

# Determination of industrial energy demand in Turkey using MLR, ANFIS and PSO-ANFIS

Didem Guleryuz<sup>1</sup>

<sup>1</sup>Department of Industrial Engineering, Bayburt University, Bayburt, Turkey  
Email: [dguleryuz@bayburt.edu.tr](mailto:dguleryuz@bayburt.edu.tr)

\*Corresponding Author: Didem Guleryuz, Email: [dguleryuz@bayburt.edu.tr](mailto:dguleryuz@bayburt.edu.tr)

**How to cite this paper:** Didem Guleryuz (2021). Determination of industrial energy demand in Turkey using MLR, ANFIS and PSO-ANFIS. Journal of Artificial Intelligence and Systems, 3, 16–34. <https://doi.org/10.33969/AIS.2021.31002>

**Received:** October 12, 2020  
**Accepted:** December 30, 2020  
**Published:** January 14, 2021

Copyright © 2021 by author(s) and Institute of Electronics and Computer. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



## Abstract

Energy is one of the most critical inputs in social and economic development, is an essential factor in increasing living standards and creating sustainable development. Since energy is an indispensable input in all sectors, energy dependence contributed to the necessity of countries' energy policy. It has critical importance to predict energy demand to determine energy policies. According to Turkey's annual energy consumption, the industry sector has consumed the most energy in the last five years. The prediction and analysis of energy demand with economic data in the industrial sector is an essential indicator of the economic development relationship between energy demand and industry. In this study, Multiple Linear Regression (MLR), Adaptive Neuro-Fuzzy Inference System (ANFIS), and optimized ANFIS with Particle Swarm Optimization (PSO) methods are employed to forecast energy demand for Turkish industrial sectors. The indicators which affect energy consumption were determined to estimate the energy demand. The 30-year dataset between 1990 and 2019 was split as training and test set. MLR, ANFIS and PSO-ANFIS were compared according to performance evaluation, and the most proper model was identified. The coefficient of determination ( $R^2$ ) for PSO-ANFIS, MLR, and ANFIS models are 0.9951, 0.9889, and 0.9932 in the training stage, and 0.9423, 0.9181, and 0.8776 in the testing stage, respectively. The study results indicated that the PSO-ANFIS model showed superior prediction capability with the least estimation error than MLR and ANFIS models. Consequently, parameters tuned PSO-ANFIS is able to predict the industrial energy demand in Turkey with high accuracy.

## Keywords

Industrial Energy Demand, Forecasting, ANFIS, PSO-ANFIS, MLR.

## 1. Introduction

After the industrial revolution, world energy need has increased with new

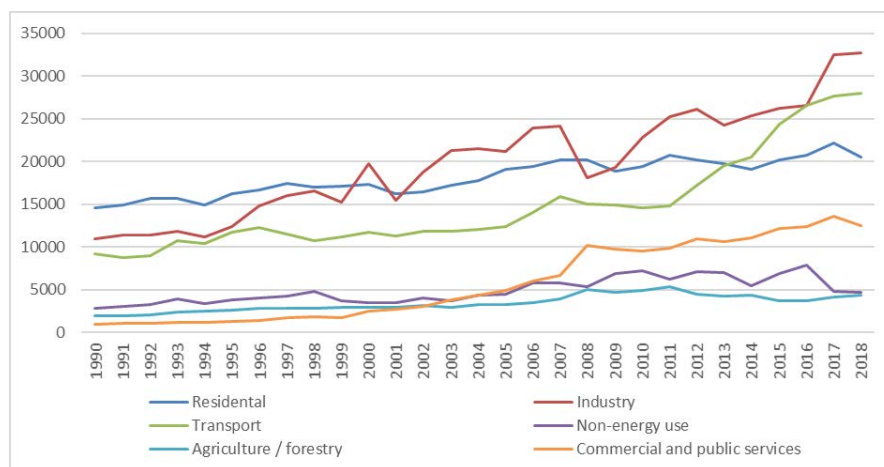
inventions and applying them to production processes. As previously ignored in an economy's production function, energy consumption is now considered crucial in the production process. The rapid increase in the population after the industrial revolution and the struggle to increase the countries' economic welfare also pioneered establishing systems that will ensure long-term energy consumption. With the established systems', economies have increased their competitive power in international markets thanks to the uninterrupted energy cost.

In this process, while the factors that restrict human movements (epidemics, natural disasters, wars, etc.) cause contractions in production and trade, examining the price and consumption of energy inputs will allow correct interpretations to understand and follow this contraction.

In the era of the industrial revolution, on the one hand, the continuous need for energy and the systems which meet these needs provided the development of international trade and the welfare of the countries; on the other hand, the emissions resulting from these energy sources constituted the bill of this system. As a result of these emissions, the world has warmed up rapidly, and the warming has turned into a state called global warming, damaging the climates and ecosystem. Countries have agreed on common international goals to reduce the global temperature rise in the context of global warming [1]. To achieve this goal, it has become necessary to peruse the effects of policies promoting energy-saving and efficiency on national GDP and economic growth.

Therefore, the variable affecting demand for analyzing the energy consumption in Turkey must determine multi-dimensional. Even though many of its energy needs have to be imported from abroad, Turkey has undertaken a significant bridge and exit point to move oil and gas from Central Asia to world markets. However, Turkey constitutes a burden to the economy's growth in the country's industrial energy consumption, energy consumption due to exports of these products formed to contribute to Turkey's economy positively.

While innovative breakthroughs in the industry caused by research and development activities can increase the country's exports, it may decrease the required energy consumption. This multi-dimensional review constitutes an energy economics literature. In energy consumption by sector in Turkey, it has been shown to occur with a large share of the industrial sector. Therefore, this study focused on energy needs in the industrial sector. Fig. 1 shows the sectoral energy consumption between 1990 and 2018 in Turkey.



**Figure 1.** Total final energy consumption by sector, Turkey 1990-2018 [2].

With this study, the links between industrial energy demand and other indicators are analyzed using ANFIS, MLR, and PSO-ANFIS approaches, then a novel estimation model is developed with optimized parameters. In previous studies, these relationships are generally examined by econometric analysis.

There is extensive literature on predicting energy demand in industrial sectors. When the literature on energy demand forecasting is examined, mostly classical regression methods are used, and forecasting studies have been made in the light of indicators affecting energy use in the industrial sector. Today, new estimation methods are developed using fuzzy systems and machine learning methods that estimate energy demand.

