

# Prediction of NO<sub>x</sub> Emissions with A Novel ANN Model in Adana

Senem Sezer  Halime Yakisik  Furkan Kartal  N. Nur Fucucu  Yagmur Dalbudak  Sevi Yasar   
Ugur Ozveren 

Marmara University, Department of Chemical Engineering, Istanbul, Turkey

## ABSTRACT

NO<sub>x</sub> emissions are one of the typical air pollutants that has drawn worldwide attention. NO<sub>x</sub> emissions from air cause detrimental effects on the environment and human health such as lung cancer, asthma, allergic rhinitis, and mental diseases. Therefore, real-time NO<sub>x</sub> monitoring has been very popular research topics in atmospheric and environmental science. However, the spatial coverage of monitoring stations within Adana is limited and thus often insufficient for exposure. Moreover, NO<sub>x</sub> monitoring stations are also lacking to reveal the influences of meteorological and air pollutant effects. In this study, artificial neural network (ANN), which is a biological mimicked computer algorithm that simulates the functions of neurons using artificial neurons, has been used to present a quantitative determination of the NO<sub>x</sub> emission in Adana through the influences of temperature (°C), wind rate (km/h), and SO<sub>2</sub> (µg/m<sup>3</sup>) on NO<sub>x</sub> emissions. The high R<sup>2</sup> values in testing dataset lead to the conclusion that the artificial neural network model provides predictions. The developed model in study is a useful tool for the design and planning of air pollution control policies as well as reducing economic cost. The developed model in study is a useful tool for the design and planning of air pollution control policies as well as reducing economic cost.

### Keywords:

Adana, Air Pollution, NO<sub>x</sub> Emissions, Artificial Neural Networks, Prediction

## INTRODUCTION

Air quality is becoming an essential environmental issue in nowadays [1-3]. Researchers have been interested in air quality due to severe health problems caused by air pollution as it is one of the factors that effects the nature of human beings [1, 4]. Air pollution have been considered as one of the most important environmental concern especially for urban areas due to their high population and industries [4, 5].

There are a number of pollutants that are also known ground-level air pollutants, namely PM<sub>10</sub>, O<sub>3</sub>, CO, NO<sub>x</sub>, NO, NO<sub>2</sub>, H<sub>2</sub>S and SO<sub>2</sub> [6]. Nitrogen oxide, generally known as NO<sub>x</sub>, is a common term for compounds include oxygen and nitrogen such as NO<sub>3</sub>, NO, N<sub>2</sub>O, N<sub>2</sub>O<sub>3</sub>, NO<sub>2</sub>, N<sub>2</sub>O<sub>4</sub>, and N<sub>2</sub>O<sub>5</sub> [7, 8]. Combustion emissions and vehicle exhaust are the main source of NO<sub>x</sub> [9, 10]. NO<sub>x</sub> emissions can cause acid rains and photochemical smog that has an effect on human health (such as infection, cancer, etc.) and air quality [9]. Inhaling NO<sub>x</sub> is associated with increased health problems including respiratory disease, difficulty in breathing, and premature death [11]. Moreover, accumulation of NO<sub>x</sub> that is come from the air into water, lowering the

oxygen concentration in the water, causes acidification and eutrophication of lakes, which can damage the survival of aquatic plants and other organisms [8].

Prediction of NO<sub>x</sub> has great importance for these environmental and health concerns. Current air quality monitoring systems provide a great amount of dataset and therefore, loss of data is occurred commonly due to instrument break down, failure of data transmission and maintenance, etc. Therefore, these monitoring techniques have failed to measure the NO<sub>x</sub> concentration for monitoring air quality and exhibited low precision in long-term predictions according to industrial and meteorological changes [2].

In recent years, artificial neural network structures (ANNs) have been widely used as a well-known technique of capturing nonlinear relations for air pollutants prediction and air quality monitoring [6]. ANN modeling is a proper mathematical approach for demonstrating highly complex relationships and can be generalized accurately when new input parameter was presented [5]. Compared to the traditional modeling

### Article History:

Received: 2020/02/21

Accepted: 2020/11/14

Online: 2020/12/31

**Correspondence to:** Uğur Özveren,  
Department of Chemical Engineering,  
Marmara University, Istanbul, Turkey  
Tel: +90 5358837701  
Fax: +90 2163480293  
E-Mail: ugur.ozveren@marmara.edu.tr

techniques, ANN is a data driven, self-adaptive, black-box method, which learns from examples.

