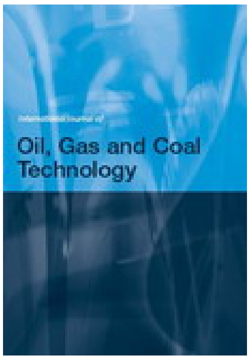


International Journal of Oil, Gas and Coal Technology

ISSN online: 1753-3317 - ISSN print: 1753-3309
https://www.inderscience.com/ijogct

Estimating coal consumption in Turkey using machine learning methods

Hande Erdoğan, Mehmet Kayakuş, Mustafa Terzioğlu

DOI: 10.1504/IJOGCT.2022.10052695

Article History:

Received:	18 April 2022
Accepted:	06 June 2022
Published online:	18 April 2023

Estimating coal consumption in Turkey using machine learning methods

Hande Erdoğan and Mehmet Kayakuş*

Department of Management Information Systems,
Akdeniz University,
Antalya, Turkey
Email: handeerdogan@akdeniz.edu.tr
Email: mehmetkayakus@akdeniz.edu.tr
*Corresponding author

Mustafa Terzioğlu

Accounting and Tax Department,
Akdeniz University,
Antalya, Turkey
Email: mterzioglu@akdeniz.edu.tr

Abstract: In this study, coal consumption was estimated by using machine learning techniques. In the model in which variables related to coal consumption were used, the R^2 value was 0.811 for the artificial neural network and 0.853 for the support vector regression, and it was observed that the model made successful predictions. The answers to important points such as the level of coal consumption in the future, the impact of the consumption on the climate, and the size of investment in the clean energy resources required for the energy needed if coal consumption is abandoned will be considered in this study to guide the researchers and decision-makers. [Received: April 20, 2022; Accepted: June 6, 2022]

Keywords: coal; coal consumption; artificial intelligence; machine learning; estimation; Turkey.

Reference to this paper should be made as follows: Erdoğan, H., Kayakuş, M. and Terzioğlu, M. (2023) 'Estimating coal consumption in Turkey using machine learning methods', *Int. J. Oil, Gas and Coal Technology*, Vol. 33, No. 1, pp.20–36.

Biographical notes: Hande Erdoğan is an Assistant Professor in the Management Information Systems at the Akdeniz University. She holds a PhD in Department of Business Administration from the Akdeniz University at Antalya in Turkey. Her work focuses on decision support systems and operations research.

Mehmet Kayakuş is an Assistant Professor in the Management Information Systems Department, Akdeniz University. He holds a PhD in Computer Engineering from the Süleyman Demirel University at Isparta in Turkey. His current research is focused on decision support systems, artificial intelligence, machine learning, and computer science.

Mustafa Terzioğlu is a Lecturer in Accounting and Tax at the Akdeniz University. In 2000, he received her BS from the Uludağ University in Turkey. He received his Master's degree from the Adnan Menderes University in Turkey in 2004. His work focuses on finance and decision support systems.

1 Introduction

Although there have been alternative developments regarding energy resources, which are an important element of economic and social development globally in the 21st century, fossil fuels such as coal, natural gas and liquid fuel still continue to be the primary energy source in world energy production. Although the use of renewable energy sources continues to increase, it is thought that the ranking of fossil fuels, especially coal, as an important source in satisfying the need of the global energy will not change in the long term due to its high reliability, accessibility, economy, and ease of handling and storage (Benalcazar et al., 2017). However, it is also known that the long-term intensive use of this fossil fuel, which has an important role in the economic and social development of countries, is a serious concern in posing environmental problems in terms of greenhouse gas emissions (Wang et al., 2018). Despite environmental agreements such as the Kyoto protocol and the Paris Climate Agreement, countries continue to use coal due to factors such as the unsteady processes experienced by oil producing countries, the instability in the prices of energy resources especially oil, rising gas prices and the inability to supply energy resources sufficiently and punctually (Apergis and Payne, 2010b; Kumar and Jain, 2010).

The increase in population, the desire to have better living standards and the investments in industrialisation increases the energy consumption significantly as well as the economic growth throughout the world (Yu et al., 2012). The increase in the energy consumption of the countries raise the demand for energy and this situation directly affects the social, environmental, and economic development of the countries. Countries that must focus on environmental conservation strategies while developing energy strategies to meet their energy needs on the one hand, have to concentrate more on energy planning as their global energy demands increase (Wolde-Rufael, 2010). As a result of the rapid economic growth of the countries, the increase in the imbalance between coal supply and demand affects the confidence in the sector and economic development. On the other hand, unplanned consumption of coal and environmental problems hinder the economic development of countries (Duan and Luo, 2021).

Turkey which has become one of the largest energy markets in the world, where energy consumption is increasing rapidly due to its young and ever-increasing population and developing economy has a total coal reserve of 20.84 billion tons according to the statistics of British Petroleum-2021 (Canyurt and Öztürk, 2008). Besides, although Turkey is heavily dependent on imported oil and natural gas, lignite coal accounts for 43% of total primary energy production (Canyurt and Öztürk, 2008; Corporation, 2021; Genç Kavas, 2021). Assuming that the current production conditions will continue, it is predicted that coal reserves worldwide will be depleted in approximately 114 years. The fact that coal has a longer lifespan than other energy sources and that it is spread over a wide geography highlights why countries should concentrate on the effective management of coal (Genç Kavas, 2021).

