Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2022) 30: 1251 – 1268

doi:10.55730/1300-0632.3847

An adaptive search equation-based artificial bee colony algorithm for transportation energy demand forecasting

Durmuş ÖZDEMİR*, **Safa DÖRTERLER**
Department of Computer Engineering, Kütahya Dumlupınar University, Kütahya, Turkey

Received: 10.06.2021 • **Accepted/Published Online:** 05.04.2022 • **Final Version:** 31.05.2022

Abstract: This study aimed to develop a new adaptive artificial bee colony (A-ABC) algorithm that can adaptively select an appropriate search equation to more accurately estimate transport energy demand (TED). Also, A-ABC and canonical artificial bee colony (C-ABC) algorithms were compared in terms of efficiency and performance. The input parameters used in the proposed TED model were the official economic indicators of Turkey, including gross domestic product (GDP), population, and total vehicle kilometer per year (TKM). Three mathematical models, linear (A-ABCL), exponential (A-ABCE), and quadratic (A-ABCQ) were developed and tested. Also, economic variables were generated using the "curve fitting" technique to see TED's projections for the year 2034, under two different scenarios. In the first scenario, the results of linear, exponential, and quadratic models according to 2034 TED estimates were 40.1, 31.6, and 70.5 million tons of oil equivalent (Mtoe), respectively. In the second scenario, the results of linear, exponential, and quadratic models according to the TED estimates for 2034 were found as 40.0, 31.5, and 66.5 Mtoe, respectively. The presented models, A-ABCL, A-ABCE, A-ABCQ for the solution of the TED problem, produced successful results compared to the studies in the literature. Besides that, according to global error metrics, developed models generated lower error values than C-ABC. Furthermore, consumption estimation values of A-ABCL and A-ABCE were lower than A-ABCQ. According to A-ABCQ model estimations for both scenarios, the TED value would increase approximately three times from 2013 to 2034.

Key words: Adaptive artificial bee colony, transportation energy demand estimation, metaheuristic algorithms, optimization

1. Introduction

Energy consumption in the transportation sector has a significant share in the government's energy resources management processes. In Turkey, the total energy consumption in 2018 was 143.666 million tons of oil equivalent (Mtoe). According to the Ministry of Environment and Urbanization report, the energy consumption value in the transportation sector is 19.8%, and this value ranks fourth among all sectors. The transportation energy consumption of Turkey between the years 2002 and 2019 grew steadily except in 2009 [1]. It could be said that due to its rising energy demand and limited energy resources, Turkey is significantly dependent on energy imports. Currently, nearly three-quarters of the total energy demand of Turkey in industrial, transportation, commercial areas, etc. has to be imported [2]. This demand could increase by over 80% until 2030. Nowadays, the total energy demand of Turkey is obtained from oil as 29.2%, natural gas as 28.7%, solid fuel as 28.4%, and

*Correspondence: durmus.ozdemir@dpu.edu.tr

the rest of it from wind, solar, hydro energy, etc.[1]. In 2020, Turkey's dependence on oil imports was 92.9%. Approximately 29.1% of it came from Iraq, 21.2% came from Russia, 8.2% came from Kazakhstan, 7.9% came from Saudi Arabia, 7% came from Norway, 6.1% came from Nigeria, 5.1% came from India, and 3.5% came from Israel [3]. Compared to the European Union and other neighbors, it is concluded that Turkey is dependent on imports in energy. Because energy production costs are high for Turkey, successful planning for balancing energy production and consumption is necessary [4, 5].

In the transportation area, energy dependency is more apparent. With the increase in economic development and social welfare, people have changed their travel habits by using individual vehicles instead of public transport. In addition, along with economic development, there was a significant increase in commercial shipments in all areas. According to data from the Turkish Statistical Institute (TSI[1].), the number of registered motor vehicles such as automobiles (54.3%), pick-up trucks (16.3), motorcycles (14.8%), tractors (8%), trucks (3.5%), mini-busses (2%), and buses (0.8%) increased from 8,903,843 in 2003 to 25,105,532 in 2021 (as shown in Figure 1a). There has been a 182% increase in the number of vehicles in 19 years. Especially in the individual vehicle category, there was an increase of 54.3%. Also, the rate of individual vehicle users is expected to increase gradually. According to a study, there has been approximately 95% decrease in the tendency of using public transportation during the COVID-19 pandemic in Turkey and this situation increased individual vehicle use, which causes more fuel consumption [6]. The transportation sector is an essential part of any modern government's economic system, and literature shows that the transportation sector has a vital force on economic growth [7]. This sector constitutes a high proportion of the energy demand, particularly for petroleum products. In the transportation sector of Turkey, the overall share of gasoline, diesel, and substitute fuels in energy consumption were 26.1%, 35.5%, and 38.4%, respectively (as shown in Figure 1b)[2].

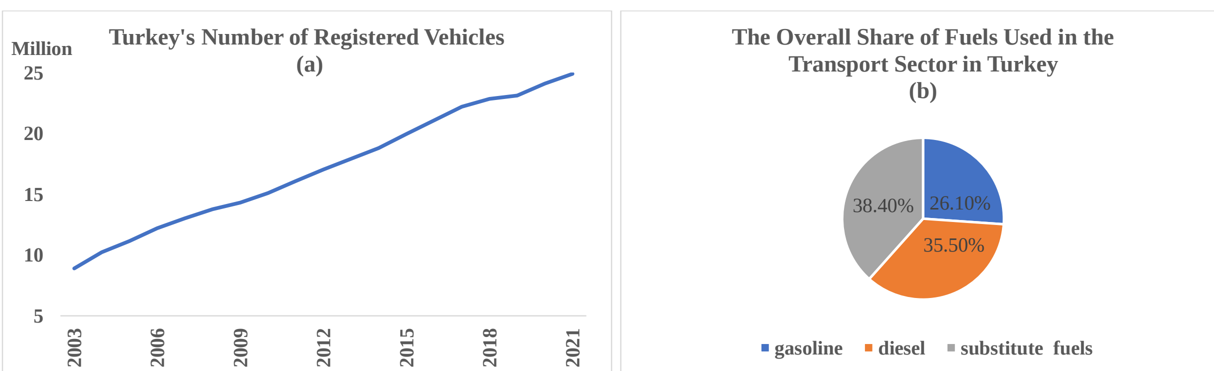

Figure 1. Number of registered vehicles and the overall share of fuels used in the transport sector in Turkey.

In Turkey, from 2017 to 2018, total imports of gasoline products increased by 17.99%, crude gasoline imports increased by 53.70%, and fuel imports increased by 0.55% [8]. Transportation energy demand models are typically based on social and economic factors, such as population, GDP, total vehicle-km, wages, ownership of vehicles, population growth, and the number of trips. Therefore, the modeling of energy consumption demand in the transportation sector differs from modeling other sectors' energy demand.

The current study presents a comparative analysis of C-ABC and A-ABC algorithms to solve the transportation energy demand problem. Although the study was carried out using Turkey's data sets, application

[1]TSI (2022). Turkish Statistical Institute [online]. Website https://www.tuik.gov.tr [accessed 10 Jan 2022].
[2]Eurostat (2021). European Community Statistical [online]. Website https://ec.europa.eu/eurostat [accessed 10 July 2021].

areas can be expanded using different countries' data sets. Linear, exponential, and quadratic mathematical regression estimation models were developed, and all models were compared. In the developed models, GDP, population, and TKM data were used as input parameters. In order to create future projections about the transportation energy consumption of Turkey, estimations were presented under two different scenarios. Thus, this study aims to solve the transportation energy demand estimation problem, which is a real-world problem, and obtain the optimum result.

