defined by the class tag, as shown in Figure 2, into two separate classes by specifying the most appropriate hyperplane (Fidan, Uzunhisarcıklı, & Çalığıuşu, 2019).

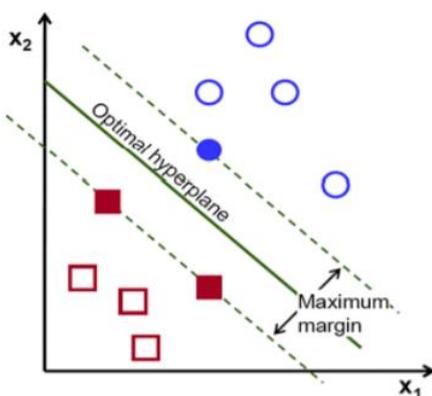


Figure 2. Optimal Hyperplane and Margin of SVM (Fidan et al.)

SVM can be divided into Linear and Nonlinear SVRs depending on the state of data discrimination.

Linear SVM is the simplest SVM model that can only be applied to distinguishable linear data. Considering that the training data set is $D = \{(\vec{x}_i, y_i), i = 1 \dots N\}$ and it consists of N elements $y_i \in \{-1, +1\}$, class tag $\vec{x}_i \in R^n$ is any example in ϵ and n dimensional space. In the $f(\vec{x}) = \vec{w}^T \vec{x} + b$ expression \vec{w}^T indicates the normal function of the decision function, \vec{x} expression indicates the points on this line and b indicates the trend value. The goal is to find \vec{w}^T and b with the help of training data, and ultimately to train the system (Küçükşille & Ateş).

Functions that can separate linearly distinctive data sets are very limited for practical applications. This is because no results can be obtained by using linear in most practical applications. In such cases, SVM moves the input space to a higher level of space through various transformation processes and tries to perform the linear separation (Çomak, 2008)

3. Application

Turkey's energy demand was estimated using 15 independent variables in this study. In this study, the data such as hydraulic, imported coal, hard coal, natural gas, sun, wind, geothermal, biomass, asphaltite, fuel oil, electricity imports, Turkey's temperature average, Turkey's monthly population, the number of days worked, the number of vacation days between 2016-2020 were used as the independent variable.

There was a total of 52 data in the study. Since the number of data was low, the number of data was increased by using different cross-validation methods and different groupings. In this method, the data are initially divided randomly as test and training data. While training data are used in the model installation phase, test data are not used in model setup, and the accuracy of the model is tested on this new data set (Bishop, 1995).

In the study, decimal scale normalization method was used to improve the performance of machine learning methods and increase the accuracy rate. In this technique, the decimal point of the values of the attribute is moved. This movement of decimal points completely depends on the maximum value among all values in the attribute. The decimal Scale normalization formula is,

$$v^i = \frac{v}{10^j} \quad (2)$$

Where, v_i is the scaled values, v is the range of values, j is the smallest integer $\text{Max}(\text{lvil}) < 1$.

In the study, as seen in Figure 3, the artificial neural networks model was designed using a backpropagation algorithm consisting of an input layer, two hidden layers and an output layer. There are 15 neurons in the input layer and 1 neuron in the output layer. The number of neurons in the hidden layer was determined by trial and error method to give the best results for the training and verification clusters, and it was designed to have 3 neurons in each hidden layer. Sigmoid activation function was

used as hidden layer activation function, and linear activation function was used as output layer activation function.