All these facts indicate that there will be serious problems in meeting the coal demands due to the limited fossil fuel reserves in Turkey, where the energy sector is heavily dependent on oil and coal-based fossil resources (Canyurt and Öztürk, 2008; Say and Yücel, 2006). Accurate and rational estimation of coal consumption in Turkey, which always has a balance problem between energy production and consumption; has great importance both theoretically and economically in determining vital energy policies, planning the effective growth of the mining sector, and designing and implementing environmental policies for sustainable energy management (Canyurt and Öztürk, 2008; Duan and Luo, 2021; Kumar and Jain, 2010).

In this study, in behalf of to guide decision makers and researchers in the management of economic growth and energy policies, it is aimed to predict the coal consumption in order to plan Turkey's coal management strategies correctly using artificial neural networks (ANNs) and support vector regression which are known to give more accurate and rational results than statistical, linear and/or nonlinear methods and then it is intended for comparing the results of both methods and examining the findings.

The next sections of this study are planned as follows. In the literature section, some publications examining the subject of this study were included. Dataset and methods were explained in the methodology section and the analysis results of these methods were shared in the findings section. In the conclusion section, the results are discussed, and the importance of this research is emphasised.

2 Literature

In the world energy statistics report published by The British Petroleum Company PLC, it is stated that the acceleration has been started to decrease in fossil energy consumption, but especially coal consumption in developing countries, continues to grow (Council, 2020; Duan and Luo, 2021). In emerging economies such as Turkey, China, and India, while considering the relationship between energy consumption and economic development, countries have to estimate their energy consumption with the pressure to carry out cost-effective and sustainable activities in realistic approach (Bloch et al., 2012). In this section of the study, some of the studies carried out to estimate coal consumption on a national or regional basis are included. Some of studies are given below.

Duan and Luo (2021) used the multivariate verhulst grey model (MVGM (1, N)) to examine the effects of population and GDP on coal consumption in China, and the MVGM (1, 3) model to test the validity of the model. As a result of their studies, they extended the univariate Verhulst model to multivariate MVGM. In their study, they underlined that with the development of energy management and emission control, they aim to apply MVGM (1, N) to estimate coal use in the whole country and to assist policy makers in making decisions (Bloch et al., 2012). Adebayo et al. (2021) investigated South Africa's coal consumption and environmental sustainability. In their study, they examined the role of financial development and globalisation using carbon emissions, coal consumption, financial development, globalisation, and economic growth data between 1980 and 2017. In addition to the auto-regressive distributed lag model (ARDL) approach, the authors used the Bayer and Hank combined co-integration, fully modified ordinary least squares (FMOLS) and dynamic ordinary least squares (DOLS) method and they compared the methods. Stating that the results of the FMOLS and DOLS approaches provide supporting evidence for the long-term results of ARDL, the authors suggested to decision-makers that South Africa should adopt policies that encourage energy consumers to prefer renewable energy. In their study, which aimed to predict coal consumption in Iran using socio-economic variables (Adebayo et al., 2021). Shakibaei et al. (2021) proposed a new hybrid method (BGWAN) by utilising the bat optimisation algorithm and grey wolf optimisation algorithm. Using the data from 1981–2019 to model and test the method, the authors stated that the performance of the BGWAN method was very high and could be used as a useful method to monitor coal consumption or energy demand in Iran (Shakibaei et al., 2021).

According to Jia et al. (2020) created a GM (1, 1) prediction model by examining the 20-year coal consumption data for the period 1999–2018 in the Gansu region of China and tested the accuracy of the model. Using the Markov chain estimation method to increase the reliability of the results, the authors also examined the accuracy of this model which they developed and observed that this new method significantly increased the accuracy of the model. Finally, the authors used the grey-Markov chain model to predict 2020–2035 Gansu coal consumption. Using the grey-Markov chain model and scenario analysis method to reduce coal consumption, the authors compared the results of the two methods. Revealing that the coal industry in Gansu was in line with the long-term development trend, the authors argued that the model with combined techniques has better predictive effect and many practical applications (Jia et al., 2020).

Including Turkey in the BRICS countries, Karakurt et al. (2020) used the economic and demographic data of all these countries in their studies to predict coal consumption with regression analysis. In their study, they determined gross domestic product, total population, and urban population as independent variables and coal consumption as a dependent variable. They achieved statistically high predictive power in their model. In their study, they also stated that the most important independent variable for Turkey, the country that is the subject of our paper, is the urban population size (Karakurt et al., 2020).

Chai et al. (2019) aimed to analyse the trend of coal consumption by examining the coal consumption and the Chinese economy in their study. In the first phase of their study, they used the annual coal consumption, GDP, and population data of China for the period 1965–2016 and performed a comparative analysis. Examining the relationship between China's coal consumption and economic growth, they found that the relationship was consecutive inverted U-shapes and reached inflection points in 1980, 1983 and 2013. With this analysis, they indicated that changes in four factors such as income, industrial structure, energy structure and energy efficiency have a significant impact on changes in coal consumption levels. However, in order to determine the specific effect value, they performed breakpoint regression analysis with 2001–2016 data. They used GDP per capita, the ratio of tertiary industry in GDP, the ratio of coal consumption in total energy consumption, and energy consumption per unit industry value added as independent variables, and per capita coal consumption data as dependent variable. Then, using the Logarithmic Mean Divisia Index (LMDI) method, they searched the formation mechanism of the Chinese coal consumption law. Finally, through partial least squares regression (PLSR), they examined the effect value of each factor and estimated future coal consumption to determine whether the third inverted U could be sustained (Chai et al., 2019). Analysing the short-term (2016–2020) and long-term (2020–2030) effects in the Shandong region, Ma et al. (2019) estimated coal consumption in their study using the ARIMA model and metabolism GM (1, 1) model. Determining that both models have high predictive power, the authors emphasised that coal consumption in the Shandong region will continue to increase as a result of their studies, and accordingly, energy policies should take this fact into account (Ma et al., 2019). Estimating India's coal consumption between 1995–2017 using metabolic grey model and the backpropagation (BP) network methods and two combined models derived from these methods, Li et al. (2019) estimated that coal consumption would increase by an average of 2.5% annually for the period 2018–2030 (Li et al., 2019).

