Energy 224 (2021) 120090

Contents lists available at ScienceDirect

Energy

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

Prediction of transportation energy demand: Multivariate Adaptive Regression Splines

Mohammad Ali Sahraei [a, *], Hakan Duman [b], Muhammed Yasin Çodur [a], Ecevit Eyduran [c]

[a] Erzurum Technical University, Faculty of Engineering and Architecture, Civil Engineering Department, 25200, Erzurum, Turkey
[b] Iğdır University, Iğdır Vocational School, Department of Transportation Services, 76000, Iğdır, Turkey
[c] Department of Business Administration, Quantitative Methods, Faculty of Economics and Administrative Sciences, Iğdır University, Iğdır, Turkey

ARTICLE INFO

Article history:
Received 23 May 2020
Received in revised form
8 January 2021
Accepted 8 February 2021
Available online 17 February 2021

Keywords:
Transport energy demand
Multivariate Adaptive Regression Splines
Predictive model

ABSTRACT

Energy usage in the transportation sector has been increasing in Turkey. Good management of energy is important as well as a reliable prediction of the energy demand in the transportation sector. The main objective of this research is to predict transport energy demand using Multivariate Adaptive Regression Splines (MARS) as a nonparametric regression technique. Transport energy demand was modeled for the period 1975–2019 based on a mix of factors including the gross domestic product (GDP), population, vehicle-km, ton-km, passenger-km and oil price. Five models were established and compared with real data collected from the Ministry of Energy and Natural Resources (MENR). Five MARS models including pairs of predictors, i.e. oil price-GDP, oil price-population, oil price-ton, oil price-vehicle and oil price-passenger, were evaluated comparatively in the prediction of energy demand. Among the candidate models, the third MARS model, which had the lowest RMSE, SD ratio, AICc values and the highest R^2, Adjusted R^2 and especially GR^2 value, was selected as the best predictive model. In conclusion, it could be suggested that the third MARS model produced the highest predictive performance in the prediction of energy demand by two predictors, ton and oil price.

1. Introduction

Energy policy and planning have become essential elements in the public planning of all nations today. The significance of energy policy and planning is connected to environmental advantage, energy security and industrial competitiveness. Transportation in Turkey is still utilizing traditional fossil energy varieties, including diesel, gasoline and Liquefied Natural Gas (LNG). These actions produce millions of tons of greenhouse gases annually. Appropriate energy policy and planning in the transportation field may decrease the demand for fossil energy and consequently reduce the emission of greenhouse gases. Depending on fossil energy usage, the transportation sector accounts for nearly 49% of the national greenhouse gas emissions [1].

In the services and industry sectors in Turkey, a total of more than 100 million Tons of Oil Equivalent energy (TOE) were used in 2014 and around 86 million TOE was used in the industrial sector, whereas the services sector utilized more than 14.50 million TOE. Sectors with the highest shares in overall energy utilization were electricity distribution and generation, which accounted for around 42.4%, the share of the manufacturing industry was 38.4%, and the transportation sector utilized 9.5%. Looking at the annual consumption rate by fuel variety, natural gas was the highest utilized type of fuel in 2014. Lignite was used to be around 1.6 million TOE, though coal usage created almost 16 million TOE in total. When energy usage is analyzed as to sectors, the manufacturing industry was the sector with the highest electricity utilization with 78.30 million MWh in 2014. In the services and industrial sectors, 79.7% of the energy was applied for the generation of a provision of services and goods, and 13.2% of this quantity was used for plugging and lighting electrical equipment at workplaces [2,3].

Realistic prediction of energy requirement is possibly the initial and the most important stage for the government to establish a suitable strategic program, and to allocate adequate quantities of resources for numerous actions in order to meet energy requirements [4,5]. Determining the energy requirements in Turkey has been carried out through the Ministry of Energy and Natural

* Corresponding author.
E-mail addresses: ali.sahraei@erzurum.edu.tr (M.A. Sahraei), hknd1977@gmail.com (H. Duman), mycodur@erzurum.edu.tr (M.Y. Çodur), ecevit.eyduran@gmail.com (E. Eyduran).

https://doi.org/10.1016/j.energy.2021.120090

Resources (MENR). Energy requirement predictions are depending on several elements such as industrialization, development, conservation, technology and urbanization. The statistics are modified each year in light of the performance over the previous year [6–8].

The majority of predictive modeling techniques can be classified into two main groups of historical and causal data-based techniques. In the causal techniques, the energy demand modeling in the transportation sector is generally related to the GDP, population, annual vehicle-km, income, the number of vehicles, import, export, and so on. The magnitude of specific elements is occasionally challenging to evaluate, and the influence of these variables on energy demand modeling and prediction may not be similarly significant. Artificial neural networks (ANNs) [9–12], regression models [13–17], Fuzzy logic [18–20] and Genetic algorithm models [21–23] are the most frequent causal methods utilized to predict the energy requirement. In contrast, techniques depending on the historical data utilize the earlier values of a variable to forecast the future values of that variable. Time series [24–28], Grey prediction [29–32] and Autoregressive Integrated Moving Average (ARIMA) [33,34] are among these methods. Among the above methods, ARIMA's approach, AI methods including ANN, Support Vector Regression (SVR), genetic algorithms, fuzzy logic and Multivariate Adaptive Regression Splines (MARS) are the well-known forecasting methods utilized by energy researchers.

As a modified form of the Classification and Regression Tree (CART) algorithm, the MARS data mining algorithm offers extremely effective implementations in many application areas because of its flexible operation process. Some academic studies on the MARS algorithm have been performed and several types of research are presented to further support its functionality. Research conducted by Lewis and Stevens [35] described the functionality of MARS for modeling time series data where lagged values of the time series were treated as explanatory parameters. Recently, MARS predictive modeling has been implemented for modeling a range of data, for instance, speech modeling by Haas and Kubin [36], mobile radio channels forecasting by Ekman and Kubin [37], intrusion recognition in information systems security by Mukkamala and Sung [38] and to identify pesticide transport in soils by Yang et al. [39]. This method has also been utilized to anticipate the average monthly foreign exchange rates by Abraham [40], to model credit scoring by Lee and Chen [41], and for data mining on breast cancer patterns by Chou et al. [42]. In all of the above-mentioned studies, promising outcomes have been achieved.

In the case of energy modeling and prediction, MARS data mining algorithm was applied to a wide variety of data recently, for example, Electricity [43–45], water resource modeling [46–48], building energy consumption prediction [49–52], natural energy consumption forecasting [53–56] and renewable energy consumption [57]. More descriptions about energy modeling using the MARS approach will be provided in the following section.