In previous studies, various approaches have been used to model industrial energy demand. These methods are the ARIMA method, Artificial Intelligence Based Prediction Models, Machine Learning (ML) based methods, Econometric models, and simulation models [3–6].

Some of the studies conducted to investigate the energy demand using different methods are shown in Table 1.

**Table 1.** Studies related to Energy Estimation.

| Reference            | Year | Application                 | Methods                                    |
|----------------------|------|-----------------------------|--|
| Barak and Sadegh [4] | 2016 | Energy Consumption          | ARIMA <sup>a</sup> -ANFIS hybrid algorithm |
| Yayar et al. [8]     | 2011 | Electric Energy Consumption | ANFIS and ARIMA                            |
| Akay and Atak [9]    | 2007 | Electricity Demand          | Grey Prediction                            |
| Demirci [10]         | 2020 | Natural Gas Demand          | ANFIS                                      |
| Yang et al. [11]     | 2016 | Electricity Demand          | ANFIS                                      |

|                             |      |                                 |   |
|-----------------------------|------|---------------------------------|---|
| Ahmad et al. [12]           | 2014 | Energy Demand                   | ANN-SVM <sup>b</sup>                    |
| Vinagre et al. [13]         | 2017 | Electrical Energy Consumption   | SVM                                     |
| Kasule and Ayan [14]        | 2020 | Electricity Consumption         | PSO-ANFIS, GA-ANFIS                     |
| Bayramoglu et al. [3]       | 2017 | Energy Demand                   | ANFIS                                   |
| Aghay et al. [15]           | 2017 | Electrical Energy Consumption   | Gene Expression Programming             |
| Huang et al. [16]           | 2019 | Renewable Energy Consumption    | Gaussian Process Regression             |
| Beyca et al. [17]           | 2019 | Natural Gas Consumption         | MLR, ANN <sup>c</sup> , SVR             |
| Liu et al. [18]             | 2016 | Energy Consumption              | Grey Prediction                         |
| Alcaraz and Villalvazo [19] | 2017 | Natural Gas Shortages           | Econometric Analysis                    |
| Sahaikh and Ji [20]         | 2016 | Natural Gas Demand              | Logistic Modelling Analysis             |
| Naji et al. [21]            | 2016 | Energy Consumption in buildings | ANFIS, SVM                              |
| Jawad et al. [5]            | 2020 | Electricity Load Forecasting    | SVM, KNN <sup>d</sup> , RF <sup>e</sup> |

a. Autoregressive Integrated Moving Average, b. Support Vector Machines

c. Artificial Neural Network, d. K-Nearest Neighbors, e. Random Forest

Yayar et al. [8] predicted the demand of the electrical energy in Tokat via ANFIS and ARIMA methods. Various prediction experiments have been conducted using the model in which electrical energy consumption is the dependent variable. For the first half of 2011, the electricity demand was predicted using ANFIS and ARIMA approaches. The developed models were evaluated by considering the performance criteria to show these techniques' prediction accuracy. The study results showed that ANFIS had better results than ARIMA in estimating electricity consumption for the data set [8].

Bayramoglu et al. [3] used ANFIS to estimate primary energy demand between 2016-2030 in Turkey. Various indicators are used, such as growth rate, population, and energy prices as an independent variable. The results show that, Turkey's energy demand has increased as in other OECD countries [3].

Kasule and Ayan [14] developed an ANFIS model to predict Uganda's electricity consumption. The parameter optimization of the ANFIS method has been done by the PSO algorithm and genetic algorithm (GA). As a result of the study, GA-ANFIS showed the best prediction performance compared to PSO-ANFIS and MLR [14].

Yang et al. [11] developed a novel forecasting model. Future data were

predicted using half an hour of electrical power data. Thanks to the improved model, seasonal inconveniences are eliminated. It was observed that the hybrid methods' results were preferable than the results obtained by using other methods individually [11].

Azadeh et al. [22] proposed an ANFIS for natural gas demand prediction, which is vital to decision-making processes in the energy sector. Before the ANFIS method was applied, the data were normalized to ensure standardization between the data. As a result of the study, it has been observed that the ANFIS method has better prediction results than traditional ANN and time series methods for the natural gas demand data of Iran [22].

Beyca et al. [17] forecasted Istanbul's natural gas consumption and used three machine learning methods: ANN, MLR, and Support Vector Regression (SVR). According to the developed model outputs, the polynomial cubic kernel function integrated SVR provided better results from the ANN technique for Istanbul's natural gas consumption [17].

Hamzacebi [23] proposed an ANN to forecast net electricity energy demand in Turkey by sectors until 2020. ANN can predict the output by using more than one input variable and model nonlinear relationships in the data structure. Therefore, the ANN model was used. As a result of the study, the highest net electricity consumption in transportation sector, housing sector, industrial sector and agricultural sector respectively [23].

In addition, some studies that have been performed ANFIS in the literature with optimization methods are listed below.

Moayedi et al. [24] have been optimized by the ANFIS with GA and PSO methods to calculate the friction capacity ratio ( $\alpha$ ) on shafts. The data set was gained from the production line for training phase. As a result of the study, it was seen that GA - ANFIS model has higher reliability to calculate the friction capacity ratio on shafts [24].

Basser et al. [25] aimed to identify the most appropriate protective spur dike parameters to reduce abrasion depth around the main spur sets. A new hybrid approach has been developed in the prediction phase that combines the ANFIS and PSO (ANFIS-PSO) to estimate the optimum parameters. As stated in the paper, the recommended hybrid method was more effective than ANFIS individually [25].

Shamshirband et al. [26] suggested an integrated model for accurate estimation of mercury released from fossil fuel power plants to assess environmental pollution and reduce health hazards. The mercury in the power plant boilers was estimated employing the PSO-ANFIS. The input variables of the model are coal properties and boiler parameters. The training data set was obtained from eighty-two points in

the factory. As a result of the study, it was seen that the optimized ANFIS method has better estimate results [26].

Numerous studies have presented that both ANFIS and PSO-ANFIS are precious methods for predicting engineering applications [24,25,27,28]. However, PSO optimized ANFIS model was not used in the forecasting of the energy economy. This study's difference from other studies is that the ANFIS parameters are optimized with the PSO algorithm; therefore, parameter optimization provides more appropriate predictive models. In this study, five indicators affecting the industrial energy demand were determined. Previous studies conducted to determine the industrial energy demand, gross domestic product per capita, and production capacity were independent variables, and energy demand was dependent variables [19,29,30]. Five different indicators that affect energy demand, different from other studies, were used as inputs. These indicators are gross domestic product per capita (constant 2010 billion US \$), manufactures exports (% of merchandise exports), merchandise exports (Current US\$), CO<sub>2</sub> emissions from manufacturing (Mt CO<sub>2</sub>), research and development expenditure (% of GDP). In particular, the addition of research and development expenditure and CO<sub>2</sub> emissions from manufacturing in input parameters will allow it to be seen in the efficiency between sustainability and energy demand. Also, using the hybrid PSO-ANFIS approach to determine the industrial energy demand will contribute to the literature.