There are several studies on prediction of ground-level air pollutants have been conducted with an ANN model. Sofuoglu et al. [12]. utilized an ANN model that examines meteorological parameters (wind speed and temperature) and measured particulate matter concentrations as input variables, to forecast  $\text{SO}_2$  concentrations in Izmir. Baawain and AS Al-Serhi [6] proposed a rigorous method of preparing air quality data to achieve more accurate air pollution prediction models which based on an artificial neural network (ANN). Lal Benjamin et al. [5] developed two air quality prediction models using feed-forward neural network to predict  $\text{NO}_x$  concentrations in Ujjain city that belongs to India. However, there aren't any study for the air quality prediction in Adana.

The main aim of the present study is to reveal the impact of meteorological and air pollutant influence for  $\text{NO}_x$  prediction in Adana. Therefore, a method based on artificial neural networks (ANNs) model generation is presented to forecast monthly  $\text{NO}_x$  concentrations in Adana province. Meteorological influencing factors including wind rate (km/h), temperature ( $^{\circ}\text{C}$ ), and  $\text{SO}_2$  concentrations ( $\mu\text{g}/\text{m}^3$ ) which is related traffic and industrial effect were selected as input variables to develop neural network model that is capable of estimating  $\text{NO}_x$  concentrations ( $\mu\text{g}/\text{m}^3$ ) as output parameter.

## MATERIAL AND METHODS

### Material

Adana is the sixth largest city of Turkey with 1,854,270 inhabitants in 2017 [13]. The city of Adana is also localized in the interface of developed and relatively underdeveloped cities in Turkey. Therefore, high unusual urbanization rate (75%) is occurred due to migration from rural areas and the lack of resources and it can cause urban problems. As conclusion, the  $\text{NO}_x$  yearly average for Adana is supposed as the higher levels compared to the other locations of Turkey. The monthly variations of average  $\text{NO}_x$  and  $\text{SO}_2$  emissions in Adana between 2014 and 2018, which was obtained from Republic of Turkey Ministry of Environment and Urbanization, "Air Quality Monitoring Stations" website [14] have been presented in Table 1. The monthly average of wind rate and temperature values in Table 1. have been acquired by using World Weather Online website [15]. The dataset was selected by randomly by using MATLAB SOFTWARE and was divided into three parts: training, validation, and testing.

**Table 1.** The monthly variations of average  $\text{NO}_x$  emissions.

Year	Month	$\text{SO}_2$ ( $\mu\text{g}/\text{m}^3$ )	Wind Rate (km/h)	Temperature ( $^{\circ}\text{C}$ )	$\text{NO}_x$ ( $\mu\text{g}/\text{m}^3$ )
2018	January	5.214576	8.4	9	34.87
	February	3.762245	7.7	12	20.45
	March	3.762245	7.5	15	20.45
	April	3.781792	7.8	18	21.88
	May	1.761529	8.4	24	7.87
	June	9.736235	9.3	26	10.67
	July	7.713672	10.1	28	9.68
	August	3.622233	9.6	28	8.67
	September	10.90257	8.2	27	15.83
	October	10.40373	7.6	27	5.80
	November	11.93729	8.4	19	5.89
	December	11.40444	9.2	13	47.97
2018	January	6.260445	8.4	9	24.61
	February	7.510149	7.7	12	20.02
	March	8.140121	7.5	15	16.22
	April	10.08224	7.8	18	12.96
	May	8.284132	8.4	24	15.83
	June	12.01866	9.3	26	13.43
	July	6.105595	10.1	28	3.43
	August	4.848124	9.6	28	4.32
	September	6.035191	8.2	27	9.75
	October	7.267949	7.6	27	21.57
	November	4.333663	8.4	19	24.14
	December	10.53031	9.2	13	34.90
2017	January	1.535483	9.7	6	30.88
	February	1.707932	8	8	27.10
	March	3.868253	7.9	12	22.96
	April	6.838653	8.4	17	20.92
	May	5.529328	7.8	21	17.10
	June	6.997	8	26	12.17
	July	4.274731	9.4	30	11.75
	August	2.379242	9.5	29	12.84
	September	5.648059	8.2	27	18.59
	October	5.347702	7.8	21	26.04
	November	3.267917	8.4	15	36.67
	December	5.146672	7.7	12	56.06
2016	January	12.18788	9.4	7	33.45
	February	14.43249	7.6	12	33.21
	March	9.846644	9.1	14	21.49
	April	11.53489	7.8	19	17.36
	May	10.24737	8.5	21	13.36
	June	8.026418	8.6	26	14.27
	July	8.55995	8.9	29	9.26
	October	60.60736	7.2	22	30.41
	November	37.57438	7.7	15	44.48
	December	13.45699	9.8	7	36.70