The main objective of this research is to develop a suitable transport energy model that is capable of predicting transport energy demand in Turkey. Since finding appropriate variables that affect transport energy demand as well as their interrelationships are complicated, some pairs of the GDP, population, vehicle-km, ton-km, passenger-km, and oil price were considered as model inputs in the MARS approach by eliminating multicollinearity problem within the scope of improving MARS predictive performances. The input data for 1975–2019 are utilized for the development of the models in this research. The results of this research are to demonstrate that the suggested method produces the best predictive performance for predicting the future transportation energy demand of Turkey.

This paper is organized as follows. Section 2 provides a literature review of the existing research about transport energy demand modeling. In Section 3, data collection procedure and the methodology are presented based on the MARS software, model calibration, and validation. Model development and discussion are described in Section 4, followed by the conclusions in the section 5.

2. Literature review

Energy planning and demand in Turkey are performed by the MENR, as described before. The MENR utilizes a detailed model for the evaluation of energy demand that requires a large amount of information for the overall and sectoral energy usage estimations. In addition, many models related to energy have been created from many types of research utilizing several different techniques [58,59]. A summary of the existing literature concerning energy demand prediction models with several methods is provided below.

2.1. Prediction methods

Different methods for the prediction and estimation of energy demand are listed in the literature. In this regard, the most beneficial techniques for predicting energy demand (i.e. for commercial and renewable energy) are evaluated by Suganthi and Samuel [61], and are classified based on the utilized methods in each research. Ant colony and particle swarm optimization (PSO), as well as support vector regression method, are new methods that are adapted for energy demand prediction. In addition, methods including MARKAL (i.e. acronym for MARKet ALlocation) and the Long-range Energy Alternatives Planning system (LEAP) are also utilized at the regional and national levels for energy demand management. Accordingly, Suganthi and Samuel [61] discovered that the ARIMA method is connected with ANN as well as other soft computing methods to enhance the precision of energy demand prediction. Grey prediction is another method that is applied effectively for energy demand evaluation. Fuzzy logic, genetic algorithms, SVR, and PSO are promising methods in predicting commercial and renewable energy resources. It is envisioned that such methods will assist energy organizers in a precise plan for the future and use renewable and sustainable energy sources to a greater magnitude. In addition, Suganthi and Samuel [61] described that macroeconomic variables for energy modeling are essential for each country. Sophisticated modeling methods including genetic algorithms, grey prediction, fuzzy logic, PSO, SVR can be utilized by macro planning experts for the prediction of the precise energy demand.

Amasyali and El-Gohary [62] provided a review of the latest studies in the field of data-driven building power usage prediction. The scope of a set of models was examined regarding building varieties (non-residential and residential), prediction period (hourly, sub-hourly, monthly, daily, and also yearly), as well as varieties of power usage predicted (cooling, heating and lighting). The outcomes of this evaluation reveal several areas that may need considerably more research focus, such as lighting building power usage prediction, residential building power usage prediction, as well as long-term building power usage prediction. Reviewing the demand predicting methods for short-term electric loads was carried out by Abu-El-Magd and Sinha [63], where both online and offline loads were investigated. Ghods and Kalantar [64] reviewed the methods for predicting long-term electric loads including some conventional methods and Artificial Intelligence (AI) approaches. Other studies in this area are examined and reviewed by Mukherjee [65], Bajay [66], Swan & Ugursal [67], and Hyndman [68].

2.2. Prediction of transportation energy demand

Shunping et al. [69] reviewed the international and domestic researches on transportation energy usage, including the relationship among transport power usage as well as social economy, external cost determination of energy usage, and forecasting of the power need of the transportation industry.

Zhang and Yang [70] implemented a mixture of forecasting techniques into the prediction of China's transportation energy demand, focusing on enhancing forecasting precision. A correctly-weighted unique model was initially established, mixing such models as error modification, nonlinear regression, and multiple regression. The outcomes demonstrated that the mixture forecasting techniques can improve prediction precision.

Sonmez et al. [71] proposed three mathematical methods to compute the transport energy demand of Turkey utilizing the artificial bee colony algorithm. Within the construction of the models, total annual vehicle-km, population and GDP were employed as potential variables. For transport energy demand calculation, quadratic, exponential and linear types of mathematical expressions were utilized. The methods that were established were utilized in two feasible scenarios in order to predict the transport energy demand of Turkey for the period of 2014–2034. The outcomes acquired from circumstances pointed out that the power demand of Turkey will be double that of 2013 by 2034.

Canyurt et al. [72] established an approach utilizing the genetic algorithm method to calculate future transportation energy demand in Turkey. The genetic algorithm transportation energy demand (GATENDM) method is established depending on socio-economic indicators including export and import, GDP, population, and transport variables such as bus, car, and also truck sales. The experimental results of the proposed method show that it can be utilized as a substitute solution and an evaluation method in forecasting future transport energy usage values.

Grey forecasting method was implemented by Lu et al. [73] to determine the development trends of the CO_2 emissions, vehicular energy usage as well as quantity of automobiles in Taiwan for the period 2007–2025. The findings demonstrated that the CO_2 emissions, energy need and vehicle fleet by the road transport system will continue to increase by 3.23%, 3.25% and 3.64%, respectively, from 2007 until 2025. The related CO_2 emission will probably be between 61.1 and 73.4 million metric tons in the low- and high-scenario profiles.

Yang [74] evaluated the energy usage and need of various transport modes, such as water, road, railway, and air, and forecasted the energy usage by the mentioned modes for 2010–2020, by examining historical information including passenger turn volume, freight turnover volume, and energy usage of the mentioned transportation modes.

2.3. Multivariate adaptive regression splines method

As a nonparametric regression method, the MARS algorithm is created to utilize for cubic or linear basis functions. Typically, the method is a flexible and fast statistical tool operated via an integrated linear and non-linear modeling method [46,48]. Furthermore, it has the ability to employ a set of basic functions utilizing various predictor parameters. This is essential for predicting demand depending on interactions among distinct parameters and the demand information. Even though the use of MARS method in energy demand prediction is extremely rare, although the method has been verified to be remarkably precise in numerous engineering problems. In Ontario (Canada), the MARS method was utilized via a semi-parametric technique, for predicting short-term oil costs by Morana [54] and for analyzing the behavior of short-term (hourly) energy cost information via lagged input combinations by Zareipour et al. [55]. Sigauke and Chikobvu [43] evaluated the MARS method for predicting electricity demand in South Africa; this confirmed its ability to yield a considerably lower Root Mean Square Error (RMSE) compared with regression-based techniques. Nevertheless, despite its increasing global use, e.g. Abraham & Steinberg [75], Sharda et al. [48], Deo et al. [46] and Deo and Kisi [76], the use of MARS method for predicting electricity demand is still investigated.

