

Forecasting of outdoor air quality index using adaptive neuro fuzzy inference system

Haripriyan Uthayakumar, Perarasu Thangavelu, Saravanathamizhan Ramanujam*

Department of Chemical Engineering, A.C. Tech, Anna University, Chennai-600025, India

ARTICLE INFORMATION

Article Chronology:

Received 25 August 2021
Revised 07 September 2021
Accepted 20 September 2021
Published 29 September 2021

Keywords:

Adaptive neuro fuzzy inference system (ANFIS); Air pollution; Air quality index (AQI)

CORRESPONDING AUTHOR:

rsthathamizhan@gmail.com
Tel: (+91) 4422359237
Fax: (+91) 4422359105

ABSTRACT

Introduction: The estimation of air pollution level is well indicated by Air Quality Index (AQI), which tells how unhealthy the ambient air is and how polluted it can become in near future. Hence, the predictions or modeling of AQI is always of greater concern among researchers and this present study aims to develop such a model for forecasting the AQI.

Materials and methods: A combination of Artificial Neural Network (ANN) and Fuzzy logic (FL) system, called Adaptive Neuro-Fuzzy Inference System (ANFIS) have been considered for model development. Daily air quality data (PM_{2.5} and PM₁₀) and meteorological data (temperature and humidity) over a period of March 2020 to March 2021 were used as the input data and AQI as the output variable for the ANFIS model. The performances of models were evaluated based on Root Mean Square Error (RMSE), Regression coefficient (R²) and Average Absolute Relative Deviation (AARD).

Results: A total of 100 datasets is split into training (70), testing (15) and simulation (15). Gaussian and Constant membership functions were employed for classifications and the final index consisted of 81 inference (IF/THEN) rules. The ANFIS Simulation result shows an R² and RMSE value of 0.9872 and 0.0287 respectively.

Conclusion: According to the results from this study, ANFIS based AQI is a comprehensive tool for classification of air quality and it is inclined to produce accurate results. Therefore, local authorities in air quality assessment and management schemes can apply these reliable and suitable results.

Introduction

The tremendous growth of industrialization and urbanisation, particularly in developing nations, has resulted in higher levels of air that have deleterious effects on physical, biological, and economic systems. CO_x, SO_x, NO_x, SPM, O₃, particulate matter (PM_{2.5}, PM₁₀) and other air pollutants that are detrimental to the human health can be found in contaminated air [1]. The

Air Quality Index (AQI) is a daily average effect of ambient air pollutants measured at monitoring stations and calculated using an index or grading scale. They categorize air quality based on the concentrations of various contaminants and describe it with AQI is a critical tool for the general population to understand how poor or excellent the air quality is for their health in a simple manner. As a result, air quality models are crucial in estimating the quality of the

Please cite this article as: Uthayakumar H, Thangavelu P, Ramanujam S. Forecasting of outdoor air quality index using adaptive neuro fuzzy inference system. Journal of Air Pollution and Health. 2021; 6(3): 161-170.

atmosphere since they provide more domain knowledge and more accurate forecasts [3]. Air pollutant prediction models can provide early warnings, and their efficient use can reduce the number of monitoring and data collection stations by a significant amount [4].

This paper uses Adaptive Neuro Fuzzy Inference System (ANFIS) that eliminates the basic problems in Artificial Neural Network (ANN) using the learning potentiality of fuzzy system design for automatic if-then rule generation and parameter optimization. The combination of ANN and Fuzzy system is called the Neuro fuzzy system. ANFIS is the hybrid approach of neural and fuzzy system has better smoothness than ANN. Several research have been conducted to analyze air quality using the fuzzy inference system, an analytical tool was first described, but later applied to environmental challenges like air quality evaluation [5]. Later in a study, the air pollutants such as Particulate Matter ($PM_{2.5}$), Ozone (O_3), Carbon Monoxide (CO), Sulphur Dioxide (SO_2) and Nitrogen Dioxide (NO_2) were taken as inputs and the air quality index was predicted as good, moderate, or unhealthy air by the ANFIS system [6]. The authenticity of various soft computing approaches such as artificial neural network (ANN), Support Vector Machine (SVM), evolutionary ANN and SVM, the fuzzy logic model, neuro-fuzzy systems, the deep learning model to predict the AQI were reviewed and discussed [7].