In their study, which aims to forecast coal demand between 2017 and 2030 in order to balance energy supply and demand, reduce carbon emissions and support sustainable development in South Africa, Ma et al. (2018) applied linear, nonlinear and combined models, and stated that all three models had high reliability in predicting long-term coal consumption as a result of their analysis, and the trend of coal consumption would decrease in the relevant forecast period, and thus South Africa's dependence on coal would decrease (Ma et al., 2018).

Benalcazar et al. (2017) used a multilayer neural network model to estimate global coal consumption between 2020–2030. The authors stated that the findings of the model were in line with the US Energy Information Administration and BP publications, so its predictive ability was high (Benalcazar et al., 2017). Li et al. (2015) estimated the coal consumption of China between 1987 and 2012 with five different models, using the data of GDP, coal price, industrial structure, total population, energy structure, coal production and urbanisation rate in their study. Comparing the results of different models with the actual coal consumption, the authors stated that the particle swarm optimisation demand estimation model (PSO-DEM) method gave the nearest results with the least estimation error (Li et al., 2015). Zhang and Han (2013), in their study aiming to predict coal consumption using ANN model by using the data of 2000–2011 years, firstly trained the ANN model with 2000–2010 data and made the estimation of year 2011 with the model, and then realised the estimation of coal consumption for the last 12 years (Zhang and Han, 2013).

Using China's energy consumption statistics from 1998 to 2006, Feng et al. (2012) estimated total energy, coal energy and clean energy consumption for the years 2007–2012 using the grey model (1, 1) method. When they compare the results with the data of 2006, the authors determined that the annual averages of total energy, coal energy and clean energy consumption covering the forecast period will increase. The authors also stated that China's energy needs were constantly increasing, and with the increase in coal energy consumption, clean energy consumption would also increase (Feng et al., 2012). Kumar and Jain (2010) applied three different time series models used to separately estimate India's consumption of crude oil, coal, and natural gas. The results obtained by the authors with the models they used were compared with the projection of the Planning Commission of India and the results showed that the predictive power of the time series models was very high (Kumar and Jain, 2010). Canyurt and Öztürk (2008), aimed to estimate Turkey's fossil fuel demand and consumption of coal, oil and natural gas energy resources using population, gross national product, import and export data by genetic algorithm estimation models. In their study, they stated that Turkey's fossil fuel demand has increased significantly; and determined that coal, oil, and natural gas consumption values may increase approximately 2.82, 1.73 and 4.83 times between 2000 and 2020. The authors also compared the values obtained by the method they applied with the estimates of the World Energy Council Turkish National Committee (WECTNC) (Canyurt and Öztürk, 2008).

3 Materials and methods

3.1 Dataset

The data used in the study, including the years 1970–2019, were gathered from the World Bank. In this study with 50 data in total, the input variables used in the model generated to estimate the coal consumption are given below:

- average temperature
- population
- GDP per capita (current US$)
- industry (including construction), value added (% annual growth)
- electrical energy consumption (kWh per capita)
- consumer price index (% annual).

While setting the model of this study, the relevant variables were obtained by considering the previous research in the literature and Turkey's economic and demographic structure. The leading energy source used especially in industry in Turkey is electrical energy produced from lignite coal. For this reason, both GDP per capita and industry (including construction) value added data are preferred in the study. In the regions where the winter season is harsh, the electricity used from lignite coal is consumed for heating purposes in the households with medium and low-income levels at a high rate, and for air conditioning in the cities. It is thought that the price of the coal consumed by the households affects the amount of use. For this reason, mean temperature, consumer price index and population are also added to the model as variables. Thus, a model was setup that considers both industrial and household consumption.

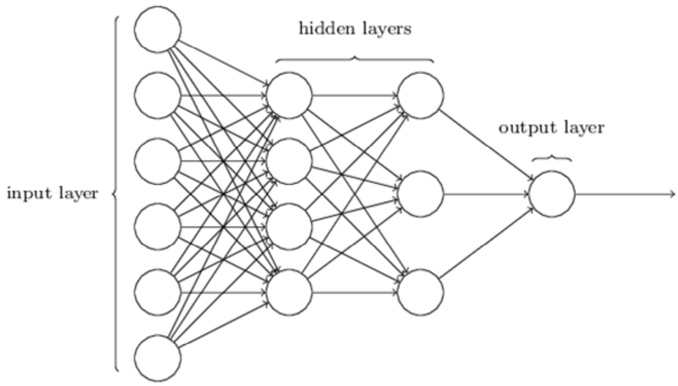
3.2 Methods

3.2.1 Artificial neural networks

ANN is an artificial intelligence learning technique developed to imitate the working mechanism of the human brain and to perform basic functions of the brain such as learning, remembering, generalising and deriving new information (Haykin, 2010). A basic ANN model consists of input, hidden and output layers. Figure 1 shows an ANN structure.

The input layer is the information that is given to the network as input and influences the value to be learned. The output layer is the information desired to be learned, predicted, or classified in the network. The hidden layer is the layer between the input layer and the output layer, where learning operations are performed. The number of layers and neurons varies according to the nature of the problem. A large number of layers will prolong the learning process and increase the computational complexity (Luo et al., 2018).