Özmen et al. [47] utilized a lately developed technique known as Robust Conic MARS (RCMARS), depending on MARS and Conic MARS (CMARS), in order to predict precipitation amount due to its capability of modeling complicated unstable data in Turkey. In CMARS, which was formulated as an effective option for MARS, the model intricacy is penalized in the form of Tikhonov regularization and examined as conic quadratic programming. On the other hand, in RCMARS, Conic MARS is processed additionally by including the existence of uncertainty in future situations and robustifying it with a strong optimization approach. Moreover, the efficiency of the RCMARS precipitation method was also compared with that of MARS and CMARS. The outcomes pointed out that RCMARS develops much more stable and precise precipitation models in comparison with CMARS and MARS. Cheng and Cao [50] suggested utilizing an artificial intelligence (AI) and Evolutionary MARS (EMARS) method, to efficiently anticipate the energy functionality of buildings. EMARS is a hybrid of artificial bee colony and MARS. The suggested method utilized eight input variables as well as two output variables including Heating Load (HL) and Cooling Load (CL). A validation strategy identified EMARS to be the greatest method for forecasting HL and CL in comparison with other techniques.

Oduro et al. [77] confirmed the feasibility of utilizing parameters including acceleration, vehicle speed, ambient temperature, power and load to anticipate (NOX) emissions in order to perform management plans for air quality modeling. Accordingly, Oduro et al. [77] suggested utilizing the non-parametric Boosting-MARS (B-MARS) algorithm to enhance the precision of the MARS modeling to efficiently anticipate NOX emissions of automobiles in accordance with onboard measurements and the chassis dynamometer testing. The model technique offers more efficient outcomes of the evaluation and gives much better predictions of NOX emissions. The outcomes consequently recommend that the B-MARS technique is a beneficial and relatively precise approach for predicting NOX emissions and it may be followed by regulatory institutions.

Li et al. [78] recommended MARS predictive modeling as a reasonably basic nonlinear regression method for predicting solar energy output. Accordingly, MARS provides an easier approach compared to the nonlinear method including classification and regression tree (CART), ANN, k-nearest neighbors (KNN), and support vector machines (SVM). The MARS method was utilized on the daily output of a grid-connected 2.1 kW system to give the 1-day-ahead mean daily prediction of the energy output. The evaluations with a broad range of prediction methods demonstrate that the MARS method is capable to produce a trustworthy predictive performance.

Chen et al. [79] established a mesoscopic fuel usage evaluation model based on the MARS approach that can be applied to an eco-routing system. This suggested model provides a framework that uses real-world driving data, and clusters road links by free-flow speed. This particular model consists of predicting parameters that were seldom or never taken into consideration before, including the number of lanes and free-flow speed. The results of the statistical evaluations reveal that the independent parameters entered into the constructed model impact the fuel consumption rates of automobiles; however, the value and direction of the effects are reliant on the form of road networks, particularly free-flow speeds of networks. Despite the fact that the real-world driving information utilized to reveal statistical relationships are particular to one area; the framework created can be quickly adjusted and utilized to discover the fuel consumption relationship in other areas.

Another research by Özmen et al. [53] was conducted for the responsibility location of Baskentgaz that is the Local Distribution Companies (LDCs) of Ankara, Turkey. Predictive methods such as CMARS and MARS are implemented for one-day ahead usage of residential area consumers. The methods not only enable us to examine both techniques but also allow us to additionally evaluate the impact of real daily maximum and minimum temperatures compared to the Heating Degree Day (HDD) equal to their average. Utilizing one-day ahead models with daily information during 2009–2012, the daily usage of each day in 2013 has been forecasted and the outcomes are investigated with Linear Regression and ANNs. The results of this particular research reflected that CMARS and MARS techniques for the natural gas industry were determined to be two new alternatives.

In the research, data-driven methods for predicting short-term electricity demand (24 h) in Queensland, Australia, are adopted utilizing 0.5, 1.0, and 24-h predicting horizons by Al-Musaylh et al. [44]. These methods are depending on the ARIMA, SVR, and MARS techniques. To identify the SVR and MARS method inputs, the partial autocorrelation function is utilized for historical data in the training period to discriminate the substantial (lagged) inputs. Furthermore, individual input data is evaluated to create the ARIMA method. For the 0.5 and 1.0 h short-term predicting horizons, the MARS method was found to be much better in predictive performance than the ARIMA and SVR methods. On the contrary, the SVR method is exceptional to the MARS and ARIMA methods for predicting daily demand. Consequently, the SVR and MARS methods can be considered as more appropriate methods for short-term prediction, compared with the ARIMA method.

Ardakani et al. [80] predicted the yearly energy usage of the Nordic residential areas by 2020 as a function of environmental and socio-economic variables, and to provide a structure for the predictors in a separate nation. Accordingly, three models have been applied including the multiple linear regressions, MARS and the ANN approaches. The achieved results demonstrated that the ANN has an excellent ability to predict domestic energy utilization and indicating the significance of predictors in comparison to the classical regression method. The methods also pointed out that modifications in urban population, unemployment rate, workforce and amount of CO_2 emissions from the residential areas may result in substantial variations in Nordic domestic field energy utilization.

3. Material and method

3.1. Data collection

Candidate variables were chosen and appropriate periodical information was gathered from various government agencies. In general, three sources were utilized for data collection including the Turkish statistical institute [81], the world bank [82], and the general directorate of Turkish highways [83]. Accordingly, the magnitude of transport energy demand was collected from the Turkish Statistical Institute. Other variables, such as vehicle-km (i.e. the unit of traffic acquired by the movement of one motorized automobile throughout 1 km), ton-km (i.e. the unit of traffic acquired by carrying one ton of load over 1-km length) and passenger-km (i.e. the unit of measurement representing the transport of one traveler by an identified mode of transportation (air, rail, road, sea, inland waterways, etc.) around 1 km), were gathered from the general directorate of Turkish highways. The magnitude of GDP (i.e. the measure of a nation's total goods and services it generates for a given year) and oil price (i.e. one of the crucial indicators in terms of worldwide and national economic performance) were collected from the Turkish statistical institute. Lastly, the population which is the number of individuals living in a specific area with a border obtained from the world bank.

Based on the period of the data collection (1975–2019), population and GDP have increased from 39277258 to 44634 to 83154997 and 740000, respectively. It clearly shows that the mentioned variables have increased around 2 and 16 times from 1975. In addition, transport energy in 1975 was around 5148 million tons of oil equivalent (mtoe) and it reached up to 28 441 mtoe in 2019. In research conducted by Çodur and Ünal [2], ANN's method with the same independent variables from 1975 to 2016 were utilized to predict the transportation energy demand in Turkey. A part of historical data associated with the above-described variables from 2000 until 2019 is shown in detail in Table 1.

3.2. MARS data mining algorithm

MARS data mining algorithm developed by Friedman [84] as a non-parametric regression technique flexibly determines the linear, nonlinear, and interaction effects between sets of responses and predictors for classification and regression type problems. The model's nonlinearity is provided with various regression slopes for each predictor's intervals. The slopes of possible regression lines are set with knots, which are the links between separate regression splines [85].