**Table 1 (cont).** The monthly variations of average NO<sub>x</sub> emissions.

Year	Month	SO <sub>2</sub> (μg/m <sup>3</sup> )	Wind Rate (km/h)	Temperature (°C)	NO <sub>x</sub> (μg/m <sup>3</sup> )
Air Quality Monitoring Station: Meteoroloji 2015	January	8.707439	9.7	8	21.19
	February	6.15348	10.2	10	16.43
	March	8.704655	9.1	13	11.42
	April	11.22918	8.4	16	5.53
	May	7.167712	8.7	22	9.83
	June	4.619178	11.1	24	5.49
	July	4.435859	10.6	28	3.82
	August	6.281179	10	29	3.03
	September	8.634721	7.9	27	5.85
	October	8.059739	7.7	23	10.56
	November	8.29913	8.3	17	12.66
	December	9.163185	8.2	11	13.50
Air Quality Monitoring Station: Vallik 2014	January	22.57541	7.8	10	66.68
	February	6.273418	7.9	11	58.34
	March	4.833347	9.5	13	45.08
	April	2.458409	7.6	17	24.94
	May	4.003284	7.4	21	17.92
	June	4.180194	8	25	14.65
	July	3.801008	9.9	27	10.78
	August	4.205149	8.5	28	12.70
	September	3.965452	8.3	25	20.83
	October	4.155049	7.5	21	39.53
Air Quality Monitoring Station: Meteoroloji 2014	January	11.08059	7.8	10	52.07
	February	9.447555	7.9	11	37.00
	March	3.383083	9.5	13	32.24
	April	2.053013	7.6	17	20.16
	May	2.468017	7.4	21	22.68
	June	2.852342	8	25	21.32
	July	3.486149	9.9	27	23.54
	August	3.21794	8.5	28	38.03
	September	3.877626	8.3	25	59.84
	October	4.591569	7.5	21	48.37

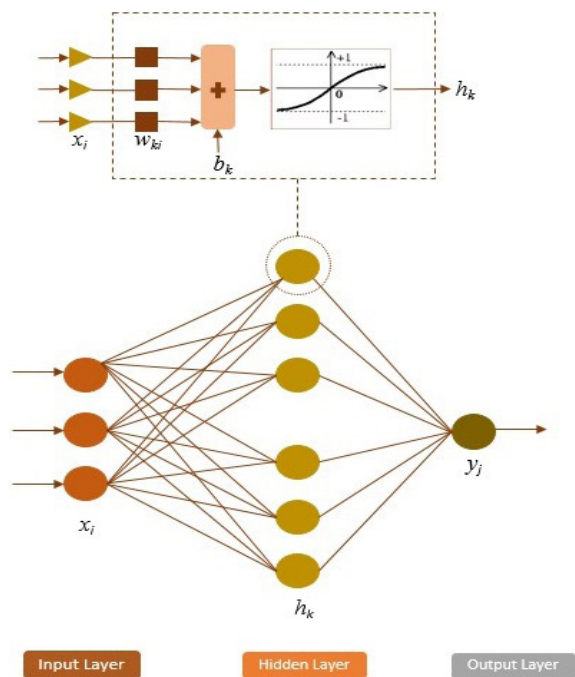
## Material

### Artificial neural network

ANN is a set of algorithms includes the functions of neurons that simulates massively parallel-distributed information processing system, inspired by neuroscience. The most commonly used ANN is multi-layer perceptron (MLP) based feed-forward back propagation model. MLP is a layered neural network structure (Fig. 1) with more than a single layer including one or several neuron(s). In this structure, the input of a layer is the output of the previous layer which is achieved through the activation function. As a result, the equation to determine the output of an artificial node *j* is given by:

$$h_k = f\left(\sum_{i=1}^n x_i \cdot w_{ki} + b_k\right) \quad (1)$$

where  $h_k$  is the calculated value called output from the  $i^{\text{th}}$  node at the previous layer,  $w_{ki}$  is the weight of the artificial node between  $i^{\text{th}}$  and  $k^{\text{th}}$  node, and  $b_k$  is the bias value. The net weights  $w_{ki}$  and biases  $b_k$  are calculated based on training algorithm. Along training procedure, training algorithm is updated the weights and deviations to reduce the error between predicted and actual values of the model. In this study we used back propagation algorithm as a training algorithm, which is accomplished through adjusting the gradient weights to minimizing the difference between target output and network output respectively.