The paper is organized as follows: Section 2 presents a literature review. The details of the C-ABC are explained in Section 3. In Section 4, the details of the A-ABC and the differences between A-ABC and C-ABC are explained. In Section 5, the developed model for transport energy demand is presented. The experimental results of the developed model are presented in Section 6. In Section 7, the estimation of TED of Turkey by 2034 is discussed based on two different scenarios. Finally, policy recommendations are presented in the conclusion section.

2. Literature review

Turkey is one of Europe's major economies, with a GDP of about $761.425 billion, and it is one of the 19th most significant economies in the world[3]. The energy consumption of Turkey was about 83 Mtoe in 2000. This figure increased by about 27% to 1023 Mtoe in 2017. Because of the rising economy, increasing population, and developments in the transport sector, TED, population, and vehicle kilometer increased around 2.6, 1.25, and 6.1 times, respectively, over the last two decades[4].

In the literature, there are two main methods used to solve the TED problems discussed in the research. The first one is the artificial intelligence (AI) application method, and the second is regression analysis (RA) [9, 10]. According to the literature, artificial neural network (ANN) models can be used to plan future projections of TED [11, 12]. In addition to ANN approaches, optimization algorithms such as genetic algorithm (GA), ant colony optimization (ACO), particle swarm optimization (PSO), and harmony search algorithm (HSA) of energy demand forecasts are used effectively.

Yasin and Codur [11] presented a TED study using economic and demographic parameters with the ANN method. In the study, they compared and analyzed seven different models. They found that oil price, population, and vehicle-km input parameters were the most successful model in the study. Amiri et al. [13] evaluated the capabilities of the linear regression model and neural network model in predicting household TED by considering the number of motorized trips and the travel distance. Leo et al. [14] determined the long-term trend of energy demand using population and GDP parameters through the regression analysis method. In this way, they proposed a model that explains the relationships between residential, transport, and commercial energy demands. Statistical testing confirmed the effectiveness of the proposed regression models. Chai et al. [15] developed autoregressive integrated moving average (ARIMA) model and multiple regression models for China using parameters such as GDP, total road turnover, highway mileage, urbanization rate, etc.

According to the literature review, the GA is also used to solve the TED problem. Sahraei and Codur [16] used hybrid versions of various algorithms, namely GA-ANN, ANN-simulated annealing, and ANN-PSO, to estimate Turkey's transportation energy. They used different combinations of input parameters such as GDP, vehicle-km, population, oil price, passenger-km, and ton-km to create 11 different models in their study. They concluded that the ANN-PSO model developed with GDP, population, ton-km parameters is the most

[3]WB (2021). The World Bank [online]. Website https://data.worldbank.org/ [accessed 10 July 2021].
[4]TSI (2022). Turkish Statistical Institute [online]. Website https://www.tuik.gov.tr [accessed 10 Jan 2022].

successful model. Ray et al. [17] conducted short-term electrical power load estimation studies in China using GA-based backpropagation model. Furthermore, Canyurt & Ozturk [18] predicted fossil fuels consumption using nonlinear mathematical model with GA. In another study, GDP, electricity consumption per capita income growth rate, and consumer price index values were used as input parameters, and electricity demand of Tamil Nadu state in India was estimated with various scenarios with Hybrid GA-PSO algorithms [19]. Bilgili et al. [20] implemented the short-term electricity energy consumption prediction with four time-series methods; long short-term memory (LSTM) neural network, adaptive neuro-fuzzy inference system (ANFIS) with subtractive clustering, ANFIS with fuzzy C-means, and ANFIS with grid partition (GP), respectively. As a consequence, the LSTM model generally outperformed all ANFIS models. Also, Sahraei et al. [21] estimated energy consumption in transportation sector using Multivariate Adaptive Regression Splines (MARS). They developed five different models with the combination of parameters such as GDP, population, vehicle-km, ton-km, passenger-km, oil price. They got the best result from the model created with ton and oil price parameters.

Similarly, other AI optimization techniques and heuristic algorithms were used for energy demand estimation studies conducted in the world. Toksari [22] estimated Turkey's TED using GDP, population, and total annual vehicle-km with the hybrid algorithm of variable neighborhood search algorithms and ACO. In the study, the hybrid algorithm was modeled using two different mathematical forms. According to the findings, the quadratic model provided a better-fit solution to the observed data. Ceylan et al. [23] developed linear, quadratic, and exponential HSA models to forecast Turkey's TED by 2025. Similarly, Kaveh et al. [24] used exponential and linear HSA models to estimate TED in Iran, based on population, GDP, and the data of numbers of vehicles. In addition, Korkmaz and Akgüngör [25] proposed a new optimization technique to predict Turkey's TED by 2030 using the flower pollination algorithm (FPA). Sonmez et al. [26] used the C-ABC algorithm to estimate the total energy demand in Turkey. The researchers used GDP, population, and vehicle-km data parameters and developed linear (C-ABCL), exponential (C-ABCE), and quadratic (C-ABCQ) models for estimating the TED by 2030 under two different scenarios. As a result, the research presented that ABC can be effectively used as an alternative model to predict future trends in TED.

In this study, a new adaptive selection method was developed and added to the artificial bee colony (ABC) algorithm. The developed A-ABC was used for Turkey's TED estimation problem. Therefore, we aimed to solve the TED optimization problem and find optimal results. For these purposes, three mathematical prediction models were used, respectively linear (A-ABCL), exponential (A-ABCE), and quadratic (A-ABCQ). Error metrics scores and complexity analysis were found, and future energy demand was estimated under two different scenarios to show the accuracy and efficiency of our suggested A-ABC algorithm. Also, our results were compared with the study results of Sonmez et al. [26].

3. Canonical artificial bee colony

Karaboga suggested the ABC algorithm in 2005 for optimization problems, which simulates the intelligent food search behavior of honeybee swarm-based optimization techniques [27]. The primary purpose of the algorithm is to try to find the best solution in search space (food source with the most nectar). The ABC algorithm consists of 4 basic steps:

Step 1 As follows, the SN number of candidate solutions $(x_1, x_2, ..., x_{SN} \in X)$ are produced uniformly at random.

$$x_{ij} = x_j^{min} + rand(0,1)(x_j^{max} - x_j^{min}) \; i = 1,2,...,SN \; and \; j = 1,...,D \tag{1}$$

Problem dimension is defined by D, and x_j^{min} is minimum bound, and x_j^{max} is the maximum bounds limit of the parameter. Moreover, according to Equation 2, the fitness value of each source of food $(fitness_i)$ is determined.

$$fitness_i = \begin{matrix} 1/(1+f_i) & , & f_i \geq 0 \\ 1+abs(f_i) & , & f_i < 0 \end{matrix} \tag{2}$$

Here, the objective function f_i shows the fitness value of source x_i. In addition, the $failure_i$ counter, which is a limit value for each food source, is defined and initialized to zero.