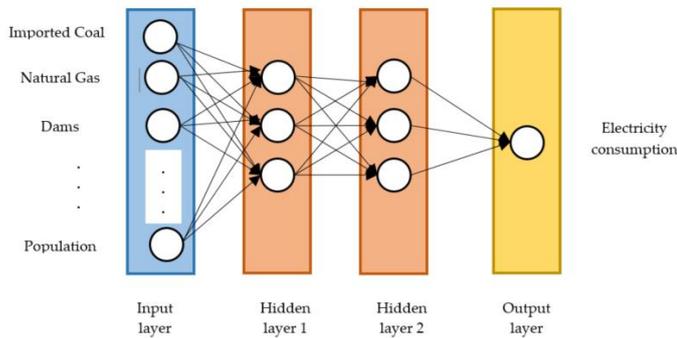


Figure 3. Artificial neural networks model

The nonlinear method was used as the SVR method. Different combinations were tested for SVR parameters of C and ϵ , and finally we determined that we obtained a good performance with ϵ 0.01 and C 500, and we decided to use it. Polynomial was preferred as the core function, as it yielded the most successful result as a result of the tests. The Polynomial kernel formula is,

$$K(x_i, x_j) = [\gamma(x_i \cdot x_j) + C]^d \tag{3}$$

To measure the success of the estimates, Coefficient of determination (R^2) in Equation 21, Mean Squared Error in Equation 22 (Bilgili et al.), Root Mean Squared Error (RMSE) in Equation 23 and Mean Absolute Error (MAE) in Equation 24 were used.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - Y_i^*)^2 \tag{4}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - Y_i^*}{Y_i} \right|^2} \tag{5}$$

$$MAE = \frac{\sum_{i=1}^n |Y_i - Y_i^*|}{n} \tag{6}$$

where Y^i is the observation value and Y_i^* is the predicted value, Y_{ave} average of observation values. Table 1 shows the comparison of the success and error performance of the models.

	Correlation	R2	MAE	MSE	RMSE	MAPE
ANN	0,936	0,913	0,009	0,003	0,003	0,049
SVR	0,900	0,851	0,016	0,009	0,011	0,074

Table 1. The comparison of the success and error performance of the models

According to the MAPE error criteria in the model created, the estimation error of the test data was 4.9% for ANN and 7.4% for the SVR. Estimation models with MAPE values below 10% are considered as models with "high accuracy" or "very good" accuracy levels (Lewis, 1982; Witt & Witt, 1992). Here, the ANN model was found to be more successful since the method with the lowest value would be considered more

successful. MSE and RMSE approaching zero indicates that the model is more successful and shows high performance (Singh, Basant, Malik, & Jain, 2009). According to MSE and RMSE, ANN was found to be a more successful model as well. MAE is a value that measures the average size of errors in a series of estimates. According to Table 1, ANN yields estimates with fewer errors, since its being close to zero indicates a less erroneous estimate.

The correlation coefficient is used to determine whether there is a relationship between the true value and the estimated value, and if any, the direction (forward / reverse) and severity of this relationship. If there is a positive relationship, one is increasing while the other one is also increasing. If there is a negative relationship, it indicates that one is increasing and the other is decreasing. If the correlation coefficient is greater than 0.8, it is interpreted that there is a very high correlation. The correlation coefficient was found to be 0.936 for ANN and 0.9 for SVR. Although both methods seem to have a strong relationship, ANN has a stronger correlation.

Coefficient of determination allows us to determine how much of the variation in one variable is explained by the other variable. The high value of this value indicates that we can explain the dependent variable with the independent variables in our model. Coefficient of determination was calculated as 0.913 for ANN and 0.851 for SVR. Although Coefficient of determination has acceptable values in the two methods, it is seen that ANN is a more successful model.

4. Conclusion

In this study, a model that estimates monthly energy consumption using 15 different independent variables was developed. Data on hydraulic, imported coal, hard coal, lignite, natural gas, sun, wind, geothermal, biomass, asphaltite, fuel oil, electricity imports, Turkey's temperature average, Turkey's monthly population, the number of days worked, the number of vacation days were used as the independent variable. In the model, the data between 2016-2020 were used, and ANN and SVR were employed as a method.

Correlation, Coefficient of determination (R^2), MAE, MSE, RMSE, MAPE statistical methods were used to measure the success and errors of the methods. According to the study, it was found that both methods were acceptable and predicted with high accuracy, but ANN was found to be more successful.