The advantages of MARS, which can construct a flexibly predictive model by employing piecewise linear regressions, are that no assumptions about distributions of the variables and the functional relationship between response and predictor variables are required for MARS algorithm, which has a two-step process, a forward selection (first stage) and a backward deletion (second stage) approach [86]. MARS begins with a constant in the first model during the forward selection stage and iteratively includes pairs of basis functions producing the smallest training error in order to progress the model. The forward pass stage typically produces an overfitted model that has maximum complexity [87]. The obtained model in the forward pass stage has a very good fit, its generalization capability can be poor when a different data set is exposed to the model, meaning that there are overfitting problems. To remove these problems, the least contributing basis functions to the MARS model are removed one at a time in a backward stepwise deletion stage (also known as the second stage of the MARS modeling) [88]. MARS is like the stepwise linear regression analysis as a new term is either chosen or not chosen at each step. Kuhn and Johnson [89] described that it is easier to interpret the obtained results in the MARS algorithm compared with Artificial Neural Networks (ANNs).

Multicollinearity problem is more sensitive in especially interactive MARS predictive modeling compared with additive MARS predictive modeling. To remove this problem, it was suggested that predictors that have multicollinearity can be excluded from the constructed model, manually [85].

In this respect, we considered pairs of predictors i.e. oil price-GDP, oil price-population, oil price-ton, oil price-vehicle and oil price-passenger. At the same time, our main goal was to predict the energy amount using a few predictors for regression type problems.

As a part of predicting energy demand from pairs of the selected predictors, the MARS data mining algorithm was formulated as follows:

Table 1
Historical data related to the transportation energy demand.

Years	GDP (10^6)$	Population	Ton-Km (Million)	Vehicle-Km (Million)	Passenger-Km (Million)	Oil Price ($)	Transport Energy (MTOE)
2000	272979	63240194	161552	56151	185681	27.64	12007
2001	200251	64192243	151421	52631	168211	28.66	11999
2002	238428	65145367	150912	51664	163327	19.48	11404
2003	311823	66089402	152163	52349	164311	33.51	12395
2004	404786	67010930	156853	57767	174312	33.05	13774
2005	501416	67903469	166831	61129	182152	48.20	13850
2006	552486	68756810	177399	64577	187593	67.92	14980
2007	675770	69581848	181330	69609	209115	58.14	17263
2008	764335	70418604	181935	69771	206098	91.75	15976
2009	644639	71321399	176455	72432	212464	41.68	15896
2010	771901	72326988	190365	80124	226913	72.89	16607
2011	832523	73443863	203072	85495	242265	92.19	18760
2012	873982	74653016	216123	93989	258874	98.48	19768
2013	950579	75928564	224048	99431	268178	97.49	21007
2014	934185	77231907	234492	102988	276073	97.49	22218
2015	859796	78529409	244329	113274	290734	48.24	24936
2016	863721	79821724	253139	119671	300852	33.62	24951
2017	851556	81101892	262 739	127997	314734	52.81	25286
2018	766510	82319724	266 502	131625	329363	64.73	25543
2019	740000	83154997	267 579	133526	339601	53.79	25857

$$y = \beta_0 + \sum_{m=1}^{M} \beta_m \prod_{k=1}^{K_m} h_{km}\left(X_{v(k,m)}\right) \tag{1}$$

Where: $\widehat{y}$ is the predicted value of the response variable (energy), β_0 is a constant, β_m is the coefficient of basis functions, $h_{km}(X_{v(k,m)})$ is the basic function, in which $v(k,m)$ is an index of the predictor in the mth component of the kth product, K_m is the parameter limiting the order of interaction. After building the most complex MARS model in the forward pass stage, the basis functions decreasing the constructed model performance were removed in the pruning process (backward pass), depending on the following generalized cross-validation error (GCV) [85]:

$$GCV(\lambda) = \frac{\sum_{i=1}^{n}\left(y_i - y_{ip}\right)^2}{\left[1 - \frac{M(\lambda)}{n}\right]^2} \tag{2}$$

Where: n is the sample size, y_i is the observed value of the response variable (energy), y_{ip} is the predicted value of the response variable (energy), $M(\lambda)$ is the penalty function for the complexity of the model covering λ terms. The evaluated MARS predictive models with no interaction effect gave the smallest GCV [85]. In the earth package of R studio, the prediction process was performed with a total of 300 folds where three cross-validations were repeated 100 times, by using the ncross = 100 option in the MARS modeling. With a small data set, one may be willing to get a wise solution by using n cross argument, which helps to repeat the whole process of producing n fold folds multiple times in the earth R package. GR^2, which is a normalized form of GCV, reveals the generalization capability of the MARS model, which gives information about how well the MARS model constructed for the training set is able to predict for the testing set. GR^2 can be calculated as follows:

$$GR^2 = 1 - GCV \big/ GCV_{null} \tag{3}$$

Where GCV_{null} is the GCV value of the model including only an intercept [90].

To obtain a good fit for the MARS model, GR^2 must agree with R^2; otherwise, one encounters an overfitting problem.

The predictive quality of all the models was assessed with the following criteria calculated [88,91,92]:

1. Pearson correlation coefficient (r) between the actual and predicted values,
2. Akaike information criterion (AIC):

$$\begin{cases} AIC = n.ln\left[\frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_{ip}\right)^2\right] + 2k, & if\ n/k > 40 \\ AIC_c = AIC + \frac{2k(k+1)}{n-k-1} & otherwise \end{cases} \tag{4}$$

3. Root-mean-square error ($RMSE$):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_{ip}\right)^2} \tag{5}$$

4. Mean error (ME):

$$ME = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_{ip}\right) \tag{6}$$

5. Mean absolute deviation (MAD):

$$MAD = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - y_{ip}\right| \tag{7}$$

6. Standard deviation ratio (SD_{ratio}):

$$SD_{ratio} = \frac{S_m}{S_d} \tag{8}$$

7. Global relative approximation error (RAE):

$$RAE = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - y_{ip}\right)^2}{\sum_{i=1}^{n} y_i^2}} \tag{9}$$

8. Mean absolute percentage error (*MAPE*):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y_{ip}}{y_i}\right| * 100 \tag{10}$$

Where n is the sample size in the data set, k is the number of model parameters, y_i is the real value of the response variable (energy), y_{ip} is the predicted value of the response variable (energy), s_m is the standard deviation of the model errors, s_d is the standard deviation of the response variable (energy).

For building MARS predictive models, the earth package proposed by Milborrow [93,94] in R Studio software was used [92]. Train function of caret (Classification and Regression Training) package was also specified in the R environment in order to perform model construction, assessment and optimization process [85]. Also, the ehaGoF package was used to measure the predictive quality of the evaluated MARS models [92].