Step 2 Each worker is sent to the food source (x_i) to which the bee is responsible, and a try is made to look for a better food source (v_i) using Equation 3. If the quality of the new food resource is better, it is taken into memory as x_k. Here $\beta_{i,j}$ is a random number generated from a uniform distribution in [-1,1].

$$v_{i,j} = x_{i,j} + \beta_{i,j}(x_{i,j} - x_{k,j}) \tag{3}$$

If $v_{i,j}$ exceeds the parameter limits, it is shifted to the limit values of the j. parameter according to Equation 4. The fitness value of the newly found v_i is calculated. The choice is made between the new and the old solution using the greedy selection technique.

$$\begin{Bmatrix} x_j^{min} & v_{i,j} < x_j^{min} \\ v_{i,j} & x_j^{min} \leq v_{i,j} \leq x_j^{max} \\ x_j^{max} & v_{i,j} < x_j^{min} \end{Bmatrix} \tag{4}$$

Step 3 The probability of each solution x_i (p_i) is calculated by the ratio of the fitness value of the solution $(fitness_i)$ to the sum of the fitness values of all solutions.

$$p_i = \frac{fitness_i}{\sum_{j=1}^{SN} fitness_j} \tag{5}$$

According to Equation 5, as the fitness value increases, the probability of scout bees choosing that solution will increase. According to the probabilistic selection rule, a new solution is found using Equation 3 concerning the preferred solution, and the fitness value is calculated. If the fitness value of v_i is better than the fitness value of x_i, x_i the first one is replaced, and the $failure_i$ counter is reset.

Step 4 The $failure_i$ values of each solution x_i are checked. If $failure_i$ exceeds the limit value, it means that the solution x_i is ended. In this case, the employed bee turns into a scout bee form. Later, scout bees look for new solutions using Equation 1 instead of the solutions they gave up. The algorithm follows these stages until it reaches the end condition. Then the best solution of the algorithm is found.

4. Proposed adaptive artificial bee colony

Even though the diversification feature of the C-ABC algorithm is successful, it is known that the focusing behavior is not adequate [28]. The key explanation for this is that only one dimension is modified in the search equation. The search is conducted under the direction of a randomly chosen solution. In order to eliminate this problem, approaches have been suggested that modify several dimensions in the search equation [28–30] or use the current best-so-far solution in the search equation [31–33]. For all types of problems, both of these approaches may not work [34]. For instance, good results are achieved in multimodal problems when search equations modify multiple dimensions. Nevertheless, search equations that guide the best solution can be useful in unimodal problem types. Both approaches can also go to stagnation (trap in optimum local points) quickly related to the type of the problem.

While these two approaches seem to contribute to the performance of the algorithm, these two methods are probable to be stuck to the local minimum as the type and size of the problem changes [35]. An adaptive search equation is therefore required to make the most suitable selection based on the type of problem [28]. This article proposes a new ABC algorithm using an adaptive search equation to solve this problem. The pseudocode of the proposed A-ABC algorithm with adaptive selection capability is shown in Figure 2. The red (deleted) and blue (modified) areas in Figure 2 show the changes in the C-ABC that we modified.

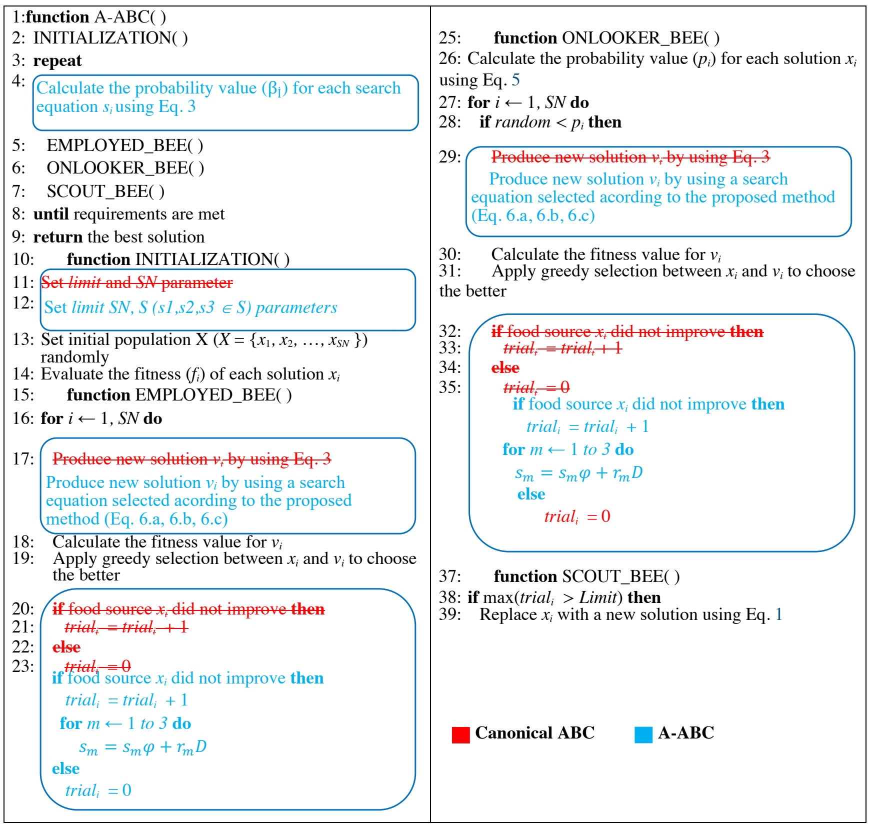

Figure 2. The differences between Canonical ABC (C-ABC) and A-ABC.

An adaptive search equation selection method in the A-ABC algorithm has been suggested rather than Equation 3 in the C-ABC algorithm. Using this method makes it possible to determine the degree to which a random solution or the best solution is used for deciding a solution. The described search equations in A-ABC are as follows:

$$x_{i,j} = x_{i,j} + rand(-c, +c)(x_{i,j} - x_{k,j}) \tag{6.a}$$

$$x_{i,j} = x_{i,j} + rand(-c, +c)(x_{i,j} - x_{best,j}) \tag{6.b}$$

$$x_{i,j} = x_{i,j} \tag{6.c}$$

The formulas state that the value of c is a positive real number and represents the scaling factor value. The magnitude of perturbation can be controlled by this parameter [29]. A probabilistic selection rule calculates the rate of utilization of these three solutions in the production of each candidate solution. Under this rule, the probability of choosing each of the search equations is determined using the following formula:

$$\beta_i = \frac{s_i}{\sum_{m=1}^{3} S_m} \tag{7}$$

Here s_i shows the success rate of each search equation. At the beginning of the algorithm, the success score is initially set at a value greater than zero for each search equation (this value was selected as 100 in this study). Whenever a new candidate solution (v_i) is better than the reference solution (x_i), the utilization rate of each equation used in this candidate solution is added to the success score as follows:

$$s_i = s_i\phi + r_i D \tag{8}$$

Where ϕ is the evaporation coefficient, r_i is the rate of use of the relevant equation, and D is the problem dimension. Although any search equation in the algorithm study process can produce good results at a particular time, it may not produce good results in later stages. Since the search equation selected with the ϕ coefficient gives unsuccessful results, the probability of selection is decreased by reducing the success rate.

According to the probabilistic selection rule and the roulette wheel technique, the search equation used in all dimensions of each candidate solution is chosen. The appropriate search equation and usage rate for a problem can be calculated adaptively with these techniques. In this proposed method, a candidate solution sample is shown in Figure 3.

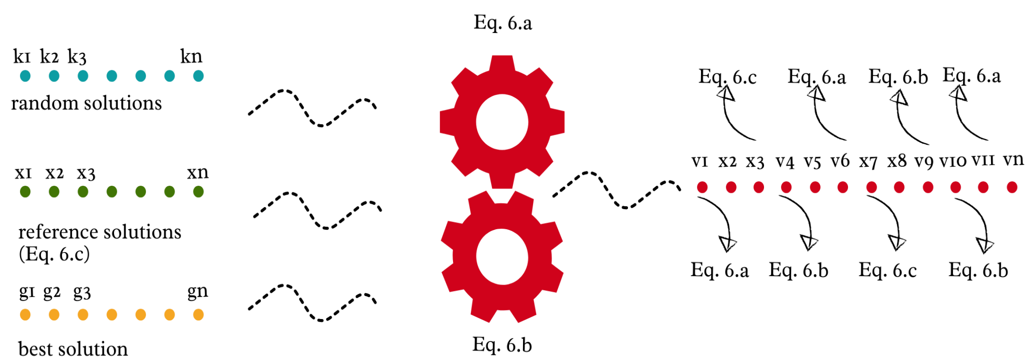

Figure 3. An example of a candidate solution generation by the proposed adaptive search equation.