**Figure 1.** The structure of a multilayer perceptron neural network.

Backpropagation algorithm is probably the most popular learning algorithm for supervised learning of artificial neural networks using gradient descent based on generalizing the Widrow-Hoff learning rule. In this study, backpropagation algorithm, which is widely applied in a variety of engineering applications, was applied as the learning algorithm of ANN model.

### Levenberg-Marquardt Algorithm

The Levenberg-Marquardt (LM) is the most used regularization algorithm for backpropagation algorithm to minimize the Mean Squared Error (MSE) of proposed neural network structure. This optimization algorithm is a kind of pseudo second order method and determines the best direction to move the weights with minimization methods accurately when model network topology has

small size. Additionally, the Levenberg–Marquardt algorithm is an iterative process to implement second-order training methods apply the Hessian matrix to determine weight values. The model validation criteria of the proposed ANN model are required to predict NO<sub>x</sub> concentration according to other published correlations on the base of the determination coefficient (R<sup>2</sup>) examination.

### Validation of the model

There are several types of correlation validation criteria exist with their own range of usability to examine the developed model efficiency. These parameters represent prediction accuracy of a statistical estimating method, which expresses generalization performance. Three statistical parameters are employed to make correlation for prediction of NO<sub>x</sub> concentration, which are computed as follows:

Correlation Coefficient (R<sup>2</sup>):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{est} - y_{exp})^2}{\sum_{i=1}^n (y_{est} - \bar{y}_{exp})^2} \cdot 100\% \quad (2)$$

Mean absolute percentage error (MAPE):

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{Y_{actual} - Y_{predicted}}{Y_{actual}} \right| \quad (3)$$

Root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_{predicted} - Y_{actual})^2} \quad (4)$$

## RESULTS AND DISCUSSION

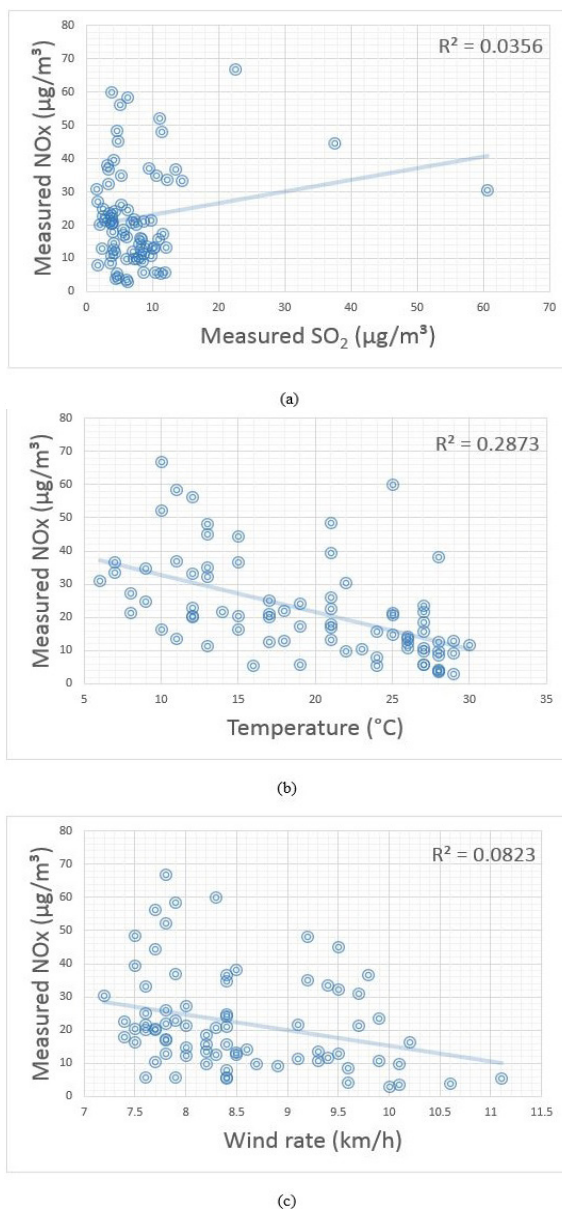
### The Impact of Input Parameters

Fig. 2 were generated by plotting the individual input parameters that are SO<sub>2</sub> (µg/m<sup>3</sup>), temperature (°C) and wind rate (km/h) of the air quality prediction against the corresponding measured NO<sub>x</sub> values.