4. Model development and discussion

Summary results of the goodness of fit criteria for five candidate MARS models are given in Table 2. Among the candidate models, the third MARS model with ton and oil price predictors was found as the best predictive model, which had the smallest RMSE, SD ratio, AICc values and the highest R^2, Adj. R^2 and especially GR^2 values. Also, the least difference between R^2 and Adj. R^2 was obtained for the third MARS predictive model. The reason was that the MARS algorithm shows a tendency of including only significant predictors in the constructed model. It was followed by the second MARS model with pop and oil price predictors. The worst MARS model was determined to be the model with GDP and oil price predictors (MARS-1). Except for the first MARS predictive model, it was observed that generalization capabilities (real predictive capabilities) of the remaining MARS models were very good due to their GR^2 and CVR^2 estimates, meaning that there was no overfitting problem. In this case, we could rely on the predictive ability of the last four models. The handled MARS models were constructed based on the smallest GCV values. The third MARS model had only 5 terms obtained in the backward deletion phase of the MARS modeling.

Fig. 1 indicates the agreement between observed and predicted values for the handled MARS models. The best agreement was obtained in the third MARS predictive model (Fig. 1), which supported the results given in Table 2.

The third MARS model was found as the best model with the highest R^2, CVR^2, Adjusted R^2 and GR^2 values. The third MARS predictive model was:

Energy demand = 12643−0.05013477 * max(0, 161552 - ton) + 0.2191397 * max(0, ton - 161552) - 0.08205221 * max(0, ton - 181330) - 35.86154 * max(0, oil - 48.24)

We can rewrite the MARS model above in BF notation.

Energy demand = 12643−0.05013477 * BF1 + 0.2191397 * BF2 - 0.08205221 * BF3 - 35.86154 * BF4.

Basis functions in MARS prediction equation of the energy demand mentioned above can be explained as follows:

BF1 = max(0, 161552-ton) where 161552 is the cutpoint value for BF1.
BF2 = max(0, ton-161552) where 161552 is the cutpoint value for BF2.
BF3 = max(0, ton-181330) where 181330 is the cutpoint value for BF3.
BF4 = max(0, oil-48.24) where 48.24 is the cutpoint value for BF4.

Here, the following interpretations are made using the cutpoint values of the third MARS model selected as the best predictive performance. In the third MARS model, the coefficient (0.2191397) of the BF2 basis function is positive, negative coefficients of BF1 (-0.05013477), BF3 (- 0.08205221), and BF4 (- 35.86154) functions are available. Among the basic functions i.e. BF1, BF2, BF3 and BF4, the BF2 has the greatest coefficient, and the smallest coefficient in the third MARS predictive model was detected to be BF4 with -35.86154.

When the value of ton predictor is equal to 161552, the effect of BF1 and BF2 on energy is masked, meaning that there is no effect of ton predictor on the energy demand. When the value of the ton predictor is greater than 161552, no effect of ton predictor on the energy demand was found for BF1; however, there is an increasing effect of ton predictor on energy demand for BF2.

When the value of ton predictor is greater than 181330 in the third MARS predictive model.

a The influence of ton predictor on energy response variable was masked for BF1 where max (0, 161552-ton) = 0
b The increasing effect of ton predictor on energy was determined for BF2 due to the positive coefficient (0.2191397) of BF2 when the value of ton predictorreached from 181330 to the possible maximum level,
c The decreasing effect often predictor on energy demand was noted as a result of the negative effect of BF3 on energy.

When the value of ton predictor is less than 181330 in the third MARS predictive model.

a The influence of the basis functions BF1 and BF3 on energy demand was masked regardless of corresponding coefficients.
b BF2 had an increasing effect on energy demand due to its positive coefficient.

When the value of ton predictor is 161552 < ton <181330 in the third MARS predictive model.

Table 2
The goodness of fit criteria for candidate MARS models.

	GR^2	CVR^2	RMSE	SD_{ratio}	MAPE	R^2	Adj.R^2	AICc
Model 1 (GDP-oil price)	0.83972	0.8124433	1878.1459	0.297	7.600	0.9117724	0.895	494.742
Model 2 (pop-oil price)	0.97859	0.9706394	746.0166	0.118	5.136	0.9860799	0.984	432.825
Model 3 (ton-oil price)	0.97965	0.9684095	669.1806	0.106	5.405	0.9887996	0.987	428.695
Model 4 (vehicle-oil price)	0.97807	0.9682204	815.4999	0.129	5.934	0.9833661	0.981	435.900
Model 5 (passenger-oil price)	0.97233	0.9540166	780.2990	0.123	5.889	0.9847711	0.982	438.527