5. Model development for TED with A-ABC

In the literature, the factors affecting the transportation energy demand of a country are GDP, population, and TKM data. These parameters were used in our model because they have a high correlation with the TED variables [11, 22, 25]. Sonmez et al. [26] proposed a solution to the TED problem using the C-ABC method. Therefore, data from Sonmez et al.'s [26] study were used in our study to present the performance of the A-ABC algorithm and to compare with the findings of C-ABC. Turkey's GDP, population, total annual vehicle-km, and observed TED data between 1970 and 2013 are given on the Kaggle[5]. All these data were collected from TSI[6], General Directorate of Highways[7], highway statistics, and literature [26].

[5]Kaggle Inc. (2022). Kaggle [online]. Website https://www.kaggle.com/safadorterler/datasets [accessed 7 Jan 2022].
[6]TSI (2022). Turkish Statistical Institute [online]. Website https://www.tuik.gov.tr [accessed 10 Jan 2022].
[7]GDH (2022). General Directorate of Highways [online]. Website https://www.kgm.gov.tr [accessed 10 Jan 2022].

5.1. Training and testing data set

Data (TED, GDP, population, TKM) from 1970 to 2013 was used to develop and test recommended forecast models. The data used for training and testing was shown with 0 and 1, respectively. Thirty-six years of data were used for training models, while eight years of data were used for testing models.

Turkey is in continuous development, and economic data has increased over the years. While the GDP in Turkey was 19.04 billion USD in 1970, it increased by 4000% to 800.5 billion USD in 2013. Due to the high difference in the distribution range of the data, the parameters were normalized as in Equation 9. It helped a more accurate comparison with the research of Sonmez et al.

$$X_{normalized} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{9}$$

5.2. Mathematical regression models and objective function

Three predictions models, linear, exponential, and quadratic functions, have been suggested to predict Turkey's TED.

$$E_{linear} = w_1 + w_2X_1 + w_3X_2 + w_4X_3 \tag{10}$$

$$E_{exponential} = w_1 + w_2e^{(X_1W_3)} + w_4e^{(X_2W_5)} + w_6e^{(X_3W_7)} \tag{11}$$

$$E_{quadratic} = w_1 + w_2X_1 + w_3X_2 + w_4X_3 + w_5X_1X_2 + w_6X_1X_3 + w_7X_2X_3 + w_8X_1^2 + w_9X_2^2 + w_{10}X_3^2 \tag{12}$$

The main purpose of TED forecasting is to find an optimum result by processing the available data. The values of X_1, X_2, and X_3 in the mathematical regression models (Equations 10–12) show GDP, population and TKM values, respectively. According to these values, weight values (w_i) are calculated to make the most appropriate TED forecast for the given years. The objective function used is given in Equation 13.

$$minf(v) = \sum\nolimits_{r=1}^{R} (E_r^{observed} - E_r^{estimated})^2, \tag{13}$$

where $E_r^{observed}$ and $E_r^{estimated}$ represent true and predicted values, while R holds the number of observations.

6. Experimental results

This section presents a series of experiments demonstrating the success of A-ABC. Turkey's TED models were developed using data observed between 1970 and 2013 by using an A-ABC algorithm. Appropriate values of A-ABC parameters were automatically set by irace, an offline parameter configuration tool [36]. Accordingly, the number of populations (SN) is 50, the limit value is 300, and the c parameter is 0.5. The data was tested 20 times, and the best results were obtained. Each run of A-ABC was stopped when the maximum number of iterations (40,000) was reached. The algorithm (A-ABC) was coded with JAVA programming language and run on a core i7, 2.40 GHz, 8 GB RAM computer. The performances of the models we proposed in the study were evaluated and interpreted with global error measurement metrics; later, the models were examined in terms of runtimes. In addition, the proposed A-ABC algorithm was compared with the C-ABC algorithm by performing complexity analysis. The proposed A-ABC algorithm determined the weight parameters of the mathematical regression models. The weight parameters are presented in Table 1.

Table 1. Coefficients of linear, exponential, and quadratic models.

	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}
Linear	0.018518	0.074928	0.235717	0.527270						
Exponential	0.407035	–0.160733	–14.030545	0.034310	2.991443	–0.278724	–0.290483			
Quadratic	–0.002376	0.043393	0.568335	0.005052	–0.101242	0.129145	0.090784	0.011216	–0.113127	0.269105

6.1. Global error measurement metrics and runtime analysis of the models

The performance of the models was evaluated by global error measurements. Among these measurements, absolute error (AE), sum of absolute error (SAE), root mean square error ($RMSE$), mean absolute error (MAE), mean squared error (MSE), the mean absolute range normalized error ($MARNE$), and standard deviation of absolute error (Std_AE) are scale-dependent measures, while the absolute fraction of variance (R^2), mean absolute percentage error ($MAPE$), and the standard deviation of percentage absolute error (Std_APE) are scale-independent error measures [37, 38].

$$AE = |E_{obs} - E_{pre}| \tag{14}$$

$$APE = |(E_{obs} - E_{pre})/E_{obs}| \tag{15}$$

$$RMSE = ((\sum_{i=1}^{n} AE^2)/n)^{0.5} \tag{16}$$

$$SAE = \sum_{i=1}^{n} AE \tag{17}$$

$$MAE = 1/n \sum_{i=1}^{n} AE \tag{18}$$

$$MAPE = 1/(n-1) \sum_{i=2}^{n} APE \tag{19}$$

$$MSE = 1/n \sum_{i=1}^{n} AE^2 \tag{20}$$

$$MARNE = (1/n \sum_{i=1}^{n} AE)/max(E_{obs}) \tag{21}$$

$$Std\_AE = ((\sum_{i=1}^{n} (AE_i - MAE)^2)/n)^{0.5} \tag{22}$$

$$Std\_APE = ((\sum_{i=2}^{n} (AE_i - MAE)^2)/n-1)^{0.5} \tag{23}$$

Here n is the number of data points; AE and APE are the absolute error and the absolute percentage error, respectively. E_{pre} and E_{obs} are the predicted and observed values, respectively. Global errors measurement scores for training and testing data are presented in Table 2 for each model of the two algorithms. The minimum error values obtained from each model for training and test data are shown in blue and red, respectively. Also, the algorithm that produces a better result for each global error from the compared algorithms is shown in bold.