It is concluded that linear models for individual parameters may not represent the most proper solution to precisely estimate measured NO<sub>x</sub> emissions. This complexity due to the nonlinearity can be handle using artificial intelligence based on data-driven modeling in the form of ANNs.

### Neural Network Model

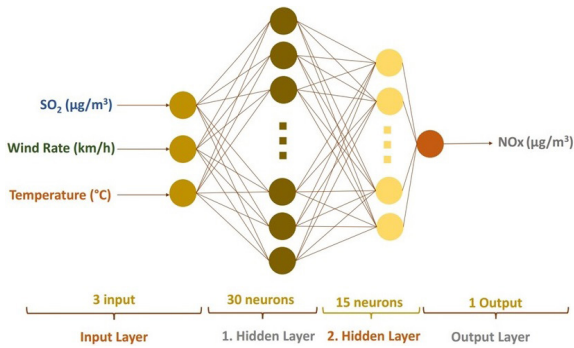
In this study, the feed-forward multi-layer neural network is used due to its ability to model very effectively any measurable input-output relationship to any desired degree of accuracy. The ANN model consists of three input variables: i) SO<sub>2</sub> (µg/m<sup>3</sup>), ii) wind rate (km/h), iii) temperature (°C). The input data was split into test, training and validation sets.



**Figure 2.** Graphics of the measured NO<sub>x</sub> values against input values that are (a) SO<sub>2</sub> (µg/m<sup>3</sup>), (b) wind rate (km/h) and (c) temperature (°C).

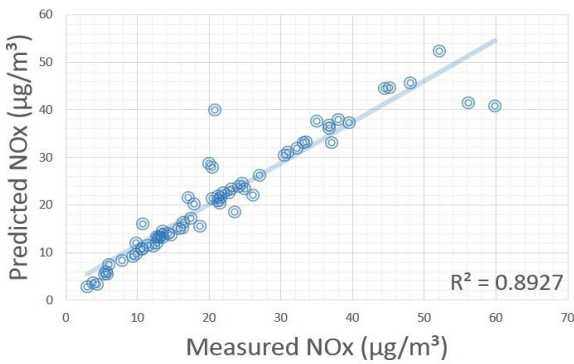
In the absence of any accurate theory to calculate the number of artificial neurons and hidden layers, the network structure related model parameters were calculated heuristically by using trial and error methods. Therefore, the applicability of 350 different ANN models was examined with various activation functions and topology including different number of hidden layers. As seen on Fig. 3, the developed ANN model includes four layers that are input layer, output layer, and two hidden layers.

The first hidden layer has thirty neurons and the second hidden layer consists of four neurons. The tangent sigmoid (tansig) functions were used for the neurons input layer, first and second hidden layers respectively, and line-



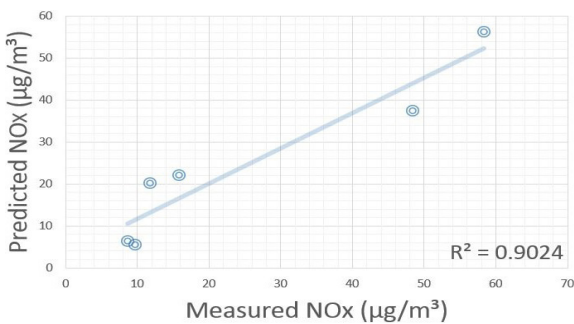
**Figure 3.** Artificial Neural Network Model.

ar function was utilized for output layer. In order to use an artificial neural network model, one needs first to train the proposed model with training dataset. Fig. 4 shows the values obtained training datasets which reached an overall score of 0.8927 when using the Levenberg-Marquardt regularization method. According to this figure, the correlation between the predicted and measured values was high enough to assert that the developed ANN model demonstrates a good agreement with the training datasets.



**Figure 4.** Comparing the predicted and measured NOx values that obtained training data sets.

The validation dataset was used to stop training early if the model performance fails. As seen on Fig. 5,  $R^2$  is higher than 0.90 for the validation set. The network structure is capable to generalize the prediction of  $\text{NO}_x$  emissions.