Figure 1 ANN structure



ANN architecture has two architectures, feed forward and back-propagation. In a feedforward architecture, there is a one-direction flow of information. The information received from the input layer is transmitted to the hidden layer. The output value is transferred to the output layer by processing the information in the hidden layers. The back-propagation ANN architecture, on the other hand, is the process of backward propagation of the error by looking at the error rate between the real system output of the values obtained in the output layer. Thus, the inputs are transmitted both in the forward and backward direction (Principe et al., 1999).

Figure 2 ANN working principle

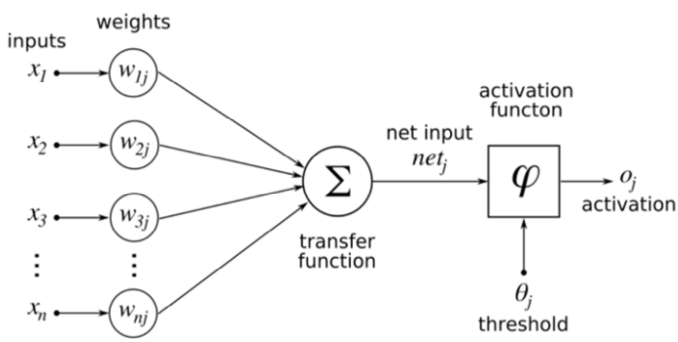


Figure 2 shows the working principle of ANN. The parameters used to determine the effect of the input values on the output value are called weights. Learning process in ANN is possible by finding the best values of the weights. The process of getting the optimal values of the weights is the training phase of the network. The weights are multiplied by the input values and proceed to the output. In order to reduce the error value to the minimum value, the weight values are recalculated as the error propagates backwards. The transfer function is a function that calculates the net input of an artificial neuron cell by adding the inputs multiplied by the weights. The activation function processes the net value coming to the cell and produces the activation output of the neuron in response to this input. The activation function has an impact on the accomplishment and performance of the network. It is important to choose an activation function while backward propagation to compute derivatives at each layer. The most used activation functions are sigmoid activation function and rectified linear unit (RELU) activation function (Yao et al., 2005).

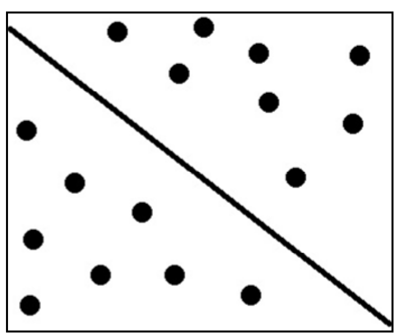
3.2.2 Support vector regression

Support vector machines (SVM) is a supervised learning method used in classification and regression analysis proposed by Vapnik in 1992 (Vapnik, 1995). Compared to other traditional learning methods, this method has much better performance and ability to solve nonlinear problems. An adaptation of SVM for regression, which is commonly used for classification problems, was proposed by Smola and Schölkopf (2004).

The purpose of regression analysis is to specify a mathematical function to accurately predict desired outcomes (target). Regression problems can be classified as linear and nonlinear problems. Since nonlinear regression problem is more difficult to solve, SVR was mainly developed for solving nonlinear regression problems (Karal, 2018). Trying to minimise the error between the actual and predicted value, the SVR tries to fit the best line within a threshold value which is the distance between the hyperplane and the boundary line (Tan et al., 2015).

For linear regression analysis, assume that the data consisting of N elements is $\theta = \{x_i, y_i\}$, $i = 1, 2, \ldots, N$ where $y_i \in \{-1, 1\}$ is the lanel values and $x_i \in R_d$ is the feature vector. As seen in Figure 3, in case of linear separation, these two-valued data can be separated directly by an extreme plane. The purpose of SVRs is to ensure that this extreme plane is equidistant from the sample group in two separate classes (Aci and Avci, 2017).

Figure 3 Linear SVR



For nonlinear regression analysis, consider a dataset of training points $\{(x_1, y_1), \ldots, (x_l, y_l)\}$ with feature vector $x_i \in R_d$ and target output $y_i \in \{-1, 1\}$.The nonlinear SVR is shown in Figure 4.

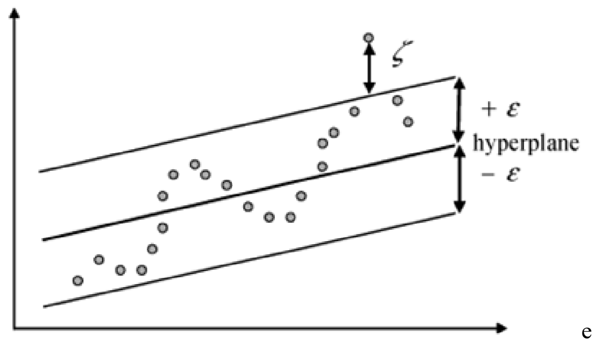
As shown in equation (1), the nonlinear relationship between input and output data is formulated with a linear function.

$$f(x) = \left(w^T \cdot \Phi(x)\right) + b \tag{1}$$

$f(x)$ are the predicted values; Φ, nonlinear mapping function; and w ($w \in R_n$) and b ($b \in R$) are adjustable coefficients. DVR standard form is defined as $C > 0$ and $\varepsilon > 0$ as in equation (2).

$$Min_{w,b,\xi,\xi^*} \frac{1}{2} w^T w + C \sum_{i=1}^{\ell} \xi_i + \xi_i^* \tag{2}$$

Figure 4 Nonlinear SVR



Source: Wu et al. (2004)

Constraints

$$\begin{cases} w^T \Phi(x_i) + b - y_i \le \varepsilon + \xi_i \\ y_i = w^T \Phi(x_i) - b \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0; i = 1, 2, \ldots, \ell \end{cases} \tag{3}$$