When Table 2 was examined, the algorithm A-ABC that we suggested in our study had the lowest error measurements for both train and test data according to all global error measurement values. In

Table 2. Comparison results of TED models and algorithms.

model	Linear				Exponential				Quadratic			
	train		test		train		test		train		test	
	A-ABC	C-ABC	A-ABC	C-ABC	A-ABC	C-ABC	A-ABC	C-ABC	A-ABC	C-ABC	A-ABC	C-ABC
R^2	**0.9602**	0.9580	**0.9641**	0.9485	**0.9639**	0.9563	**0.9518**	0.9411	**0.9611**	0.9501	**0.9643**	0.9311
RMSE	**0.0464**	0.0476	**0.0477**	0.0572	**0.0442**	0.0486	**0.0553**	0.0612	**0.0458**	0.0519	**0.0476**	0.0661
SAE	**1.2150**	1.2386	**0.2946**	0.3572	**1.1154**	1.1378	**0.3328**	0.3719	**1.2136**	1.4747	**0.3216**	0.4760
MAE	**0.0337**	0.0343	**0.0367**	0.0446	**0.0310**	0.0319	**0.0416**	0.0465	**0.0337**	0.0409	**0.0402**	0.0595
MAPE	**0.1950**	0.2143	**0.0988**	0.1100	**0.1214**	0.1236	**0.1068**	0.1206	**0.1491**	0.2662	**0.1503**	0.1605
MSE	**0.0021**	0.0023	**0.0022**	0.0032	**0.0019**	0.0023	**0.0030**	0.0037	**0.0021**	0.0026	**0.0021**	0.0043
MARNE	**0.0014**	0.0015	**0.0016**	0.0019	**0.0012**	0.0014	**0.0018**	0.0020	**0.0014**	0.0017	**0.0017**	0.0026
Std_AE	**0.0318**	0.0334	**0.0304**	0.0382	**0.0314**	0.0371	**0.0365**	0.0425	**0.0310**	0.0324	**0.0255**	0.0308
Std_APE	**0.3713**	0.4436	**0.0492**	0.0605	**0.0906**	0.1049	**0.0629**	0.0633	**0.1189**	0.4859	0.0974	**0.0781**

■ Best train data score ■ Best test data score ■ Best result

the linear model, the A-ABCL model made markedly more successful predictions for both train and test data, producing lower error values than C-ABCL. Also, we can see in Table 2 that the lowest error values ($SAE, MAE, MAPE, MARNE$ and Std_APE) were produced by the A-ABCL model in the analysis made with the test data of all models. The A-ABCE model was more successful than the C-ABCE model by producing the lowest error values in all of the global error measurements made with both train and test data. The exponential model was the most suitable for all train data compared with other models, except for the Std_AE. A-ABCQ, which we proposed in our research, produced more successful results than C-ABCQ in both train and test results in all global measurement metrics except Std_APE. In addition, A-ABCQ produced the best training score among all models according to the Std_AE measurement. Besides, when the test data analysis was examined, it was seen that A-ABCQ produced the best test score according to $R^2, RMSE, MSE, STD\_AE$ measurements and was the best model for test data. As seen in Table 2, A-ABCL and A-ABCQ models have the best error scores for test data (shown with red color), while A-ABCE model has the lowest error scores for training data (shown with blue color). A-ABC was made more successful predictions by generating lower error scores for both test and train data. The A-ABCE model produced fewer errors, and this model was more compatible with training data.

When the measurements made with the test data were evaluated, we observed that A-ABCL and A-ABCQ models make more successful predictions and are the most compatible models for test data. The performance of TED estimation models for the train set is given in Figure 4a. According to Figure 4a, all models of A-ABC made similar predictions as observed TED. The performance of the TED estimate for test data is presented in Figure 4b. While A-ABCL and A-ABCQ had the lowest error values in the measurements made with test data, the A-ABCE model produced results close to other models. As a result, linear and quadratic models performed better than the exponential model and made predictions closer to observed TED in test data.

The models were compared in terms of runtime, and the results were given respectively. A-ABC linear model is 2848 ms, A-ABC exponential model is 3067 ms, A-ABC quadratic model is 3278 ms, C-ABC linear model is 1966 ms, C-ABC exponential model is 2148 ms, and the C-ABC quadratic model is 2308 ms. Quadratic models of both algorithms were slower than other models. Exponential and linear models, respectively, followed quadratic models. The C-ABC algorithm worked approximately 30% faster than A-ABC. However, the proposed A-ABC algorithm produced more accurate and efficient results. The analyses made support this situation.

In terms of average relative errors, the results of our model and different metaheuristics models in the literature are compared in Table 3. The proposed model (A-ABC) offered a more successful performance than other metaheuristic algorithms.

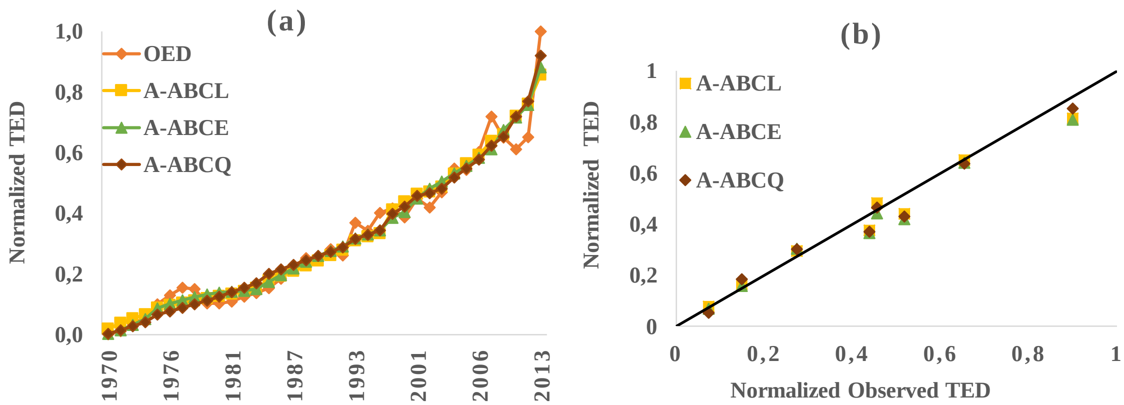

Figure 4. Performance of TED prediction models for the training (a) and test data set (b).

Table 3. Comparison of the literature and present study.

Reference	**Method**	**Forecasting Type/Country**	**Average Relative Errors (%)**
Present Study	A-ABC	Transport Energy/TURKEY	1.66
Sonmez et al. [26]	C-ABC	Transport Energy/TURKEY	2.06
Ceylan et al. [23]	Harmony Search	Transport Energy/TURKEY	13.41
Kaveh et al. [24]	Improved Harmony Search	Transport Energy/IRAN	3.32
Kaveh et al. [24]	Charged System Search	Transport Energy/IRAN	3.31
Amjadi et al. [39]	Particle Swarm Optimization	Electricity/IRAN	3.92
Canyurt and Ozturk [18]	Genetic Algorithm	Coal/TURKEY	3.22

6.2. Complexity analysis

The A-ABC algorithm was proposed to solve the TED problem instead of C-ABC in our study. The proposed A-ABC algorithm produced more accurate results than the C-ABC algorithm. The complexity of the proposed A-ABC algorithm is presented below. When calculating the algorithm's time complexity, the computational cost at each step of the algorithm at each iteration is considered asymptotically. These computational costs of each step are the cost of employed bees step (T_e), onlooker bees step (T_o), scout bee step (T_s), selecting search equation components step (T_c). Also, the equation pool size is presented as "es". The non-local computational complexity of the A-ABC algorithm is given in Equation 24.

$$\begin{aligned} T_{ABC} &= (T_e + T_o + T_s + T_c)\, x\, (itr)_{max} \\ &= (O(SN_{max}\, x\, D) + O(SN_{max}\, x\, D)) \\ &+ O(es\, x\, log_{es} + O(D))\, x\, itr_{max} \\ &= O(SN_{max}\, x\, D)\, x\, itr_{max} \end{aligned} \tag{24}$$

Also, the complexity of A-ABC for the N-element vector is $O(SN_{max}\, x\, D)\, x\, itr_{max}$, while the complexity of C-ABC for the same size vector is also $O(SN_{max}\, x\, D)\, x\, itr_{max}$. Thus, the time complexity of C-ABC for the proposed approach can be noted as $O(SN_{max}\, x\, D)\, x\, itr_{max}$. Therefore, the modification made in the proposed algorithm does not create extra complexity. The convergence graph of A-ABC and C-ABC is given in Figure 5.