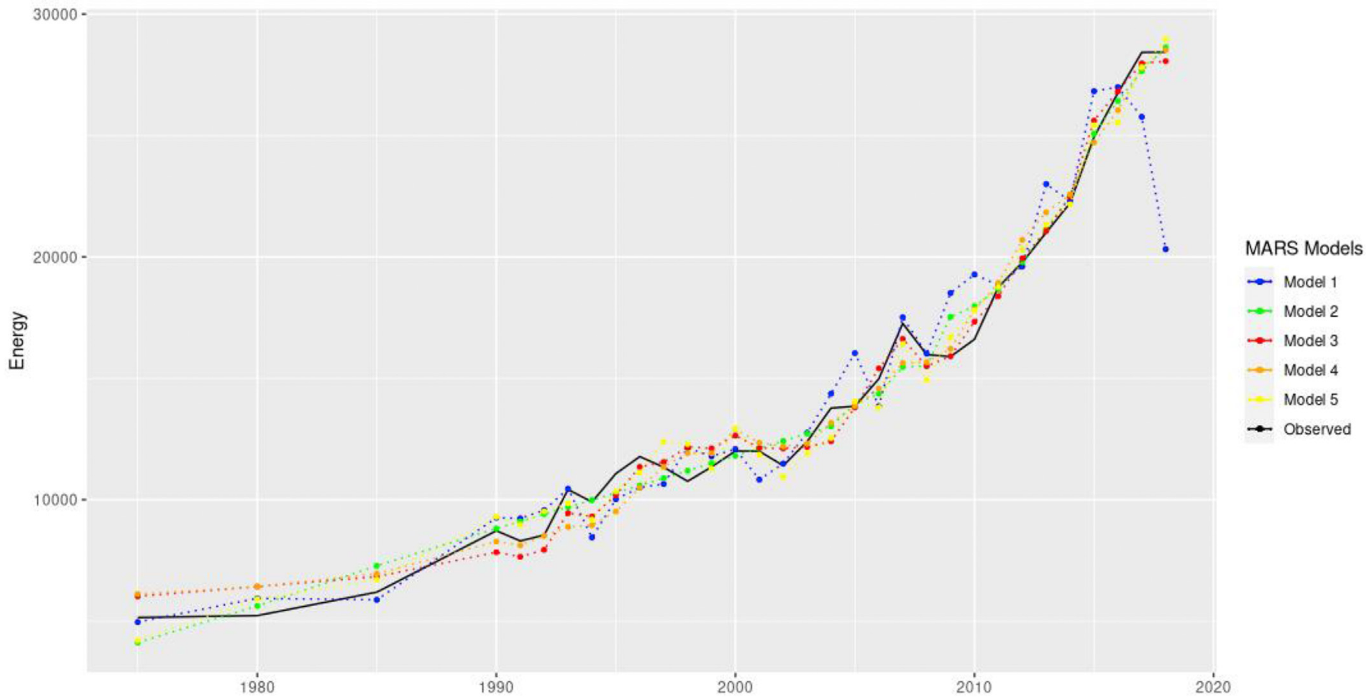

Fig. 1. The agreement graph between observed and predicted values for each model.

Table 3
Forecasted transport energy demand for Model 3.

Years	Oil price ($)	Ton-km (Million)	Energy Demand (MTOE)
2020	74.82	277137	26139
2023	82.28	292014	27054
2025	86.23	303386	27674
2030	94.52	334476	29285

a The influence of BF1 and BF3 basis functions on energy demand was masked.
b The increasing effect of BF2 on energy demand is obtained due to its positive coefficient.

When the values of oil price and ton predictors are considered to be oil price $\leq$48.24 and 161552 < ton $\leq$181330 in the third MARS predictive model, the predicted energy demand value is on the increase because BF1 = 0, BF3 = 0, BF4 = 0 and BF2 are gradually increasing. From MARS 3 predictive model with interaction effect, it was understood that the effect of a ton on energy was independent of the oil price predictor.

Table 3 shows the forecasted transportation energy demand magnitudes based on the oil price and Ton-km values (model 3). Accordingly, the oil price will reach 74.82, 82.28, 86.23, and 94.52 for 2020, 2023, 2025 and 2030, respectively. In addition, energy demand forecasted to be around 26139, 27054, 27674, and 29285 (MTOE) for a period of 10 years between 2020 until 2030.

5. Conclusion

This study seeks an application of the MARS approach as a tool for modeling and prediction techniques for the energy demand of the transportation sector in Turkey. Transport energy demand is analyzed for the period 1975–2019 using GDP, population, vehicle-km, ton-km, passenger-km and oil price as variables. Among the evaluated predictive MARS models, the worst model was determined to be the first MARS predictive model with GDP and oil price predictors, while the third MARS predictive model, which had the highest R^2, Adj. R^2 and especially GR^2 values, with ton and oil price predictors, were found as the best predictive model,. Furthermore, the least difference between R^2 and Adj. R^2 estimates were obtained for the third MARS predictive model. Thus, the government needs to motivate people to utilize public transportation as much as possible, as well as such vehicles as bicycles for short-range travel. These cheap options will require less fuel and create lower emissions. Since road transport is the greatest consumer of energy within the whole transportation sector, therefore, the appropriate approach for standard energy consumption for automobiles in order to decrease energy usage must be provided. The technique utilized in this research demonstrates that the MARS method is beneficial for investigating and predicting energy usage in the transportation sector. As for future research, more different variables might be considered to estimate the transportation energy demand, for example, technological developments, environmental effects, and so on. As a result, the use of the predictive MARS modeling may be a flexibly considerable option for the prediction of energy demand with the scope of regression type problems.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Al-Mofleh A, Taib S, Salah WA. Malaysian energy demand and emissions from the transportation sector. Transport 2010;25(4):448–53.
[2] Yasin Çodur M, Ünal A. An estimation of transport energy demand in Turkey via artificial neural networks. Promet - Traffic & Transp 2019;31(2):151–61.
[3] Muratori M, Moran MJ, Serra E, Rizzoni G. Highly-resolved modeling of personal transportation energy consumption in the United States. Energy 2013;58:168–77.