Meeting the energy demand in a sustainable way is vital for the development and progress of the countries. Accurate estimation of energy consumption ensures investment planning and appropriate use of resources. The industrial sector and energy policy will be shaped according to this information.

More successful and stable estimates can be made by researching and using different variables for estimating energy consumption.

References

- Akay, D., & Atak, M. (2007). Grey prediction with rolling mechanism for electricity demand forecasting of Turkey. *Energy*, *32*(9), 1670-1675.
- Ali, S., & Smith-Miles, K. A. (2006). *Improved Support Vector Machine Generalization Using Normalized Input Space*, Berlin, Heidelberg.
- Asilkan, Ö., & IRMAK, S. (2009). İkinci el otomobillerin gelecekteki fiyatlarının yapay sinir ağları ile tahmin edilmesi. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, *14*(2), 375-391.
- Ataseven, B. (2013). Yapay Sinir Ağları İle Öngörü Modellemesi. *Öneri Dergisi*, *10*(39), 101-115.
- Bayramoğlu, T., Pabuçcu, H., & Boz, F. Ç. (2017). Türkiye için anfis modeli ile birincil enerji talep tahmini. *Ege Akademik Bakis*, *17*(3), 431-445.
- Bilgili, M., Sahin, B., Yasar, A., & Simsek, E. (2012). Electric energy demands of Turkey in residential and industrial sectors. *Renewable and Sustainable Energy Reviews*, *16*(1), 404-414.
- Bishop, C. M. (1995). *Neural networks for pattern recognition*: Oxford university press.
- Çınar, D. (2007). *Hidroelektrik enerji üretiminin hibrid bir model ile tahmini*. Fen Bilimleri Enstitüsü,
- Çomak, E. (2008). *Destek vektör makinelerinin etkin eğitimi için yeni yaklaşımlar*. Selçuk Üniversitesi Fen Bilimleri Enstitüsü,
- Çuhadar, M., Güngör, İ., & Göksu, A. (2009). Turizm Talebinin Yapay Sinir Ağları İle Tahmini Ve Zaman Serisi Yöntemleri İle Karşılaştırmalı Analizi: Antalya İline Yönelik Bir Uygulama. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, *14*(1), 99-114.
- Demirel, Ö., Kakilli, A., & Tektaş, M. (2010). Anfis ve arma modelleri ile elektrik enerjisi yük tahmini. *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, *25*(3).
- Ekonomou, L. (2010). Greek long-term energy consumption prediction using artificial neural networks. *Energy*, *35*(2), 512-517.
- Es, H. A., Kalender, F. Y., & Hamzaçebi, C. (2014). Yapay sinir ağları ile Türkiye net enerji talep tahmini. *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, *29*(3), 495-504.
- Fidan, U., Uzunhisarcıklı, E., & Çalıklı, İ. (2019). Classification of Dermatological Data with Self Organizing Maps and Support Vector Machine. *Afyon Kocatepe Üniversitesi Fen Ve Mühendislik Bilimleri Dergisi*, *19*(3), 894-901.
- Gallant, S. I., & Gallant, S. I. (1993). *Neural network learning and expert systems*. MIT press.
- Geem, Z. W. (2011). Transport energy demand modeling of South Korea using artificial neural network. *Energy policy*, *39*(8), 4644-4650.
- Geem, Z. W., & Roper, W. E. (2009). Energy demand estimation of South Korea using artificial neural network. *Energy policy*, *37*(10), 4049-4054.
- Jung, H. C., Kim, J. S., & Heo, H. (2015). Prediction of building energy consumption using an improved real coded genetic algorithm based least squares support vector machine approach. *Energy and buildings*, *90*, 76-84.
- Kankal, M., Akpınar, A., Kömürcü, M. İ., & Özşahin, T. Ş. (2011). Modeling and forecasting of Turkey's energy consumption using socio-economic and demographic variables. *Applied Energy*, *88*(5), 1927-1939.
- Karaca, C., & Karacan, H. (2016). Çoklu regresyon metoduyla elektrik tüketim talebini etkileyen faktörlerin incelenmesi.
- Kavaklıoğlu, K. (2011). Modeling and prediction of Turkey's electricity consumption using Support Vector Regression. *Applied Energy*, *88*(1), 368-375.
- Kavaklıoğlu, K., Ceylan, H., Ozturk, H. K., & Canyurt, O. E. (2009). Modeling and prediction of Turkey's electricity consumption using artificial neural networks. *Energy Conversion and Management*, *50*(11), 2719-2727.
- Khashei, M., & Bijari, M. (2010). An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Systems with applications*, *37*(1), 479-489.
- Kocadayı, Y., ErKaymaz, O., & Uzun, R. (2017). Yapay sinir ağları ile Tr81 bölgesi yıllık elektrik enerjisi tüketiminin tahmini. *BİLDİRİ ÖZETLERİ KİTABI*, 239.
- Küçükşille, E. U., & Ateş, N. (2013). Destek Vektör Makineleri ile Yaramaz Elektronik Postaların Filtrelenmesi. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, *6*(1), 1-7.