Figure 5 shows that A-ABC converges faster and has a greater convergence rate than C-ABC. Table 4 shows the statistical results of the best fitness, worst fitness, average fitness, standard deviation, and mean runtime (ms). The values that show the best outcome are shown in bold.

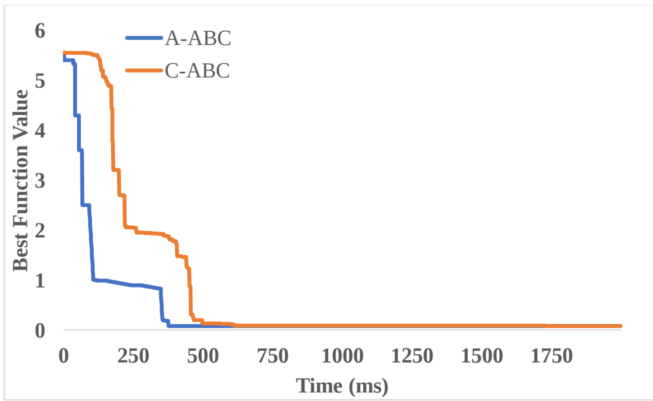

Figure 5. A-ABC and C-ABC convergence graph.

Table 4. Statistical results of the experimental study.

	A-ABC	C-ABC
Best fitness	**0.0967**	0.0971
Wors fitness	**5.5543**	5.5673
Average fitness	**0.1064**	0.1398
Standard deviation	**0.0337**	0.0470
Mean runtime (ms)	3278	**2308**

A-ABC gave better results than C-ABC in terms of the best fitness and worst fitness values. The average fitness and standard deviation metrics are important indicators to show the consistency of the algorithm. According to the statistical results of the experimental study, A-ABC produced more successful results. Similarly, the standard deviation value of A-ABC gave better results than C-ABC. When A-ABC was evaluated in terms of mean runtime, it was seen that A-ABC was behind C-ABC. However, when evaluated in terms of convergence time, A-ABC reached the best result faster (as shown in Figure 5). As a result, the suggested A-ABC algorithm is more successful and consistent than the C-ABC algorithm.

7. Scenarios of Turkey's TED

In this stage of the study, two different scenarios were used to test the A-ABC algorithm and predict Turkey's TED in 2014–2034. Also, the A-ABC was compared to Sonmez et al.'s [26] C-ABC models. The input parameters (GDP, population, and TKM) used in scenario 1 were obtained using the S-curve technique.

Scenario 1. To create the input parameters of this scenario, we used the third-order polynomial S-Curve Fitting technique [26]. The functions used for GDP, population, and TKM estimation are given below.

$$GDP_{est}(y) = 0.02427x^3 - 0.9327x^2 + 14.54x - 7.286 \tag{25}$$

$$Population_{est}(y) = -3.904\ 10^{-5}x^3 + 0.0006804x^2 + 0.9889x + 34.14 \tag{26}$$

$$TKM_{est}(y) = 0.0005816x^3 + 0.007741x^2 + 0.4939x + 8.342 \tag{27}$$

TED predictions were made with the parameters calculated according to the S-Curve technique. The A-ABCL model made approximately 25% higher predictions than the A-ABCE model; on the other hand, the A-ABCQ model made higher predictions than the other two models. Our recommended algorithm gave successful results as seen in Figures 4a and 4b according to observed transportation energy demand data. We also proved this situation statistically by global error metrics. Comparison of the energy demand estimation results of the proposed methods is given in Figure 6 for scenario 1.

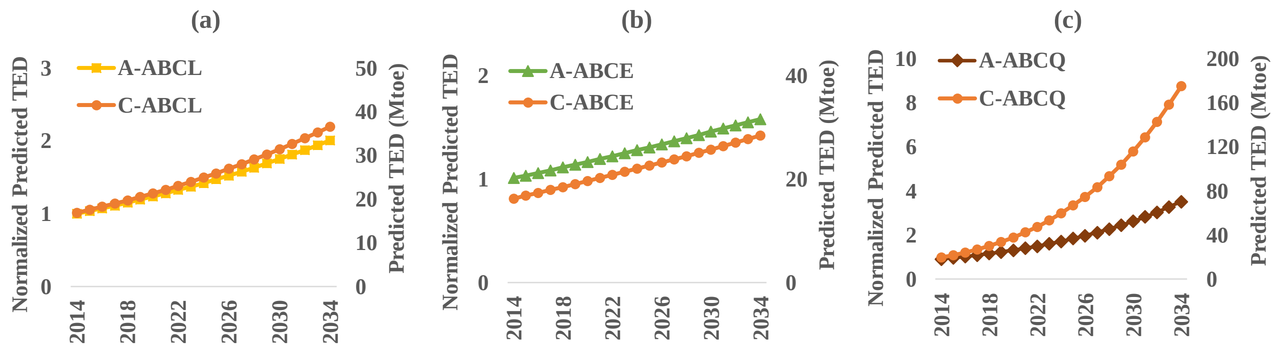

Figure 6. Future projections of TED according to scenario 1: (a) linear, (b) exponential, (c) quadratic

.

As shown in Figure 6a, the A-ABCL model predicted that in 2034 the TED value would grow by about 77% and would be 40.1 Mtoe. C-ABCL predicted a higher energy demand estimation value than A-ABCL. In Figure 6b, exponential models of the compared algorithms were lower than the linear model estimates. The A-ABCE model estimated that by 2034, the TED value would grow by approximately 40% to 31.6 Mtoe. C-ABCE model predicted about 1% growth by 2021. On the other hand, A-ABCE, which predicted higher energy demand, made more realistic predictions for 2021, assuming growth of approximately 20%. In Figure 6c, the A-ABCQ model predicted that the TED value in 2034 would grow by approximately 3.5 times from 2013 to 2034 and reach 70.5 Mtoe. However, the C-ABCQ model predicted growth of approximately 7.5 times and made much higher predictions than the A-ABCQ model. Therefore, it was seen that the A-ABCQ gave more successful results than the C-ABCQ, and all models of A-ABC were suitable for scenario 1.

Scenario 2. The input parameters used in this scenario were obtained from the literature [26]. Moreover, GDP, population, and TKM projections were confirmed by looking at the Organization for Economic Co-operation and Development (OECD) report [40], TSI[8] official projections, and a research report on traffic growth [41], respectively. The data used in scenario 2 are given in Table 5.

[8]TSI (2022). Turkish Statistical Institute [online]. Website https://www.tuik.gov.tr [accessed 10 Jan 2022].

When Table 5 is examined, the approximate values of GDP, population, and TKM in 2034 were found to be 1.24 trillion dollars, 90 million, and 274 billion, respectively. For scenario 2, linear, exponential, and quadratic models of A-ABC and C-ABC algorithms were developed by using the parameter values in Table 5. The A-ABCL model had higher prediction values than the A-ABCE model, while the A-ABCQ model had higher prediction values than the other two models. Comparison of the energy demand estimation results of the proposed methods is given in Figure 7 for scenario 2.

Table 5. Data used for scenario 2.