The following parts of the paper are the description of the data collection process, theoretical backgrounds of MLR ANFIS and PSO-ANFIS methods, applications of the developed models, determining parameters of PSO, results and discussion for the case study.

## **2. Material and Method**

### **2.1. Data Collection Process**

In this study, energy demand in the industrial sector (ED) in Turkey between 1990–2019 is predicted using gross domestic product per capita (GDP), manufactures exports (ME), merchandise exports (MCE), CO<sub>2</sub> emissions from manufacturing (CO<sub>2</sub>) and research and development expenditure (RDE) data. All data were obtained from various data sources between 1990 and 2019. Table 2 shows the definition and source of the variables.

ED was stated as the output variable while GDP, ME, MCE, CO<sub>2</sub>, and RDE variables were defined as inputs supposed to significantly impact ED. Table 3 shows the statistical information of the variables.

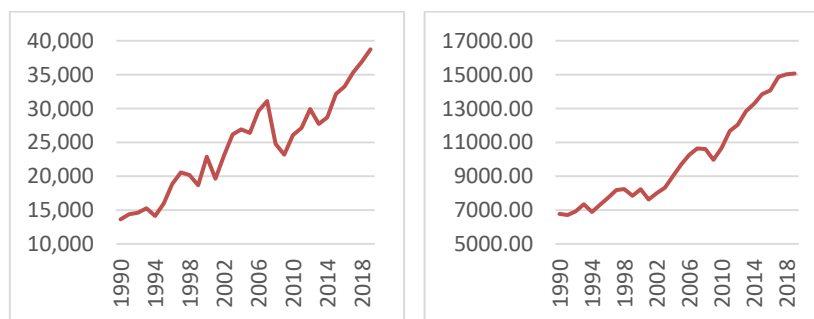
**Table 2.** Definition of the Variables.

| Variable                               | Abbreviation | Unit                          | Source                          |
|--|--------------|-------------------------------|---------------------------------|
| Energy demand in the industrial sector | ED           | Tonne of Oil Equivalent (Toe) | Turkish Ministry of Energy [31] |
| Gross domestic product per capita      | GDP          | Constant 2010 billion US\$    | WorldBank [32]                  |
| Manufactures exports                   | ME           | (% of merchandise exports)    | WorldBank [32]                  |
| Merchandise exports                    | MCE          | Current US\$                  | WorldBank [32]                  |
| CO2 emissions from manufacturing       | CO2          | Mt CO2                        | International Energy Agency[2]  |
| Research and development expenditure   | RDE          | % of GDP                      | WorldBank [32]                  |

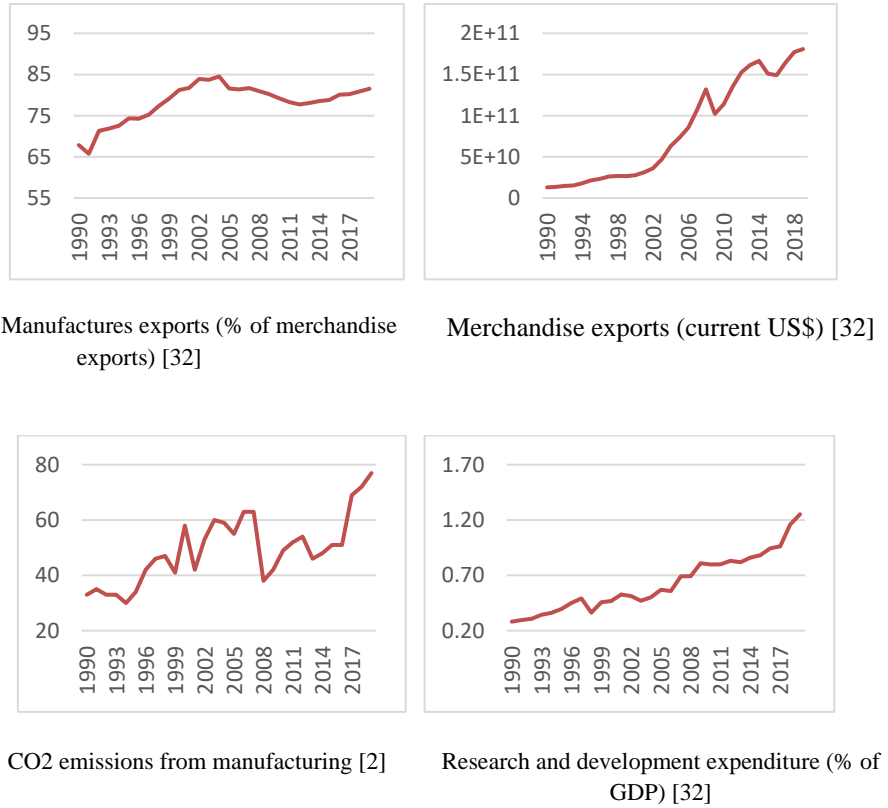
**Table 3.** The Descriptive statistics of Data Set

|                 | ED       | GDP      | ME    | MCE             | CO2   | RDE   |
|-----------------|----------|----------|-------|-----------------|-------|-------|
| <b>Mean</b>     | 24534,90 | 9991,31  | 78,14 | 81911133333,33  | 49,20 | 0,63  |
| <b>Std</b>      | 7123,36  | 2710,70  | 4,62  | 61246919341,38  | 12,32 | 0,26  |
| <b>Min</b>      | 13641,00 | 6709,09  | 65,77 | 12959000000,00  | 30,00 | 0,28  |
| <b>Mak</b>      | 38745,00 | 15068,98 | 84,55 | 180871000000,00 | 77,00 | 1,25  |
| <b>Kurtosis</b> | -0,82    | -0,98    | 0,80  | -1,60           | -0,41 | -0,32 |
| <b>Skew</b>     | 0,14     | 0,61     | -1,09 | 0,31            | 0,40  | 0,65  |

The annual change of the data of input and output variables is given in Fig. 2 graphically. When the last thirty years of energy consumption by sector in Turkey is examined, it is seen that energy consumption has increasing trend in the industrial sector. For the indicators determined regarding energy consumption, it is clear that the variables are in an upward trend.