ξ_i^*, shows the training errors above ε, and ξ_i shows the training errors below ε. After solving the above quadratic optimisation problem with inequality constraints, the parameter vector w is calculated using equation (4).

$$w = \sum_{i=1}^{\ell} (\lambda_i^* - \lambda_i) \Phi(x_i) \tag{4}$$

λ_i^* and λ_i are Lagrange multipliers. Thus, the DVR formula is obtained as in the equation (5).

$$f(x) = \sum_{i=1}^{\ell} (\lambda_i^* - \lambda_i) K(x_i, x_j) + b \tag{5}$$

For an SVR, choosing a suitable kernel is imperative for the success of the learning process. SVR cores include linear core, polynomial nucleus, index core, dual neural network core, and radial basis function (RBF) cores. In these cores, the RBF kernel is the most popular core function and has better stable overall (Hsu and Lin, 2002).

RBF kernel:

$$K(x_i, x_j) = \exp\left(-\gamma \left|(x_i \cdot x_j)\right|^2\right) \tag{6}$$

For comparing the performance of ANNs, support vector regression and multiple linear regression models, adjusted coefficient of determination (R^2), mean absolute error (MAE), mean squared error (MSE), root mean square error (RMSE), and mean absolute percentage error (MAPE) criteria are used. Equations for these criteria are given below.

$$R^2 = 1 - \frac{\sum \left(y_i - y_i^* \right)^2}{\sum \left(y_i - y_{ave} \right)^2} \tag{7}$$

$$MAE = \frac{\sum_{i=1}^{n} \left| y_i - y_i^* \right|}{n} \tag{8}$$

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - x_i \right)^2 \tag{9}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_i - x_i \right)^2} \tag{10}$$

$$MAPE = \frac{\sum_{i=1}^{n} \left| \frac{y_i - y_i^*}{y_i} \right|}{n} \cdot 100 \tag{11}$$

where y_i is the observation value and y_i^* is the predicted value, y_{ave} average of observation values. According to these criteria, high R^2 and low MAE, MSE, RMSE and MAPE values determine the model that fits well.

4 Results and discussion

In this study, ANNs and support vector regression models were used to calculate the annual coal consumption in Turkey. In both models, the normalisation process was fulfilled using the decimal scaling method to be able to operate on the data and for ease of calculation. The decimal point of values of the attribute has been moved in this method. This movement of decimal points is completely dependent on the maximum value among all values in the attribute. The decimal scale normalisation formula is, where, v_i is the scaled values, v is the range of values, j is the smallest integer $Max(|v_i|) < 1$ (Saranya and Manikandan, 2013)

$$v^i = \frac{v}{10^j} \tag{12}$$

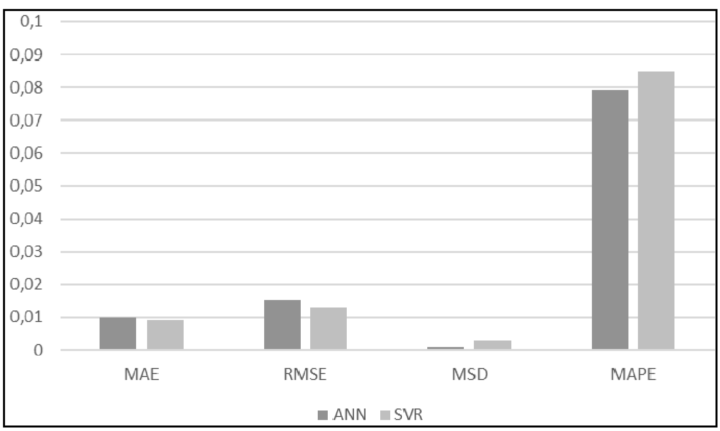
Since the normalised data is between 0 and 1, it is not possible to compare the results with the actual values. For this reason, denormalisation process was applied to compare the values obtained in the models with the real values. Thus, the actual values and the estimated values can be compared. As a result of the tests conducted in the study, it was decided to use 70% of the data for training and 30% for testing purposes. In order to work on the same data in comparison of methods, linear sampling method was preferred in data selection.

A feedback ANN model was created for the ANN method. In this model, it consists of three hidden layers and three neurons in each hidden layer. 1,000 iterations were carried out to get the best result in the model. The nonlinear method was used in the SVR method. RBF was preferred as the core function since it gave the most successful value as a result of the tests. Combinations were tried C and ε which are SVR parameters, and finally good performance was obtained with ε: 0.3 and C: 1,000 and it was decided to use these values. Comparison of the statistical values of the models is given in Table 1.