Year	GDP ($10^9)	Population (10^6)	TKM ($10^9)	Year	GDP ($10^9)	Population (10^6)	TKM ($10^9)
2014	814.80	77.32	116.08	2025	991.40	85.57	200.99
2015	829.40	78.15	122.93	2026	1009.30	86.18	206.81
2016	844.40	78.97	129.78	2027	1027.40	86.78	212.62
2017	859.50	79.77	136.63	2028	1045.90	87.35	223.58
2018	875.00	80.55	141.60	2029	1064.70	87.90	232.84
2019	890.80	81.32	151.08	2030	1083.90	88.43	240.42
2020	906.80	82.08	159.84	2031	1222.90	88.93	248.00
2021	923.10	82.82	168.59	2032	1163.40	89.41	257.50
2022	939.70	83.54	176.27	2033	1205.20	89.86	266.00
2023	956.70	84.25	183.94	2034	1248.60	90.28	274.30
2024	973.90	84.94	192.47				

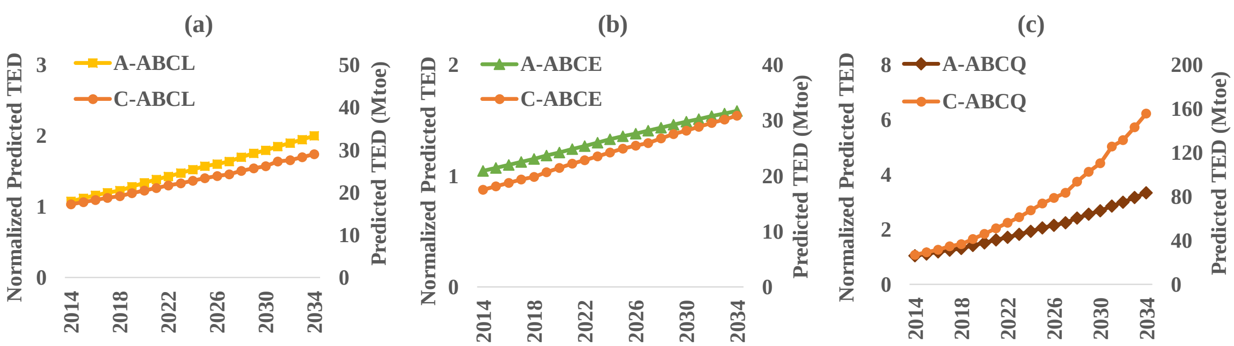

Figure 7. Future projections of TED according to scenario 2. (a) Linear (b) Exponential (c) Quadratic.

When Figure 7a (linear model estimation graph) was examined, the A-ABCL model predicted that the TED value would grow by approximately 75% in 2034, reaching 40 Mtoe. Also, unlike scenario 1, A-ABCL gave higher energy demand estimates than C-ABCL. When Figure 7b (exponential model estimation graph) was examined, the A-ABCE model predicted that the TED value would grow by approximately 38% in 2034 and reach 31.5 Mtoe. In the first years, the C-ABCE model predicted lower prediction values than A-ABCE. However, in the following years, they made closer predictions. When Figure 7c (quadratic model estimation graph) was examined, similar to scenario 1, the A-ABCQ model predicted that the TED value would grow approximately three times from 2013 to 2034 and reach 66.5 Mtoe. The C-ABCQ model predicted growth of approximately 5.5 times, making much larger predictions than the A-ABCQ model. Therefore, the A-ABCQ model, which produced the lowest errors in the global error measurement functions made with test data, gave more successful results than the C-ABCQ and all models of A-ABC were suitable for scenario 2.

8. Conclusion

Energy consumption in the transportation sector is equal to around 1/5 of the total energy consumption of Turkey. However, due to the increase in energy demand depending on the growth rate of Turkey, prediction of the transportation energy demand is essential for the efficient management of energy policies. In addition, it is necessary for the efficient use of fossil energy resources to achieve economic efficiency and an eco-friendly structure. The developed A-ABC algorithm has overcome the problem of being stuck at the local optimum points in the C-ABC algorithm by making the most appropriate choice according to the type of problem thanks to the adaptive search equation. With the developed algorithm, we suggested an alternative solution to the transportation energy demand problem. Also, we proved that the A-ABC algorithm is more successful than the classical ABC algorithm in terms of performance.

A new adaptive artificial bee colony (A-ABC) algorithm that can adaptively select an appropriate search equation was suggested in this study more efficiently estimate Turkey's TED. Also, in our study, the A-ABC algorithm and canonical artificial bee colony (C-ABC) performances were compared, and the A-ABC algorithm that we developed was found more successful. GDP, population, and TKM, which are Turkey's economic indicators, were used as input parameters. In this study, three mathematical models were developed as linear (A-ABCL), exponential (A-ABCE), and quadratic (A-ABCQ). In the research, experimental studies were carried out with a curve fitting technique on future predictions of TED under two different scenarios. The proposed A-ABCL, A-ABCE, A-ABCQ models provided better solutions than the studies in the literature for the TED problem. Besides, lower error values were obtained from proposed models with global error metrics. The A-ABCL and A-ABCE models produced significantly better predictions for both train and test data, producing lower error values than C-ABCL and C-ABCE. The A-ABCQ model produced better results than C-ABCQ for both train and test results except for the Std_APE metric. In addition, A-ABCQ produced the best training score among all models according to the Std_AE measurement. Besides, when the analysis with test data was examined, A-ABCQ produced the best test score according to R^2, $RMSE$, MSE, STD_AE measurements.

In the first scenario, the results of linear, exponential, and quadratic models according to 2034 TED estimates were found to be 40.1, 31.6, and 70.5 Mtoe, respectively. In the second scenario, the results of linear, exponential, and quadratic models according to the TED estimates for 2034 were found to be 40.0, 31.5, and 66.5 Mtoe, respectively. According to these results, consumption estimation values of A-ABCL and A-ABCE were relatively lower than A-ABCQ. In both scenarios we presented in our study, the A-ABCQ model predicted that the TED value would increase approximately three-fold from 2013 to 2034. The C-ABCQ model, on the other hand, predicted that the TED value would increase by approximately 5.5 times, unlike the A-ABCQ model.

When the results were evaluated, it can be concluded that the developed model produced more accurate values than the C-ABC model. While the estimation results of the linear and exponential models of both C-ABC and A-ABC algorithms were closer to each other, the significant difference in the quadratic model estimation results supports that the C-ABCQ was more realistic. The developed algorithm produced successful results according to the observed transportation energy demand data as shown in Figures 4a and 4b. Therefore, it was seen that the consumption estimations we offer in two different scenarios gave successful results. This situation was also statistically proven by global error measurements.

As a result of this study, a new adaptive artificial bee colony algorithm that can adaptively select an appropriate search equation was suggested to estimate TED of Turkey more efficiently. Also, the A-ABC algorithm and C-ABC were compared according to performances, and our developed algorithm (A-ABC) gave better solutions between 2014 and 2034 according to global error metrics. We expect the results we presented in the study to contribute to researchers and politicians working on energy planning and strategies in the field of transportation. In future studies, the modeling can be expanded and evaluated more comprehensively by increasing the number of input parameters. Using different algorithms such as gravity search algorithm, crow search algorithm, differential search algorithm, artificial fish-swarm, and wolf colony algorithm, the performances of the algorithms on TED can be compared in future studies. In addition, TED estimates of different countries can be made with the proposed A-ABC algorithm and models.

References

[1] Nuray A, Manap GR, Sarioglu K, Sanalan AT, Aydin S. Environmental Indicators. Ankara, Turkey: Bee Content & Communication, 2020.