Industrial sector energy demand (toe) [31]    GDP per capita (constant 2010 US\$) [32]



**Figure 2.** Trends of inputs and output for the Period 1990-2019

Since the analysis data's measurement units are different, bringing the data to the same scale will yield more meaningful results. Normalization will analyze the values of two different properties (variables) in the data set most accurately. It is more appropriate to normalize variables in the range of 0.05-0.95 to make more suggestive comparisons in big data applications [33].

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \times 0.9 + 0.05 \tag{1}$$

where  $x_{min}$  stands for the minimum value of  $x$  and  $x_{max}$  is the maximum value of  $x$ ,  $x'_i$  stands for the  $i^{th}$  normalized value, and  $x_i$  denotes the  $i^{th}$  real value. The normalization process made the graphics more understandable in Fig. 2. The current 30-year data is split into two parts, with 21 data points (70 %) training set and 9 data points (30%) test set.

## 2.2. Adaptive Neuro-Fuzzy Inference System (ANFIS)

The ANFIS model aims to solve nonlinear problems by bringing together the



Fuzzy Inference System (FIS) and Artificial neural networks (ANN). ANFIS was introduced by Jang in 1993 and is called Jang's model ANFIS.

When there is sufficient data for training, correctly defined input and output variables, an appropriate number of layers, and tools to calculate the developed model, artificial neural networks can easily apply and powerful prediction methods. Also, since artificial neural networks can establish relationships according to the data set, they can quickly solve nonlinear problems. Like artificial neural networks, the fuzzy logic system converts a nonlinear input vector to scalar outputs, but it can use numeric values and verbal expressions [27].

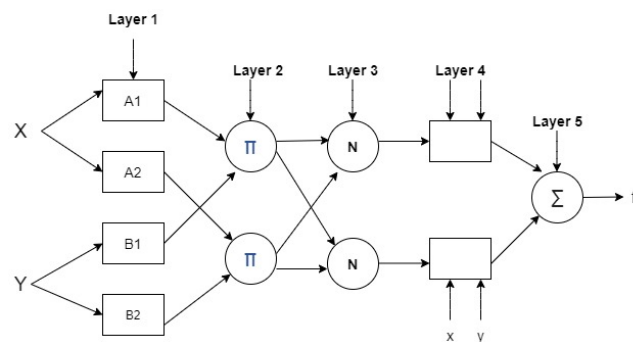
With ANFIS, which uses two methods in a hybrid way, artificial neural networks' learning and computing power are added to fuzzy logic inference systems. In contrast, artificial neural networks gain the ability of fuzzy logic inference systems to make decisions like humans and provide expert knowledge. ANFIS operates Sugeno type models. Mamdani type is a modified version of modeling, and the operations to be applied to input variables are the same, but there is a difference in output variables. In Sugeno type modeling, output variables have membership functions as a function of inputs, which must be linear or constant [10].

In order to understand the architecture of the ANFIS method easily, the structure is given according to two input parameters and one output parameter. A fuzzy rule-based on Takagi – Sugeno can be constructed like in Eq. 2 and Eq. 3[10].

$$\text{If } x_1 \text{ is } A_1 \text{ and } x_2 \text{ is } B_1 \text{ then } f_1 = a_1 x_1 + b_1 x_2 + r_1 \quad (2)$$

$$\text{If } x_1 \text{ is } A_2 \text{ and } x_2 \text{ is } B_2 \text{ then } f_2 = a_2 x_1 + b_2 x_2 + r_2 \quad (3)$$

$A_1, A_2$  are nonlinear parameters and  $B_1, B_2$  are membership functions for  $x_1$  and  $x_2$ .  $a_1, a_2, b_1, b_2, r_1, r_2$  are output parameters. ANFIS architecture has five layers consisting of different functions. These layers are known as the fuzzy layer, product, normalized, defuzzification layer, and output layer [27]. Fig. 3 shows the basic ANFIS structure.



**Figure 3.** ANFIS structure [34].

Each node in layer 1 determines the membership degree of linguistic tags. The membership relationship that includes the input and output functions of layer-1 can be shown:

$$F_i^1 = \mu A_i(x); \quad i = 1,2,3 \dots \quad (4)$$

$$F_i^1 = \mu B_i(y); \quad i = 1,2,3 \dots \quad (5)$$

Where  $F_i^1$ ,  $F_i^2$  represent the output functions and  $\mu A_i(x)$ ,  $\mu B_i(y)$  are membership functions. Each node in this layer is a fixed node that calculates the 'firing strength' ( $w_i$ ) output is the input signal product which calculated to as:

$$F_i^2 = w_i = \mu A_i(x) x \mu B_i(y) \quad i = 1,2,3 \dots \quad (6)$$

$F_i^2$  shows the output of layer-2. Normalization is employed on the layer-3.

$$F_i^3 = w = \frac{w_i}{w_1+w_2} \quad i = 1,2,3 \dots \quad (7)$$

The layer-4 with all nodes is the defuzzification layer. This layer prepares the output values according to the inference of the rules.

$$F_i^4 = \bar{w}_i f_i = w_i (a_i x_1 + b_1 x_2 + r_1) \quad i = 1,2,3 \dots \quad (8)$$

In the layer-5, the sum of outputs of each rule is calculated, and the output can be calculated as follows:

$$F_i^5 = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (9)$$

ANFIS structure has parameters that can be optimized in the layer-1 and the layer-4. The layer-1 contains nonlinear parameters, and the fourth layer contains linear parameters. Both parameters can be changed during the training phase, allowing the model to give better results. Therefore, PSO method was used for parameter optimization in this study.

### 2.3. Particle Swarm Optimization (PSO)

PSO is a heuristic approach proposed as an evolutionary computation method developed to optimize decision-making in continuous and discrete functions. Kennedy and Eberhart first used it in 1995, inspired by animals' physical and social movements as birds and fishes. PSO is a search method in which every possible solution is presented as a particle (swarm) in a population. Particles seek the solution by changing their location in a multi-dimensional search space until optimal state is reached, or computational constraints are exceeded [35].

While each individual is looking for the PSO solution (particle), a swarm is known as the particles' population. The fitness function is used to know the closeness of an individual to the solution.  $X$  is a position (direction) of the particle, and  $v$  represents particle flight speed. Each  $x$  position in the swarm is scored depending on the problem-solving approach [35].