Table 1 Comparison of statistical values

	ANN	*SVR*
R^2	0.811	0.853
MAE	0.010	0.009
MSE	0	0
RMSE	0	0
MSD	0.001	–0.003
MAPE	0.079	0.085

Figure 5 Graphical comparison of statistical values of models



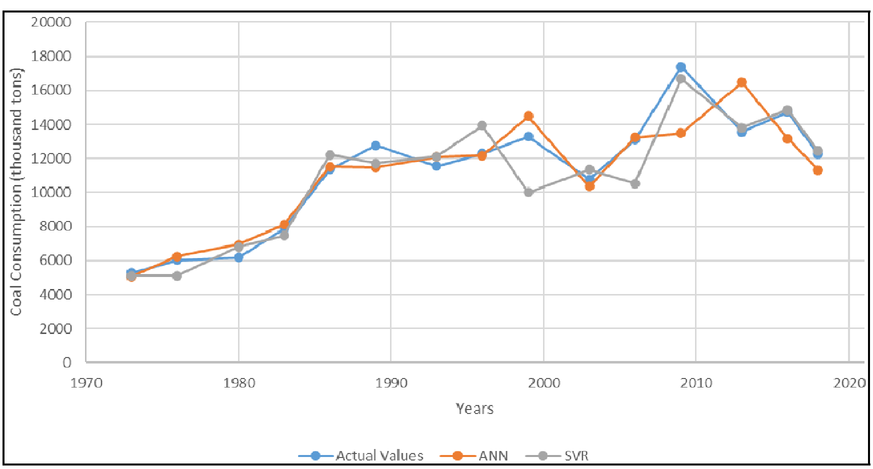
Comparison of the statistical values of the models is given in Figure 5. In the model, the estimation error of the test data according to the MAPE error criterion was 4.9% for ANN and 7.4% for SVR. Estimation models with MAPE values below 10% are considered models with ‘high accuracy’ or ‘very good’ accuracy (Lewis, 1982; Witt and Witt, 1992). Since the method with the lowest value will be considered more successful, the ANN model was deemed to be more successful in the study. MSE and RMSE values close to zero demonstrate that the model is more successful and performant (Singh et al., 2009). When the MSE and RMSE values of the models were examined, ANN was found to be more successful. MAE is a value that measures the average size of errors in a set of predictions (Willmott and Matsuura, 2005). The fact that this value is close to zero indicates that the estimation is realised with less error. ANN estimates with fewer errors than the SVR than Table 1, as its proximity to zero indicates a less erroneous estimate.

The comparison of the actual values of the models with the estimated values for the selected years is given in Table 2.

Table 2 Comparison of actual values and estimated values (thousand tons)

Year	*Actual values*	*ANN*	*SVR*
1973	5,282	5,058	5,106
1976	6,020	6,249	5,106
1980	6,173	6,958	6,805
1983	7,860	8,122	7,467
1986	11,361	11,518	12,209
1989	12,770	11,482	11,723
1993	11,549	12,106	12,118
1996	12,296	12,173	13,945
1999	13,284	14,480	10,004
2003	10,777	10,353	11,360
2006	13,088	13,231	10,532
2009	17,402	13,487	16,674
2013	13,553	16,478	13,813
2016	14,735	13,172	14,848
2018	*12,254*	*11,313*	*12,450*
Average	*11,227*	*11,079*	*10,944*

Figure 6 Actual and estimated values of methods (see online version for colours)



When the values in Table 2 are examined, it is seen that the difference between the actual values and the estimated values of the models is 148.2 thousand tons for ANN and 282.9 thousand tons for SVR. The error rate between actual values and estimated values is 1.32% for ANN and 2.52% for SVR. It is seen that these error rates are among the acceptable tolerance values. Figure 6 shows the comparison of the actual data with the estimated values.

As seen in Figure 6, it is seen that the actual values and the estimated values are very close to each other and both methods can be used for estimation. However, the results of SVR have demonstrated more accurate prediction performance.

5 Conclusions

It is an undeniable fact that the continuous use of coal, which has been a low-cost and easily obtainable energy source since the industrial revolution, is a threat to the world, considering the damage it causes to the environment. Today, when countries and organisations make decisions to reduce the use of this energy source, coal continues to be used as a mostly consumed energy source. At the Conference of the Parties (COP26) summit, countries that are most dependent on coal such as China, USA, Australia, and India, including Turkey, where this research was conducted, have not made any commitment to consume coal. Nearly forty countries that signed the agreement on quitting coal at COP26 also committed to quit coal only in electricity generation. The most important reason for this is the prediction that the welfare of the countries will decrease due to the high costs of transition from coal to renewable energy, as well as the loss of time that will occur while making this transition. While the consumption of coal is increasing, the reasons such as the decrease in reserves, inability to replace alternatives at a speed to compensate for this loss, lack of confidence in supplying alternative energy sources, large deviations in world energy prices, easy availability, and the relative abundance of coal, have being directed countries to use coal. And all of these allow the future of coal to be planned accurately and realistically, the countries to manage both social, environmental, and economic development effectively only with the right planning, and because of all this, coal gain strategic importance on a global scale (Apergis and Payne, 2010a; Ataş and Güler, 2020; Güllü and Kartal, 2021; Tamzok, 2012). Therefore, in the short and medium term, the consumption of coal among alternative energy sources, as it was in the past, is expected to show an ever-increasing acceleration considering the growing economies and population growth rates.

The latest 11th Development Plan, covering the period of 2019–2023, which has been offering plans on economic development and progress in Turkey since 1963 on issues such as economy, health, education, energy, has focused on goals especially increasing the quality of the population, emphasising information technologies-based energy management on behalf of economic growth, determining energy policies to manage the balance of power globally, concentrating on energy efficiency and preventing air pollution (The Presidency of the Republic of Turkey, 2019). It is known that the coal industry, which is one of the most important touchstones of national economies, can contribute to social and economic development with the right energy planning strategies. Based on this fact, also in our country, exact and accurate estimation of coal consumption is very important in order to effectively manage the development of the coal industry, to determine energy management control policies and to implement them (Jia et al., 2020).

For this purpose, current techniques such as ANNs and support vector regression methods are used to predict coal consumption in Turkey in this study. Estimations were realised using two methods with the data obtained from the World Bank for the years 1970–2019, the methods were compared, and it was seen that the support vector regression method (R^2 = 85.3%) gave better results with the estimation performance.