[2] Erat S, Telli A, Ozkendir OM, Demir B. Turkey's energy transition from fossil-based to renewable up to 2030: milestones, challenges and opportunities. Clean Technologies and Environmental Policy 2021; 23 (2): 401-412.

[3] Republic of Turkey Energy Market Regulatory Authority. Oil and Natural Gas Sector Report 2020. Ankara,Turkey: Energy Market Regulatory Authority Publications,2021.

[4] Ağbulut Ü. Turkey's electricity generation problem and nuclear energy policy. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects 2019; 41 (18): 2281-2298.

[5] Kaygusuz K, Guney MS, Kaygusuz O. Renewable energy for rural development in Turkey. Journal of Engineering Research and Applied Science 2019; 8 (1): 1109-18.

[6] Karataş Z. Social Impacts of COVID-19 Pandemic, Change and Empowerment. Türkiye Sosyal Hizmet Araştırmaları Dergisi 2020; 4 (1): 3-17.

[7] Mousavi MH, Ghavidel S. Structural time series model for energy demand in Iran's transportation sector. Case Studies on Transport Policy 2019; 7 (2): 423-432.

[8] Republic of Turkey Energy Market Regulatory Authority. Turkish petroleum market report 2019. Ankara,Turkey: Energy Market Regulatory Authority Publications,2020.

[9] Debnath KB, Mourshed M. Forecasting methods in energy planning models. Renewable and Sustainable Energy Reviews 2018; 88 : 297-325.

[10] Wei N, Li C, Peng X, Zeng F, Lu X. Conventional models and artificial intelligence-based models for energy consumption forecasting: A review. Journal of Petroleum Science and Engineering 2019; 181 : 106187.

[11] Codur MY, Ünal A. An estimation of transport energy demand in Turkey via artificial neural networks. Promet-Traffic&Transportation 2019; 31 (2): 151-161.

[12] Purwanto P, Eswaran C. Enhanced hybrid prediction models for time series prediction. International Arab Journal of Information Technology (IAJIT) 2018; 15 (5): 866-874.

[13] Amiri SS, Mostafavi N, Lee ER, Hoque S. Machine learning approaches for predicting household transportation energy use. City and Environment Interactions 2020; 7 : 100044.

[14] Di Leo S, Caramuta P, Curci P, Cosmi C. Regression analysis for energy demand projection: An application to TIMES-Basilicata and TIMES-Italy energy models. Energy 2020; 196 : 117058.

[15] Chai J, Lu QY, Wang SY, Lai KK. Analysis of road transportation energy consumption demand in China. Transportation Research Part D: Transport and Environment 2016; 48 : 112-124.

[16] Sahraei MA, Çodur MK. Prediction of transportation energy demand by novel hybrid meta-heuristic ANN. Energy 2022; 123735.

[17] Ray P, Panda SK, Mishra DP. Short-term load forecasting using genetic algorithm. In Computational Intelligence in Data Mining 2019; 863-872.

[18] Canyurt OE, Ozturk HK. Application of genetic algorithm (GA) technique on demand estimation of fossil fuels in Turkey. Energy Policy 2008; 36 (7): 2562-9.

[19] Anand A, Suganthi L. Hybrid GA-PSO optimization of artificial neural network for forecasting electricity demand. Energies 2018; 11 (4): 728.

[20] Bilgili M, Arslan N, Şekertekin A, Yaşar A. Application of long short-term memory (LSTM) neural network based on deep learning for electricity energy consumption forecasting. Turkish Journal of Electrical Engineering & Computer Sciences 2022; 30 (1): 140-157.

[21] Sahraei MA, Duman H, Çodur MY, Eyduran E. Prediction of transportation energy demand: multivariate adaptive regression splines. Energy 2021; 224: 120090.

[22] Toksari MD. A Hybrid Algorithm for Forecasting Transportation Energy Demand of Turkey. Available at SSRN 2022; 4009459.

[23] Ceylan H, Ceylan H, Haldenbilen S, Baskan O. Transport energy modeling with meta-heuristic harmony search algorithm, an application to Turkey. Energy Policy 2008; 36 (7): 2527-2535.

[24] Kaveh A, Shamsapour N, Sheikholeslami R, Mashhadian M. Forecasting transport energy demand in Iran using meta-heuristic algorithms. International Journal of Optimization in Civil Engineering 2012; 2 (4): 533-544.

[25] Korkmaz E, Akgüngör AP. Flower pollination algorithm approach for the transportation energy demand estimation in Turkey: model development and application. Energy Sources, Part B: Economics, Planning, and Policy 2018; 13 (11-12): 429-447.

[26] Sonmez M, Akgüngör AP, Bektaş S. Estimating transportation energy demand in Turkey using the artificial bee colony algorithm. Energy 2017; 122 : 301-310.

[27] Karaboga D. An Idea Based on Honey Bee Swarm for Numerical Optimization. Technical report-tr06, Erciyes University, ,Engineering Faculty, Computer Engineering; 2005.

[28] Yavuz G, Aydın D. Improved Self-adaptive Search Equation-based Artificial Bee Colony Algorithm with competitive local search strategy. Swarm and Evolutionary Computation 2019; 51 : 100582.

[29] Akay B, Karaboga D. A modified artificial bee colony algorithm for real-parameter optimization. Information Sciences 2012; 192 : 120-142.

[30] Brajevic I, Tuba M. An upgraded artificial bee colony (ABC) algorithm for constrained optimization problems. Journal of Intelligent Manufacturing 2013; 24 (4): 729-740.

[31] Gao W, Liu S, Huang L. A global best artificial bee colony algorithm for global optimization. Journal of Computational and Applied Mathematics 2012; 236 (11): 2741-2753.

[32] Bansal JC, Joshi SK, Sharma H. Modified global best artificial bee colony for constrained optimization problems. Computers & Electrical Engineering 2018; 67 : 365-382.

[33] Xue Y, Jiang J, Zhao B, Ma T. A self-adaptive artificial bee colony algorithm based on global best for global optimization. Soft Computing 2018; 22 (9): 2935-2952.

[34] Aydın D, Yavuz G, Stützle T. ABC-X: a generalized, automatically configurable artificial bee colony framework. Swarm Intelligence 2017; 11 (1): 1-38.

[35] Aydın D. Composite artificial bee colony algorithms: From component-based analysis to high-performing algorithms. Applied Soft Computing 2015; 32 : 266-285.

[36] López-Ibáñez M, Dubois-Lacoste J, Cáceres LP, Birattari M, Stützle T. The irace package: Iterated racing for automatic algorithm configuration. Operations Research Perspectives 2016; 3:43-58.

[37] Koutsandreas D, Spiliotis E, Petropoulos F, Assimakopoulos V. On the selection of forecasting accuracy measures. Journal of the Operational Research Society 2021; 1-18.

[38] Dörterler M, Faruk Bay Ö. Neural network based vehicular location prediction model for cooperative active safety systems. Promet-Traffic&Transportation 2018; 30 (2): 205-215.

[39] Amjadi MH, Nezamabadi-pour H, Farsangi MM. Estimation of electricity demand of Iran using two heuristic algorithms. Energy Conversion and Management 2010; 51 (3): 493-7.

[40] Johansson A, Guillemette Y, Murtin F, Turner D, Nicoletti G et al. Looking to 2060: A Global Vision of Long-Term Growth. Paris,France: OECD Publications,2012.

[41] BITRE. Traffic Growth: Modelling a Global Phenomenon. Canberra,Australia: Bureau of Infrastructure, Transport and Regional Economics, 2012.