The primary purpose of this function is to measure how close you are to the real solution. While a particle is looking for the solution, its best state at that moment is called "pbest", while the particle closest to the solution throughout the entire swarm is called "gbest". Subsequently, all particles displaced in the solution space are subjected to renewed rules at each iteration for new positions until the newly identified location is reached [25].

The stochastic and deterministic replacement rules showing the regeneration of a particle velocity and position are given in equations. The velocity and the new position vector are obtained by Eq. 10 and Eq. 11, to determine the particle's later position using the information about the particle's positions up to that point.

$$v_{id}^{k+1} = wv_{id}^{k+1} + c_1r_1(p_{id} - x_{id}^k) + c_2r_2(p_{gd} - x_{id}^k) \quad (10)$$

$$x_{id}^{k+1} = x_{id}^{k+1} + v_{id}^{k+1} \quad i=1,2,3,\dots,n \quad d=1,2,3,\dots,m \quad (11)$$

In Eq. 10 and Eq. 11,  $w$  represents the inertia constant used to balance the algorithm's local and global search capabilities, and the global search for large values of  $w$  and local search for small values.  $c_1$  and  $c_2$  are acceleration coefficients representing learning behaviors.  $r_1$  and  $r_2$  are numbers between 0 and 1,  $k$  represents the number of iterations,  $n$  is the population size,  $m$  is the number of decision variables,  $p_{id}$  is the best position,  $x_{id}$  is  $k$  position,  $v_{id}^{k+1}$  is velocity values and  $p_{gd}$  refers to the best position [25].

#### 2.4. Multiple Linear Regression (MLR)

MLP is a method to show the relationship between variables. The interaction between variables can be modeled as linear. The general formula of MLR is given in Eq. 12.

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_{k-1}X_{k-1} + e \quad (12)$$

Where  $Y$  is a dependent variable,  $Y$  is equal to  $a$  when  $X$  is zero, known as intercept.  $b$  is the change in  $Y$  caused by a unit change in  $X$  known as slope, and  $e$  is the residual. A mathematical relation is established between input and output values of a linear system, and it can be determined which parameters have the most binding effect by MLR analysis [36].

## 2.5. Performance evaluation criteria

The model's accuracy can be tested using some of the performance evaluation criteria found in the literature. Mean Square Error (MSE), Root Mean Square Error (RMSE), the coefficient of determination (R2) values were used to compare the results of the two methods and make meaningful interpretations. The performance criteria equations are given in Eq. 13 to Eq. 15, respectively [37].

$$MSE = \frac{1}{n} \sum_{t=1}^n (ED_i^{observed} - ED_i^{predicted})^2 \quad (13)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (ED_i^{observed} - ED_i^{predicted})^2} \quad (14)$$

$$R^2 = \left( \frac{\sum_{i=1}^n (ED_i^{observed} - \overline{ED_i^{observed}}) (ED_i^{predicted} - \overline{ED_i^{predicted}})}{\sqrt{\sum_{i=1}^n (ED_i^{observed} - \overline{ED_i^{observed}})^2 (ED_i^{predicted} - \overline{ED_i^{predicted}})^2}} \right)^2 \quad (15)$$

Where n is the number of observed values,  $ED_i^{observed}$  is the actual value of energy demand at the time i and  $ED_i^{predicted}$  is the predicted value of industrial energy demand at time i.

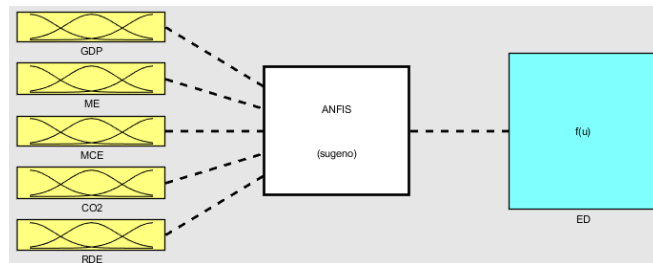
## 3. Results and Discussion

### 3.1. Results of MLR, ANFIS, PSO-ANFIS

MLR model is established to investigate the relation between five inputs and energy demand in industry as the dependent variable. The training data set was used the MLR model to predict ED (Y) using GDP (X1), ME (X2), MCE (X3), CO2 (X4), and RDE (X5) variables. The coefficients of the regression model were calculated using STATA 15.1. The following regression equation in Eq. 16 was used to predict ED.

$$ED = -0.0381 + 0.4122x_1 - 0.1594x_2 - 0.1593x_3 - 0.4229x_4 - 0.0901x_5 \quad (16)$$

ANFIS model to forecast industrial sector energy demand is developed, relying on the Sugeno fuzzy inference system. It has five inputs: GDP, ME, MCE, CO2, RDE, and the one output variable is ED. The model was implemented using MATLAB 2020a. Fig. 4 indicates the developed ANFIS model architecture.



**Figure 4.** ANFIS model for energy demand forecasting.

In order to get sufficient predictability using ANFIS, the model should have an appropriate number of clusters. The model created in this study makes fuzzy clustering using the Fuzzy c-means method. The results of this method have affected the amount of data. To state the membership functions and the number of rules, the function created based on the fuzzy c-means method known as "genfis3" in MATLAB 2020a is used in the rule extraction stage.

For the PSO-ANFIS model, ANFIS parameters have been optimized to increase the predictability by using PSO. The Gaussian function has two parameters to optimize instead of other membership functions, as suggested by many researchers. PSO optimization in ANFIS model ensures closer relationships between output and inputs. PSO parameters such as population size, the maximum number of iterations, initial inertia weight, inertia weight damping ratio, and learning coefficients should be selected appropriately to develop PSO-ANFIS [25]. The PSO-ANFIS model was performed using MATLAB 2020a software.