With this study, it has been emphasised that it is very important for policy makers in a country to correctly estimate, plan, manage and control energy accurately and has been underlined that the incomplete realisation of even one of these may lead to consequences such as insufficient energy supply in that country, increase in the costs of existing energy resources, inability to provide energy in a reliable and timely manner, and foreign dependency. For countries to manage their coal consumption, it is recommended that policy makers plan the service-production sector balance, make cross-sectoral optimisation and control the increase in investment. In addition, increasing the contribution of energy consumption to economic development and increasing energy efficiency should be among the suggestions. For this purpose, decision makers should support technology innovation for coal use, adjust the energy consumption structure, and increase coal use investments as well as investments and supports for natural gas and renewable energy resources (Chai et al., 2019). As the amount of energy use increases, countries should plan renewable energy sources, besides fossil fuels and solar, geothermal, wind, etc. It is also recommended to develop strategies to increase alternative energy sources. As the amount of energy use increases, it is recommended those countries should plan renewable energy sources besides fossil fuels, and also they should develop strategies to increase alternative energy sources like solar, geothermal, wind, etc. (Ataş and Güler, 2020). Considering the fact that domestic energy sources should be preferred in order to avoid foreign dependency in energy, the strategic importance of coal in the energy policies of the country emerges. It will be possible to use coal efficiently and cleanly by increasing research and development activities in order to increase the reserves of coal, to operate it effectively and to expand it. Effective and scientific planning and management of energy management will not only reduce foreign dependency, but also provide and support the country's economic development cost-effective (TMMOB, 2020). In addition, it is thought that decision makers will be able to accurately estimate the economic cost of giving up coal in the medium and long-term economic planning of countries in the future, based on this model. At the same time, it is thought that the model is an appropriate model in revealing the essential investment criteria for alternative renewable energy sources to coal, and with high predictive results machine learning techniques like ANNs and support vector regression will be appropriate.

References

Aci, M. and Avci, M. (2017) 'Reducing simulation duration of carbon nanotube using support vector regression method', *Journal of the Faculty of Engineering and Architecture of Gazi University*, Vol. 32, No. 3, pp.901–907.

Adebayo, T.S., Kirikkaleli, D., Adeshola, I., Oluwajana, D., Akinsola, G.D. and Osemeahon, O.S. (2021) 'Coal consumption and environmental sustainability in South Africa: the role of financial development and globalization', *International Journal of Renewable Energy Development*, Vol. 10, No. 3, pp.527–536.

Apergis, N. and Payne, J.E. (2010a) 'The causal dynamics between coal consumption and growth: evidence from emerging market economies', *Applied Energy*, Vol. 87, No. 6, pp.1972–1977.

Apergis, N. and Payne, J.E. (2010b) 'Coal consumption and economic growth: evidence from a panel of OECD countries', *Energy Policy*, Vol. 38, No. 3, pp.1353–1359.

Ataş, H. and Güler, H. (2020) 'The effect of natural gas, oil, and coal consumption of Turkey on growth: an econometric investigation', *Cukurova University Institute of Social Sciences Journal*, Vol. 29, No. 3, pp.524–539.

Benalcazar, P., Krawczyk, M. and Kamiński, J. (2017) 'Forecasting global coal consumption: an artificial neural network approach', *Gospodarka Surowcami Mineralnymi*, Vol. 33, No. 4, pp.29–44.

Bloch, H., Rafiq, S. and Salim, R. (2012) 'Coal consumption, CO_2 emission and economic growth in China: empirical evidence and policy responses', *Energy Economics*, Vol. 34, No. 2, pp.518–528.

Canyurt, O.E. and Öztürk, H.K. (2008) 'Application of genetic algorithm (GA) technique on demand estimation of fossil fuels in Turkey', *Energy Policy*, Vol. 36, No. 7, pp.2562–2569.

Chai, J., Du, M., Liang, T., Sun, X.C., Yu, J. and Zhang, Z.G. (2019) 'Coal consumption in China: how to bend down the curve?', *Energy Economics*, Vol. 80, No. C, pp.38–47.

Corporation, G.D.o.T.C.E. (2021) *Annual Report* [online] https://www.tki.gov.tr/yayinlar (accessed 23 November 2021).

Council, W.E. (2020) *BP Energy Outlook 2020 Report* [online] https://www.dunyaenerji.org.tr/wp-content/uploads/2020/09/BP-Enerji-Gorunumu-2020-Raporu-Ozeti.pdf (accessed 18 March 2022).

Duan, H. and Luo, X. (2021) 'A novel multivariable grey prediction model and its application in forecasting coal consumption', *ISA Transactions*, Vol. 120, No. 1, pp.110–127.

Feng, S., Ma, Y., Song, Z. and Ying, J. (2012) 'Forecasting the energy consumption of China by the grey prediction model', *Energy Sources, Part B: Economics, Planning, and Policy*, Vol. 7, No. 4, pp.376–389.

Genç Kavas, H. (2021) *Future Forecast of Turkey Demand with World Energy Outlook and Artificial Neural Networks*, Iksad Publications, Ankara, Turkey.

Güllü, M. and Kartal, Z. (2021) 'Forecasting of Turkey's renewable energy sources up to year 2030', *Journal of Social Sciences*, 19 May, Vol. 2, No. 2, pp.288–313.

Haykin, S. (2010) *Neural Networks: A Comprehensive Foundation*, pp.1–24, McMillan, New Jersey.

Hsu, C-W. and Lin, C-J. (2002) 'A comparison of methods for multiclass support vector machines', *IEEE Transactions on Neural Networks*, Vol. 13, No. 2, pp.415–425.

Jia, Z-q., Zhou, Z-f., Zhang, H-j., Li, B. and Zhang, Y-x. (2020) 'Forecast of coal consumption in Gansu Province based on Grey-Markov chain model', *Energy*, Vol. 199, No. C, p.117444.