[4] Limanond T, Jomnonkwao S, Srikaew A. Projection of future transport energy demand of Thailand. Energy Pol 2011;39(5):2754–63.
[5] Haldenbilen S, Ceylan H. Genetic algorithm approach to estimate transport energy demand in Turkey. Energy Pol 2005;33(1):89–98.
[6] WECTNC. World energy Council Turkish national Committee. 2000.
[7] Hepbasli A, Oturanc G, Kurnaz A, Ergin E, Genc A, Iyit N. Simple correlations for estimating the energy production of Turkey. Energy Sources 2002;24(9): 855–67.
[8] Hepbasli A, Ozalp N. Development of energy efficiency and management implementation in the Turkish industrial sector. Energy Convers Manag 2003;44(2):231–49.
[9] Anand A, Suganthi L. Forecasting of electricity demand by hybrid ANN-PSO models. Deep Learning and Neural Networks: Concepts, Methodologies, Tools, and Applications 2020:865–82. IGI Global.
[10] Hrnjica B, Mehr AD. Energy demand forecasting using deep learning. Smart cities performability, cognition, & security. Springer; 2020. p. 71–104.
[11] Masoumi A, Ghassem-zadeh S, Hosseini SH, Ghavidel BZ. Application of neural network and weighted improved PSO for uncertainty modeling and optimal allocating of renewable energies along with battery energy storage. Appl Soft Comput 2020;88:105979.
[12] Bui D-K, Nguyen TN, Ngo TD, Nguyen-Xuan H. An artificial neural network (ANN) expert system enhanced with the electromagnetism-based firefly algorithm (EFA) for predicting the energy consumption in buildings. Energy 2020;190:116370.
[13] Zhang M, Mu H, Li G, Ning Y. Forecasting the transport energy demand based on PLSR method in China. Energy 2009;34(9):1396–400.
[14] Ciulla G, D'Amico A. Building energy performance forecasting: a multiple linear regression approach. Appl Energy 2019;253:113500.
[15] Johannesen NJ, Kolhe M, Goodwin M. Relative evaluation of regression tools for urban area electrical energy demand forecasting. J Clean Prod 2019;218: 555–64.
[16] Akdi Y, Gölveren E, Okkaoğlu Y. Daily electrical energy consumption: periodicity, harmonic regression method and forecasting. Energy 2020;191: 116524.
[17] Zhang M, Li G, Mu H, Ning Y. Energy and exergy efficiencies in the Chinese transportation sector, 1980–2009. Energy 2011;36(2):770–6.
[18] Torrini FC, Souza RC, Oliveira FLC, Pessanha JFM. Long term electricity consumption forecast in Brazil: a fuzzy logic approach. Soc Econ Plann Sci 2016;54:18–27.
[19] Ahn J, Cho S, Chung DH. Analysis of energy and control efficiencies of fuzzy logic and artificial neural network technologies in the heating energy supply system responding to the changes of user demands. Appl Energy 2017;190: 222–31.
[20] Mukhopadhyay P, Mitra G, Banerjee S, Mukherjee G. Electricity load forecasting using fuzzy logic: short term load forecasting factoring weather parameter. Conference Electricity load forecasting using fuzzy logic: short term load forecasting factoring weather parameter. High Perform Dev IEEE Cornell Conf, p. 812-819.
[21] Ray P, Panda SK, Mishra DP. Short-term load forecasting using genetic algorithm. Computational Intelligence in Data Mining 2019:863–72. Springer.
[22] Kampelis N, Sifakis N, Kolokotsa D, Gobakis K, Kalaitzakis K, Isidori D, et al. HVAC Optimization genetic algorithm for industrial near-zero-energy building demand response. Energies 2019;12(11):2177.
[23] Shin Y, Kim Z, Yu J, Kim G, Hwang S. Development of NOx reduction system utilizing artificial neural network (ANN) and genetic algorithm (GA). J Clean Prod 2019;232:1418–29.
[24] Moraes LAd, Flauzino RA, Araújo MAd, Batista OE. A fuzzy methodology to improve time series forecast of power demand in distribution systems. Conference A fuzzy methodology to improve time series forecast of power demand in distribution systems. IEEE, p. 1-5.
[25] Box GE, Jenkins GM, Reinsel GC, Ljung GM. Time series analysis: forecasting and control. John Wiley & Sons; 2015.
[26] Efendi R, Ismail Z, Deris MM. A new linguistic out-sample approach of fuzzy time series for daily forecasting of Malaysian electricity load demand. Appl Soft Comput 2015;28:422–30.
[27] Raza SA, Shah N, Sharif A. Time frequency relationship between energy consumption, economic growth and environmental degradation in the United States: evidence from transportation sector. Energy 2019;173:706–20.
[28] Sa'ad S. Improved technical efficiency and exogenous factors in transportation demand for energy: an application of structural time series analysis to South Korean data. Energy 2010;35(7):2745–51.
[29] Shaikh F, Ji Q, Shaikh PH, Mirjat NH, Uqaili MA. Forecasting China's natural gas demand based on optimised nonlinear grey models. Energy 2017;140: 941–51.
[30] Xu N, Dang Y, Gong Y. Novel grey prediction model with nonlinear optimized time response method for forecasting of electricity consumption in China. Energy 2017;118:473–80.
[31] Şahin U. Forecasting of Turkey's greenhouse gas emissions using linear and nonlinear rolling metabolic grey model based on optimization. J Clean Prod 2019;239:118079.
[32] Sadri A, Ardehali M, Amirnekooei K. General procedure for long-term energy-environmental planning for transportation sector of developing countries with limited data based on LEAP (long-range energy alternative planning) and EnergyPLAN. Energy 2014;77:831–43.
[33] Shu M-H, Hung W, Nguyen T, Hsu B, Lu C. Forecasting with Fourier residual modified ARIMA model-An empirical case of inbound tourism demand in New Zealand. WSEAS Trans Math 2014;13(1):12–21.
[34] Lee D, Lee D, Choi M, Lee J. Prediction of network throughput using ARIMA. Conference prediction of network throughput using ARIMA. IEEE, p. 1-5.
[35] Lewis PA, Stevens JG. Nonlinear modeling of time series using multivariate adaptive regression splines (MARS). J Am Stat Assoc 1991;86(416):864–77.
[36] Haas H, Kubin G. A multi-band nonlinear oscillator model for speech. Conference A multi-band nonlinear oscillator model for speech, vol. vol. 1. IEEE, p. 338-342.
[37] Ekman T, Kubin G. Nonlinear prediction of mobile radio channels: measurements and MARS model designs. Conference Nonlinear prediction of mobile radio channels: measurements and MARS model designs, vol. vol. 5. IEEE, p. 2667-2670.
[38] Mukkamala S, Sung AH. A comparative study of techniques for intrusion detection. Conference A comparative study of techniques for intrusion detection. IEEE, p. 570-577.
[39] Yang C-C, Prasher SO, Lacroix R, Kim SH. A multivariate adaptive regression splines model for simulation of pesticide transport in soils. Biosyst Eng 2003;86(1):9–15.
[40] Abraham A. Analysis of hybrid soft and hard computing techniques for forex monitoring systems. Conference Analysis of hybrid soft and hard computing techniques for forex monitoring systems, vol. vol. 2. IEEE, p. 1616-1621.
[41] Lee T-S, Chen I-F. A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. Expert Syst Appl 2005;28(4):743–52.
[42] Chou S-M, Lee T-S, Shao YE, Chen I-F. Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines. Expert Syst Appl 2004;27(1):133–42.
[43] Sigauke C, Chikobvu D. Daily peak electricity load forecasting in South Africa using a multivariate non-parametric regression approach. Orion 2010;26(2).
[44] Al-Musaylh MS, Deo RC, Adamowski JF, Li Y. Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia. Adv Eng Inf 2018;35:1–16.
[45] Al-Musaylh MS, Deo RC, Li Y, Adamowski JF. Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting. Appl Energy 2018;217:422–39.
[46] Deo RC, Samui P, Kim D. Estimation of monthly evaporative loss using relevance vector machine, extreme learning machine and multivariate adaptive regression spline models. Stoch Environ Res Risk Assess 2016;30(6):1769–84.
[47] Özmen A, Batmaz İ, Weber G-W. Precipitation modeling by polyhedral RCMARS and comparison with MARS and CMARS. Environ Model Assess 2014;19(5):425–35.
[48] Sharda V, Prasher S, Patel R, Ojasvi P, Prakash C. Performance of Multivariate Adaptive Regression Splines (MARS) in predicting runoff in mid-Himalayan micro-watersheds with limited data. Hydrol Sci J 2008;53(6):1165–75.
[49] Roy SS, Roy R, Balas VE. Estimating heating load in buildings using multivariate adaptive regression splines, extreme learning machine, a hybrid model of MARS and ELM. Renew Sustain Energy Rev 2018;82:4256–68.
[50] Cheng M-Y, Cao M-T. Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines. Appl Soft Comput 2014;22:178–88.
[51] Williams KT, Gomez JD. Predicting future monthly residential energy consumption using building characteristics and climate data: a statistical learning approach. Energy Build 2016;128:1–11.
[52] Tian W. A review of sensitivity analysis methods in building energy analysis. Renew Sustain Energy Rev 2013;20:411–9.
[53] Özmen A, Yılmaz Y, Weber G-W. Natural gas consumption forecast with MARS and CMARS models for residential users. Energy Econ 2018;70:357–81.
[54] Morana C. A semiparametric approach to short-term oil price forecasting. Energy Econ 2001;23(3):325–38.
[55] Zareipour H, Bhattacharya K, Canizares CA. Forecasting the hourly Ontario energy price by multivariate adaptive regression splines. Conference Forecasting the hourly Ontario energy price by multivariate adaptive regression splines. IEEE, p. 7 pp.
[56] Bakirtas T, Akpolat AG. The relationship between energy consumption, urbanization, and economic growth in new emerging-market countries. Energy 2018;147:110–21.
[57] Yuksel S, Ubay GG. Identifying the influencing factors of renewable energy consumption in Turkey with MARS methodology. Ekonomi İşletme ve Maliye Araştırmaları Dergisi 2020;2(1):1–14.
[58] Say NP, Yücel M. Energy consumption and CO2 emissions in Turkey: empirical analysis and future projection based on an economic growth. Energy Pol 2006;34(18):3870–6.
[59] Baskan O, Haldenbilen S, Ceylan H. Estimating transport energy demand using ant colony optimization. Energy Sources B Energy Econ Plann 2012;7(2): 188–99.
[61] Suganthi L, Samuel AA. Energy models for demand forecasting—a review. Renew Sustain Energy Rev 2012;16(2):1223–40.
[62] Amasyali K, El-Gohary NM. A review of data-driven building energy consumption prediction studies. Renew Sustain Energy Rev 2018;81:1192–205.
[63] Abu-El-Magd MA, Sinha NK. Short-term load demand modeling and forecasting: a review. IEEE transactions on systems, man, and cybernetics 1982;12(3):370–82.
[64] Ladan G, Kalantar M. Different methods of longterm electric load demand