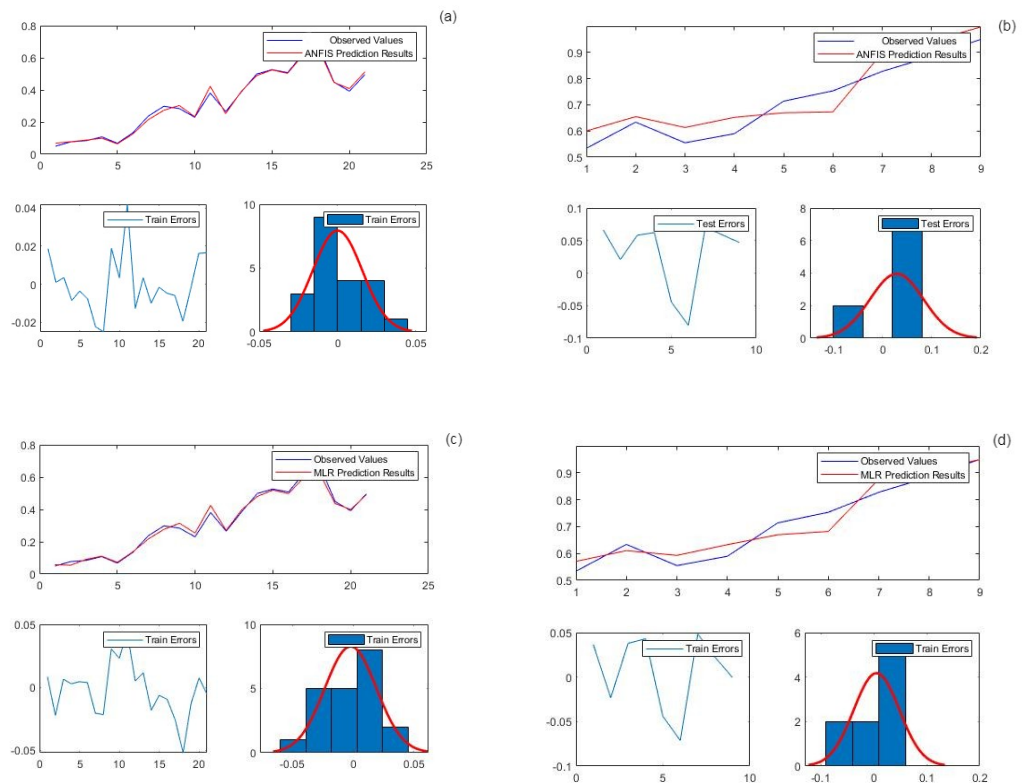
### 3.2. Forecasting accuracy evaluation

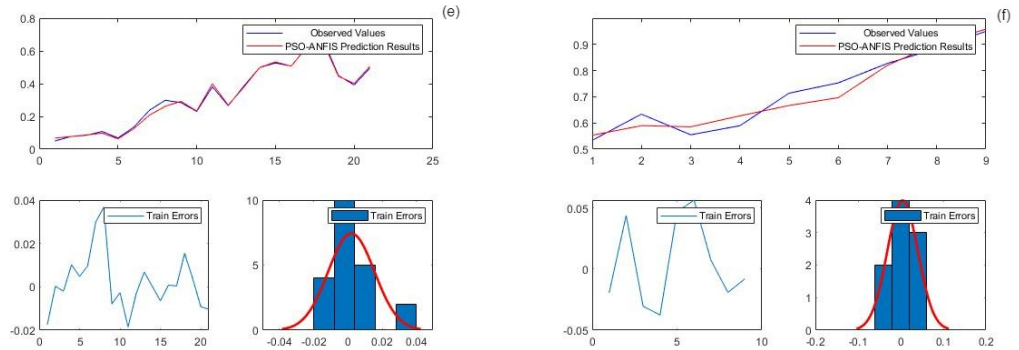
The network structure has a remarkable impression on the computational intricacy of the model and the efficiency of the predictions. In addition, the performance prediction of the ANFIS model is positively related to optimizable parameters. In this paper, the energy demand is an output variable, and five indicators affecting this demand are input variables in all developed models. Relationships between input and output are established via training data. The test data set is then used to check the accuracy of the structure. The available data on industrial sector energy demand is divided into two parts: 70% (21 data points) training and 30% (9 data points) testing parts. RMSE, MSE, and R2 values are calculated as performance evaluation criteria. The comparison of the models for performance criteria is seen in Table 4.

**Table 4.** The comparison of the developed models

| Developed Model  | Training Phase  |               |                | Testing Phase   |                |                |
|------------------|-----------------|---------------|----------------|-----------------|----------------|----------------|
|                  | MSE             | RMSE          | R <sup>2</sup> | MSE             | RMSE           | R <sup>2</sup> |
| <b>ANFIS</b>     | 0.000239        | 0.0155        | 0.9932         | 0.003500        | 0.0588         | 0.8776         |
| <b>MLR</b>       | 0.000436        | 0.0209        | 0.9889         | 0.001670        | 0.0409         | 0.9181         |
| <b>PSO-ANFIS</b> | <b>0.000178</b> | <b>0.0134</b> | <b>0.9951</b>  | <b>0.001172</b> | <b>0.03424</b> | <b>0.9423</b>  |

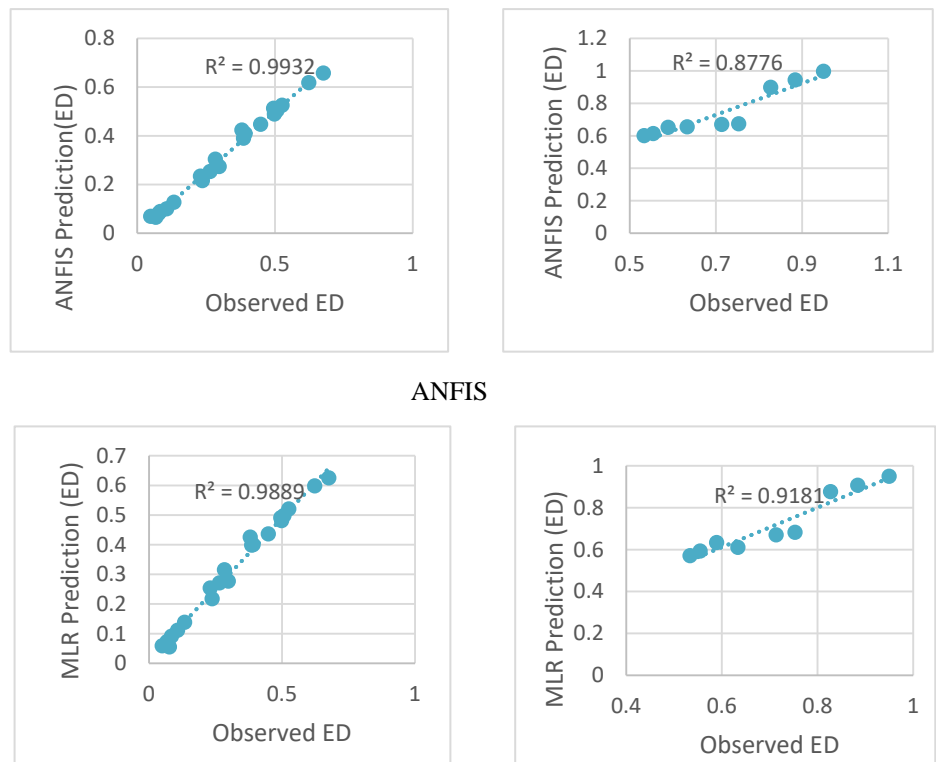
Table 4 shows that the PSO algorithm successfully enhances the performance of the ANFIS model. Since the values for the coefficient of the determination (R<sup>2</sup>) are higher than the other two methods, this situation reinforces the model's reliability in terms of the estimation process. The graphics of the training and test analysis results are given in Fig. 5.