**Figure 5.** Comparing the predicted and measured NOx values that obtained validation data sets.

Fig. 6 indicated that the prediction performance of the

**Table 2.** MAPE and RMSE results for the prediction of  $\text{NO}_x$  emissions in the training, validation and test dataset.

	Root mean squared error (RMSE)	
Training	Root mean squared error (RMSE)	4.32
	Mean absolute percentage error (MAPE)	8.96
Validation	Root mean squared error (RMSE)	6.46
	Mean absolute percentage error (MAPE)	33.99
Test	Root mean squared error (RMSE)	4.38
	Mean absolute percentage error (MAPE)	28.81

ANN model with testing datasets is better than against its training performance. The test data result indicates that the predicted value of  $\text{NO}_x$  concentration is fitted to the measured value, the correlation coefficient can meet the requirements for the estimation. Beyond that, the MAPE and RMSE results for each target were also examined. Table 2 lists the MAPE and RMSE values for the training, validation and test dataset.

## CONCLUSION

In this study, a new correlation based on ANN model approach was used for prediction of  $\text{NO}_x$  concentration ( $\mu\text{g}/\text{m}^3$ ) in Adana, which is rely on  $\text{SO}_2$  concentration ( $\mu\text{g}/\text{m}^3$ ) as an air pollutant and meteorological effecting parameters (wind rate (km/h), temperature ( $^{\circ}\text{C}$ )). The performance of the developed model ( $R^2 > 0.9568$ ) has been evaluated using measurements collected from weather stations in Adana. At the moment of writing this paper, we are not aware of other techniques for  $\text{NO}_x$  prediction in Adana.

The key merit of the proposed model in this study is an easy method to predict  $\text{NO}_x$  concentration when air monitoring station is not available. Therefore, the model could be utilized to provide air monitoring data at currently unmonitored locations in Adana, which obviate the necessity of a relatively high number of monitoring stations for describing the  $\text{NO}_x$  concentration.

## REFERENCES

1. Wang, J. and G. Song, A deep spatial-temporal ensemble model for air quality prediction. *Neurocomputing*, 314 (2018) 198-206.
2. Wen, C., et al., A novel spatiotemporal convolutional long short-term neural network for air pollution prediction. *Science of The Total Environment*, 654 (2019) 1091-1099.
3. Wang, X.-C., et al., Air pollution terrain nexus: A review considering energy generation and consumption. *Renewable and Sustainable Energy Reviews*, 105 (2019) 71-85.
4. Azid, A., et al., Prediction of the level of air pollution using principal component analysis and artificial neural network techniques: A case study in Malaysia. *Water, Air, & Soil Pollution*, 225(8) (2014) 2063.
5. Lal Benjamin, N., et al., Air quality prediction using artificial neural network. *IJCS*, 2(4) (2014) 07-09.
6. Baawain, M.S. and A.S. Al-Serihi, Systematic approach for the prediction of ground-level air pollution (around an industrial port)

- using an artificial neural network. *Aerosol and air quality research*, 14(1) (2014) 124-134.
7. Shakil, M., et al., Soft sensor for NO<sub>x</sub> and O<sub>2</sub> using dynamic neural networks. *Computers & Electrical Engineering*, 35(4) (2009) 578-586.
  8. Park, J.-H., et al., Historic and futuristic review of electron beam technology for the treatment of SO<sub>2</sub> and NO<sub>x</sub> in flue gas. *Chemical Engineering Journal*, 2018.
  9. Hoffman, S., Short-time forecasting of atmospheric NO<sub>x</sub> concentration by neural networks. *Environmental Engineering Science*, 23(4) (2006) 603-609.
  10. Shi, J.P. and R.M. Harrison, Regression modelling of hourly NO<sub>x</sub> and NO<sub>2</sub> concentrations in urban air in London. *Atmospheric Environment*, 31(24) (1997) 4081-4094.
  11. Carbajal-Hernandez, J.J., et al., Assessment and prediction of air quality using fuzzy logic and autoregressive models. *Atmospheric Environment*, 60 (2012) 37-50.
  12. Sofuoglu, S.C., et al., Forecasting ambient air SO<sub>2</sub> concentrations using artificial neural networks. *Energy Sources, Part B*, 1(2) (2006) 127-136.
  13. Unsal, F., Globalization and the mid-rank city: The case of Adana, Turkey. *Cities* 21(5) (2004) 439-449.
  14. <https://www.havaizleme.gov.tr/>, 2019
  15. <https://www.worldweatheronline.com/adana-weather-averages/adana/tr.aspx>, 2019