- Lewis, C. D. (1982). *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*. Butterworth-Heinemann.
- Limanond, T., Jomnonkwao, S., & Srikaew, A. (2011). Projection of future transport energy demand of Thailand. *Energy policy*, 39(5), 2754-2763.
- Özden, S., & Öztürk, A. (2018). Yapay sinir ağları ve zaman serileri yöntemi ile bir endüstri alanının (ivedik OSB) elektrik enerjisi ihtiyaç tahmini. *Bilişim Teknolojileri Dergisi*, 11(3), 255-261.
- Pao, H.-T. (2006). Comparing linear and nonlinear forecasts for Taiwan's electricity consumption. *Energy*, 31(12), 2129-2141.
- Shen, J., Pei, Z., Fisher, G., & Lee, E. (2006). Modelling and analysis of waviness reduction in soft-pad grinding of wire-sawn silicon wafers by support vector regression. *International journal of production research*, 44(13), 2605-2623.
- Singh, K. P., Basant, A., Malik, A., & Jain, G. (2009). Artificial neural network modeling of the river water quality—a case study. *Ecological Modelling*, 220(6), 888-895.
- Slaughter, G. E., & Hobson, R. S. (2006). *Artificial Neural Network for Temporal Impedance Recognition of Neurotoxins*. Paper presented at the The 2006 IEEE International Joint Conference on Neural Network Proceedings.
- Sözen, A., & Arcaklioglu, E. (2007). Prediction of net energy consumption based on economic indicators (GNP and GDP) in Turkey. *Energy policy*, 35(10), 4981-4992.
- Sözen, A., Arcaklioglu, E., & Özkaymak, M. (2005). Turkey's net energy consumption. *Applied Energy*, 81(2), 209-221.
- Vapnik, V. (1995). *The Nature of Statistical Learning Theory* Springer, 2000. *Google Scholar Google Scholar Digital Library Digital Library*.
- Wang, S., Yu, L., Tang, L., & Wang, S. (2011). A novel seasonal decomposition based least squares support vector regression ensemble learning approach for hydropower consumption forecasting in China. *Energy*, 36(11), 6542-6554.
- Witt, S. F., & Witt, C. A. (1992). *Modeling and forecasting demand in tourism*: Academic Press Ltd.
- Yiğit, V. (2011). Genetik algoritma ile Türkiye net elektrik enerjisi tüketiminin 2020 yılına kadar tahmini. *International Journal of Engineering Research and Development*, 3(2), 37-41.
- Yüzük, F. (2019). *Çoklu regresyon analizi ve yapay sinir ağları ile Türkiye enerji talep tahmini*. (Master thesis), Sivas Cumhuriyet Üniversitesi, Sivas.