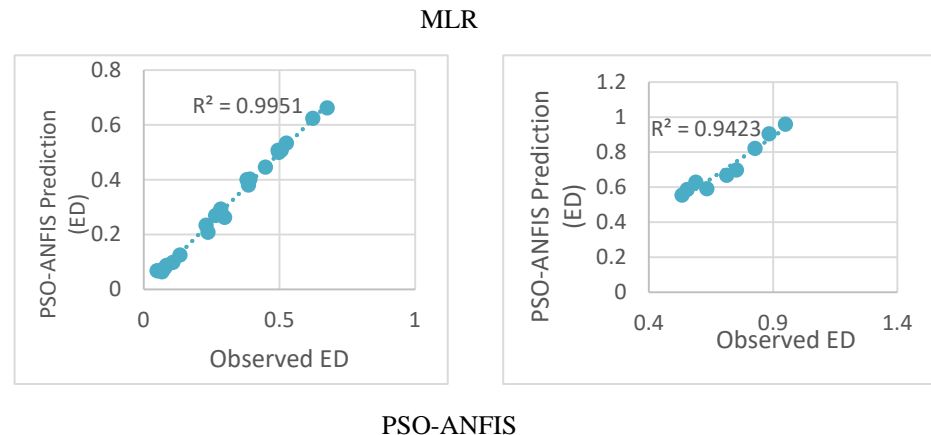


**Figure 5.** The training and test analysis results. ((a). ANFIS Training phase (b). ANFIS Testing phase (c). MLR Training phase (d). MLR Testing phase (e). PSO-ANFIS Training phase (f). PSO-ANFIS Testing phase)

The coefficient of determination ( $R^2$ ) value of PSO-ANFIS was more significant than the ANFIS and MLR for estimation, which is found as 0.9951, 0.9932, and 0.9889 for the testing phase, 0.9423, 0.8776, 0.9181 for the testing phase, respectively. As shown in Fig. 6, to get a predictive success, the actual values were graphed against predicted values, and  $R^2$  values were calculated.



ANFIS



**Figure 6.** The coefficient of the determination values.

#### 4. Conclusions and Recommendations

As energy is vital to societies' safety and well-being, determining energy demand has become an important research topic, lately. Among economic theories, estimating energy consumption has an important place in energy planning since energy is one of the most critical industrial production resources. The supply and demand for energy should be planned in the long term for sustainable development. Accurate estimates can help decision-makers know the volume and trend of energy consumption. In this paper, a prediction model is designed to enlighten the factors affecting industrial sector energy consumption. The models are applied with artificial intelligence tools such as ANFIS, PSO-ANFIS, and a traditional tool as MLR. It has been observed that the industrial energy demand, which varies depending on different economic and social parameters, has been estimated by the PSO-ANFIS method accurately. Decision-makers can use these results as a source for future sectoral energy plans. In future studies, the ANFIS can be combined with other heuristic optimization algorithms, and hybrid models can be compared with classical econometric models and machine learning methods.

#### References

- [1] IEA. IEA World Energy Report 2005. <https://www.iea.org>.
- [2] IEA. IEA World Energy Balances 2020. <https://www.iea.org/subscribe-to-data-services/world-energy-balances-and-statistics>.
- [3] Bayramoğlu T, Pabuçcu H, Çelebi Boz F, Türkiye İçin Anfis Modeli İle Birincil Enerji Talep Tahmini. Ege Akad Bakis (Ege Acad Rev 2017;17:431–46. <https://doi.org/10.21121/eab.2017328408>.
- [4] Barak S, Sadegh SS. Forecasting energy consumption using ensemble



- ARIMA-ANFIS hybrid algorithm. *Int J Electr Power Energy Syst* 2016;82:92–104. <https://doi.org/10.1016/j.ijepes.2016.03.012>.
- [5] Jawad M, Nadeem MSA, Shim SO, Khan IR, Shaheen A, Habib N, et al. Machine Learning Based Cost Effective Electricity Load Forecasting Model Using Correlated Meteorological Parameters. *IEEE Access* 2020;8:146847–64. <https://doi.org/10.1109/ACCESS.2020.3014086>.
- [6] Karaaslan A, Gezen M. Forecasting of Turkey's sectoral energy demand by using fuzzy grey regression model. *Int J Energy Econ Policy* 2017;7:67–77.
- [7] Adedeji P, Madushele N, Akinlabi S. Adaptive Neuro-fuzzy Inference System (ANFIS) for a multi-campus institution energy consumption forecast in South Africa. *Proc Int Conf Ind Eng Oper Manag* 2018;2018:950–8.
- [8] Yayar R, Hekim M, Yılmaz V, Bakırcı F. A comparison of ANFIS and ARIMA techniques in the forecasting of electric energy consumption of Tokat province in Turkey. *J Econ Soc Stud* 2011;1:87–112. <https://doi.org/10.14706/jecoss11124>.
- [9] Akay D, Atak M. Grey prediction with rolling mechanism for electricity demand forecasting of Turkey. *Energy* 2007;32:1670–5. <https://doi.org/10.1016/j.energy.2006.11.014>.
- [10] Demirci E. Anfis İle Doğalgaz Talep Tahmini; Türkiye Örneği. *Int J Soc Sci Acad* 2020.
- [11] Yang Y, Chen Y, Wang Y, Li C, Li L. Modelling a combined method based on ANFIS and neural network improved by DE algorithm: A case study for short-term electricity demand forecasting. *Appl Soft Comput J* 2016;49:663–75. <https://doi.org/10.1016/j.asoc.2016.07.053>.
- [12] Ahmad AS, Hassan MY, Abdullah MP, Rahman HA, Hussin F, Abdullah H, et al. A review on applications of ANN and SVM for building electrical energy consumption forecasting. *Renew Sustain Energy Rev* 2014;33:102–9. <https://doi.org/10.1016/j.rser.2014.01.069>.
- [13] Vinagre E, Pinto T, Ramos S, Vale Z, Corchado JM. Electrical Energy Consumption Forecast Using Support Vector Machines. *Proc - Int Work Database Expert Syst Appl DEXA* 2017:171–5. <https://doi.org/10.1109/DEXA.2016.046>.
- [14] Kasule A, Ayan K. Using PSO and Genetic algorithms to optimize ANFIS model for forecasting Uganda's net electricity consumption. *Sak Univ J Sci* 2020:324–37. <https://doi.org/10.16984/saufenbilder.629553>.
- [15] Aghay SH, Fallahpour A, Selvaraj J, Rahim NA. Long-term electrical energy consumption formulating and forecasting via optimized gene expression programming. *Energy* 2017;126:144–64. <https://doi.org/10.1016/j.energy.2017.03.009>.
- [16] Huang Y, Yang L, Gao C, Jiang Y, Dong Y. A novel prediction approach for short-term renewable energy consumption in china based on improved gaussian process regression. *Energies* 2019;12:1–17. <https://doi.org/10.3390/en12214181>.