Karakurt, İ., Aydın, G. and Amiri, M.R. (2020) 'Prediction of coal consumption of the BRICS-T countries by multiple regression analysis', *Harran Üniversitesi Mühendislik Dergisi*, Vol. 5, No. 1, pp.32–45.

Karal, Ö. (2018) 'Compression of ECG data by support vector regression method', *J. Fac. Eng. Arch.*, Vol. 1, No. 2, pp.743–756, Gazi Univ.

Kumar, U. and Jain, V.K. (2010) 'Time series models (grey-Markov, grey model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India', *Energy*, Vol. 35, No. 4, pp.1709–1716.

Lewis, C.D. (1982) *Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting*, Butterworth-Heinemann, Oxford, UK.

Li, B-B., Liang, Q-M. and Wang, J-C. (2015) 'A comparative study on prediction methods for China's medium-and long-term coal demand', *Energy*, Vol. 93, No. 2, pp.1671–1683.

Li, S., Yang, X. and Li, R. (2019) 'Forecasting coal consumption in India by 2030: using linear modified linear (MGM-ARIMA) and linear modified nonlinear (BP-ARIMA) combined models', *Sustainability*, Vol. 11, No. 3, p.695.

Luo, K., Xing, J., Bai, Y. and Fan, J. (2018) 'Prediction of product distributions in coal devolatilization by an artificial neural network model', *Combustion and Flame*, Vol. 193, No.1, pp.283–294.

Ma, M., Su, M., Li, S., Jiang, F. and Li, R. (2018) 'Predicting coal consumption in South Africa based on linear (metabolic grey model), nonlinear (non-linear grey model), and combined (metabolic grey model-autoregressive integrated moving average model) models', *Sustainability*, Vol. 10, No. 7, p.2552.

Ma, M., Wu, H., Wang, H. and Zhang, Y. (2019) 'Short-term and long-term prediction of coal consumption in Shandong Province – with ARIMA model and metabolism GM (1, 1) model', *IOP Conference Series: Earth and Environmental Science*.

Principe, J.C., Euliano, N.R. and Lefebvre, W.C. (1999) *Neural and Adaptive Systems: Fundamentals through Simulations with CD-ROM*, John Wiley & Sons, Inc., New York City, USA.

Saranya, C. and Manikandan, G. (2013) 'A study on normalization techniques for privacy preserving data mining', *International Journal of Engineering and Technology*, Vol. 5, No. 3, pp.2701–2704.

Say, N.P. and Yücel, M. (2006) 'Energy consumption and CO_2 emissions in Turkey: empirical analysis and future projection based on an economic growth', *Energy Policy*, Vol. 34, No. 18, pp.3870–3876.

Shakibaei, A., Ghasemi Nejad, A., Jalaee, S.A. and Derakhshani, R. (2021) 'A novel computational intelligence approach for coal consumption forecasting in Iran', *Sustainability*, Vol. 13, No. 14, p.7612.

Singh, K.P., Basant, A., Malik, A. and Jain, G. (2009) 'Artificial neural network modeling of the river water quality – a case study', *Ecological Modelling*, Vol. 220, No. 6, pp.888–895.

Smola, A.J. and Schölkopf, B. (2004) 'A tutorial on support vector regression', *Statistics and Computing*, Vol. 14, No. 3, pp.199–222.

Tamzok, N. (2012) 'The future of coal from the perspective of geopolitical and technological developments', *TMMOB*, Vol. 8, No. 1, pp.247–291.

Tan, P., Zhang, C., Xia, J., Fang, Q-Y. and Chen, G. (2015) 'Estimation of higher heating value of coal based on proximate analysis using support vector regression', *Fuel Processing Technology*, Vol. 138, No. C, pp.298–304.

The Presidency of the Republic of Turkey, D.o.S.a.B. (2019) *Eleventh Development Plan (2019–2023)* [online] https://www.sbb.gov.tr/wp-content/uploads/2019/07/OnbirinciKalkinmaPlani.pdf (accessed 5 March 2022).

TMMOB (2020) *Chamber of Mining Engineers Coal and Energy Report*.

Vapnik, V.N. (1995) *The Nature of Statistical Learning Theory*, Springer-Verlag, NY; Berlin/Heidelberg, Germany.

Wang, C., Li, B-B., Liang, Q-M. and Wang, J-C. (2018) 'Has China's coal consumption already peaked? A demand-side analysis based on hybrid prediction models', *Energy*, Vol. 162, No. C, pp.272–281.

Willmott, C.J. and Matsuura, K. (2005) 'Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance', *Climate Research*, Vol. 30, No. 1, pp.79–82.

Witt, S.F. and Witt, C.A. (1992) *Modeling and Forecasting Demand in Tourism*, Academic Press Ltd., Cambridge, Massachusetts, USA.

Wolde-Rufael, Y. (2010) 'Coal consumption and economic growth revisited', *Applied Energy*, Vol. 87, No. 1, pp.160–167.

Wu, C-H., Ho, J-M. and Lee, D-T. (2004) 'Travel-time prediction with support vector regression', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 5, No. 4, pp.276–281.

Yao, H., Vuthaluru, H., Tade, M. and Djukanovic, D. (2005) 'Artificial neural network-based prediction of hydrogen content of coal in power station boilers', *Fuel*, Vol. 84, Nos. 12–13, pp.1535–1542.

Yu, S., Wei, Y-M. and Wang, K. (2012) 'China's primary energy demands in 2020: predictions from an MPSO–RBF estimation model', *Energy Conversion and Management*, Vol. 61, No. 1, pp.59–66.

Zhang, Y. and Han, Y. (2013) 'Prediction of the coal consumption based on artificial neural network', *Computer Science and Application*, Vol. 3, No. 1, pp.278–281.