forecasting; a comprehensive review. Iranian Journal of Electrical & Electronic Engineering 2011;7(4):249–59.
[65] Mukherjee SK. Energy demand forecasting: a Critical review of Current Approaches. Washington DC: Energy, Water and Telecommunication Department; 1977.
[66] Bajay SV. Long-term electricity demand forecasting models: a review of methodologies. Elec Power Syst Res 1983;6(4):243–57.
[67] Swan LG, Ugursal VI. Modeling of end-use energy consumption in the residential sector: a review of modeling techniques. Renew Sustain Energy Rev 2009;13(8):1819–35.
[68] Hyndman RJ. Review of Transpower's electricity demand forecasting methods. 2011.
[69] Shunping J, Hongqin P, Shuang L, Zhang X. Review of transportation and energy consumption related research. Journal of Transportation Systems Engineering and Information Technology 2009;9(3):6–16.
[70] Zhang X-d, Yang L. Study on combination forecasting model for traffic energy demand. J Nanjing Inst Technol 2008;2.
[71] Sonmez M, Akgüngör AP, Bektaş S. Estimating transportation energy demand in Turkey using the artificial bee colony algorithm. Energy 2017;122:301–10.
[72] Canyurt OE, Ozturk HK, Hepbasli A, Utlu Z. Genetic algorithm (GA) approaches for the transport energy demand estimation: model development and application. Energy Sources, Part A 2006;28(15):1405–13.
[73] Lu I, Lewis C, Lin SJ. The forecast of motor vehicle, energy demand and CO2 emission from Taiwan's road transportation sector. Energy Pol 2009;37(8): 2952–61.
[74] Yang H. Analysis on transportation energy consumption and saving potentiality in China. Energy Policy Research 2007;5:51–5.
[75] Abraham A, Steinberg D. Is neural network a reliable forecaster on earth? a MARS query! Conference Is neural network a reliable forecaster on earth? a MARS query! Springer, p. 679-686.
[76] Deo RC, Kisi O, Singh VP. Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. Atmos Res 2017;184:149–75.
[77] Oduro SD, Metia S, Duc H, Hong G, Ha Q. Multivariate adaptive regression splines models for vehicular emission prediction. Visualization in Engineering 2015;3(1):13.
[78] Li Y, He Y, Su Y, Shu L. Forecasting the daily power output of a grid-connected photovoltaic system based on multivariate adaptive regression splines. Appl Energy 2016;180:392–401.
[79] Chen Y, Zhu L, Gonder J, Young S, Walkowicz K. Data-driven fuel consumption estimation: a multivariate adaptive regression spline approach. Transport Res C Emerg Technol 2017;83:134–45.
[80] Ardakani SR, Hossein SM, Aslani A. Statistical approaches to forecasting domestic energy consumption and assessing determinants: the case of Nordic countries. Strat Plann Energy Environ 2018;38(1):26–71.
[81] Turkstat. Turkish Statistical Institut; 2020.
[82] Bank TW. Population. 2020.
[83] GDTH. General directorate of Turkish highways. 2020.
[84] Friedman JH. Multivariate adaptive regression splines. Ann Stat 1991:1–67.
[85] Akin M, Eyduran SP, Eyduran E, Reed BM. Analysis of macro nutrient related growth responses using multivariate adaptive regression splines. Plant Cell Tissue Organ Cult 2020;140(3):661–70.
[86] Arthur CK, Temeng VA, Ziggah YY. Multivariate Adaptive Regression Splines (MARS) approach to blast-induced ground vibration prediction. Int J Min Reclamat Environ 2020;34(3):198–222.
[87] Weber G-W, Batmaz İ, Köksal G, Taylan P, Yerlikaya-Özkurt F. CMARS: a new contribution to nonparametric regression with multivariate adaptive regression splines supported by continuous optimization. Inverse Problems in Science and Engineering 2012;20(3):371–400.
[88] Zaborski D, Ali M, Eyduran E, Grzesiak W, Tariq MM, Abbas F, et al. Prediction of selected reproductive traits of indigenous Harnai sheep under the farm management system via various data mining algorithms. Pakistan J Zool 2019;51(2):421.
[89] Kuhn M, Johnson K. Applied predictive modeling. Springer; 2013.
[90] Huang H, Ji X, Xia F, Huang S, Shang X, Chen H, et al. Multivariate adaptive regression splines for estimating riverine constituent concentrations. Hydrol Process 2020;34(5):1213–27.
[91] Zhang W, Goh AT. Multivariate adaptive regression splines and neural network models for prediction of pile drivability. Geoscience Frontiers 2016;7(1):45–52.
[92] Eyduran E, Akın M, Eyduran SP. Application of multivariate adaptive regression splines in agricultural sciences through R software. Nobel Bilimsel Eserler Sertifika; 2019. p. 20779.
[93] Milborrow S. Derived from Mda: Mars by Trevor Hastie and Rob Tibshirani. Uses Alan Miller's Fortran utilities with Thomas Lumley's leaps wrapper. earth: Multivariate Adaptive Regression Splines. R package version 4.4. 4. 20164; 2016.
[94] Milborrow MS. Package 'earth'. R Software package. 2019.