- [17] Beyca OF, Ervural BC, Tatoglu E, Ozuyar PG, Zaim S. Using machine learning tools for forecasting natural gas consumption in the province of Istanbul. *Energy Econ* 2019;80:937–49. <https://doi.org/10.1016/j.eneco.2019.03.006>.
- [18] Liu X, Moreno B, García AS. A grey neural network and input-output combined forecasting model. Primary energy consumption forecasts in Spanish economic sectors. *Energy* 2016;115:1042–54. <https://doi.org/10.1016/j.energy.2016.09.017>.
- [19] Alcaraz C, Villalvazo S. The effect of natural gas shortages on the Mexican economy. *Energy Econ* 2017;66:147–53. <https://doi.org/10.1016/j.eneco.2017.06.006>.
- [20] Shaikh F, Ji Q. Forecasting natural gas demand in China: Logistic modelling analysis. *Int J Electr Power Energy Syst* 2016;77:25–32. <https://doi.org/10.1016/j.ijepes.2015.11.013>.
- [21] Naji S, Shamshirband S, Basser H, Alengaram UJ, Jumaat MZ, Amirmojahedi M. Soft computing methodologies for estimation of energy consumption in buildings with different envelope parameters. *Energy Effic* 2016;9:435–53. <https://doi.org/10.1007/s12053-015-9373-z>.
- [22] Azadeh A, Asadzadeh SM, Ghanbari A. An adaptive network-based fuzzy inference system for short-term natural gas demand estimation: Uncertain and complex environments. *Energy Policy* 2010;38:1529–36. <https://doi.org/10.1016/j.enpol.2009.11.036>.
- [23] Hamzaçebi C. Forecasting of Turkey's net electricity energy consumption on sectoral bases. *Energy Policy* 2007;35:2009–16. <https://doi.org/10.1016/j.enpol.2006.03.014>.
- [24] Moayedi H, Raftari M, Sharifi A, Jusoh WAW, Rashid ASA. Optimization of ANFIS with GA and PSO estimating  $\alpha$  ratio in driven piles. *Eng Comput* 2020;36:227–38. <https://doi.org/10.1007/s00366-018-00694-w>.
- [25] Basser H, Karami H, Shamshirband S, Akib S, Amirmojahedi M, Ahmad R, et al. Hybrid ANFIS-PSO approach for predicting optimum parameters of a protective spur dike. *Appl Soft Comput J* 2015;30:642–9. <https://doi.org/10.1016/j.asoc.2015.02.011>.
- [26] Shamshirband S, Hadipoor M, Baghban A, Mosavi A, Bukor J, Várkonyi-Kóczy AR. Developing an ANFIS-PSO model to predict mercury emissions in combustion flue gases. *Mathematics* 2019;7:1–16. <https://doi.org/10.3390/math7100965>.
- [27] Pousinho HMI, Mendes VMF, Catalão JPS. A hybrid pso-anfis approach for short-term wind power prediction in Portugal. *Energy Convers Manag* 2011;52:397–402. <https://doi.org/10.1016/j.enconman.2010.07.015>.
- [28] Kouziokas GN. A new W-SVM kernel combining PSO-neural network transformed vector and Bayesian optimized SVM in GDP forecasting. *Eng Appl Artif Intell* 2020;92:103650. <https://doi.org/10.1016/j.engappai.2020.103650>.
- [29] Hamit-Hagggar M. Greenhouse gas emissions, energy consumption and

- economic growth: A panel cointegration analysis from Canadian industrial sector perspective. *Energy Econ* 2012;34:358–64. <https://doi.org/10.1016/j.eneco.2011.06.005>.
- [30] Déés S, Karadeloglou P, Kaufmann RK, Sánchez M. Modelling the world oil market: Assessment of a quarterly econometric model. *Energy Policy* 2007;35:178–91. <https://doi.org/10.1016/j.enpol.2005.10.017>.
- [31] TME. Energy Efficiency Portal 2020. <https://enverportal.enerji.gov.tr/>.
- [32] Bank W. World Bank Open Data 2020. <https://data.worldbank.org/>.
- [33] Wang Y, Li J, Gu J, Zhou Z, Wang Z. Artificial neural networks for infectious diarrhea prediction using meteorological factors in Shanghai (China). *Appl Soft Comput J* 2015;35:280–90. <https://doi.org/10.1016/j.asoc.2015.05.047>.
- [34] Zahedi G, Azizi S, Bahadori A, Elkamel A, Wan Alwi SR. Electricity demand estimation using an adaptive neuro-fuzzy network: A case study from the Ontario province - Canada. *Energy* 2013;49:323–8. <https://doi.org/10.1016/j.energy.2012.10.019>.
- [35] Duan H, Lei GR, Shao K. Forecasting crude oil consumption in China using a grey prediction model with an optimal fractional-order accumulating operator. *Complexity* 2018;2018. <https://doi.org/10.1155/2018/3869619>.
- [36] Civelekoglu G, Yigit NO, Diamadopoulos E, Kitis M. Prediction of bromate formation using multi-linear regression and artificial neural networks. *Ozone Sci Eng* 2007;29:353–62. <https://doi.org/10.1080/01919510701549327>.
- [37] Becerra M, Jerez A, Aballay B, Garcés HO, Fuentes A. Forecasting emergency admissions due to respiratory diseases in high variability scenarios using time series: A case study in Chile. *Sci Total Environ* 2020;706. <https://doi.org/10.1016/j.scitotenv.2019.134978>.