Furthermore, the authenticity of fuzzy logic modeling in forecasting the daily maximum O_3 concentration levels in Santiago, Chile were demonstrated [8]. The concentration of various pollutants such as carbon monoxide (CO), nitric oxide (NO), $PM_{2.5}$ (particulate matter) and PM_{10} collected in Datong, Taiwan were predicted by using Genetic Algorithm-Back Propagation Neural Network (GA-BPNN), Support Vector Regression (SVR), Extreme Learning Machine (ELM) and WELM (Weighted Extreme Learning Machine) – ANFIS [9]. As a result, various

research using a fuzzy inference method to develop new air quality indices have been conducted. For example, in a study, the data for concentrations of sulphur dioxide were collected from 15 selected points of Konya city to predict the air quality using ANN and Fuzzy inference system [10]. From the above results, it is very clear that fuzzy inference system is an apt tool for air quality assessment.

The majority of the studies on air quality forecasting in the literature have been done for specific air contaminants. The principal focus of the present study is to create models for predicting daily air quality indices, which is based not only on the major air pollutants such as $PM_{2.5}$ and PM_{10} but also the meteorological parameters such as Temperature and Humidity on the AQI of Tamil Nadu. The developed model will furnish timely information to the public so that they can take precautionary actions to fortify their health. Therefore the objective of the present study is to forecast the AQI using ANFIS model (MATLAB R2020a) by taking $PM_{2.5}$, PM_{10} , temperature and humidity to be the input parameters and to study their combined effect on AQI as a case study conducted in Tamil Nadu, the Southern state of India. This approach can be globally used as an integrated tool for air quality assessment and management by local authorities [11].

Study area

Tamil Nadu (11.1271° N, 78.6569° E) is one of the southern states of India, spread over an area of 130,058 km² with 6.79 crore inhabitants as of 2021. Nearly 26 million automobiles ply Tamil Nadu roadways as a result of the enormous number of industries and exodus of people from bordering states. Its rapid population increase, combined its rapid economic growth, has concluded in an ever-increasing demand for transportation, striking unnecessary force on the city's current transport networks. Tamil Nadu's cities face severe transportation management issues, resulting in severe air pollution.

Materials and methods

Dataset

In this study, daily air quality $PM_{2.5}$ and PM_{10} and meteorological data (temperature and humidity) over a period of March 2020 to March 2021 was used as input data and AQI as the output variable for the ANFIS model. The AQI categories and their standard quality intervals are given in Table 1. For air quality evaluation; fuzzy linguistic terms and numerical intervals have been used to identify these AQI categories.

Modeling using ANFIS

ANFIS is used to formulate a heuristic pattern between the input-output according to the initial given fuzzy system and available input-output data pairs by employing learning methodologies [12]. The most

pre-dominantly employed Fuzzy Inference Systems (FIS) which are used in diverse applications are the Mamdani inference system and Sugeno inference system. The MATLAB R2020a package was applied to develop the models, and the modeling was performed in two stages: training and testing. In this work, 70% of the dataset is employed to train models and the remains, 15% is used for testing the models and the remaining 15% is used for Checking (Simulation). The training stage is a great step in the formation of networks. After developing the target network, the testing data is then applied to the network in order to obtain the results of the testing stage. It is essential to go for the pertinent sample for training and testing in order to reduce the prediction error [13]. In the present work ANFIS is constructed using Sugeno fuzzy model is shown in the Fig. 1.

Table 1. Air Quality Index ranges

AQI categories	Quality levels of health concern
0-50	Good
51-100	Moderate
101-200	Poor
201-300	Unhealthy
301-400	Severe
401-500	Hazardous

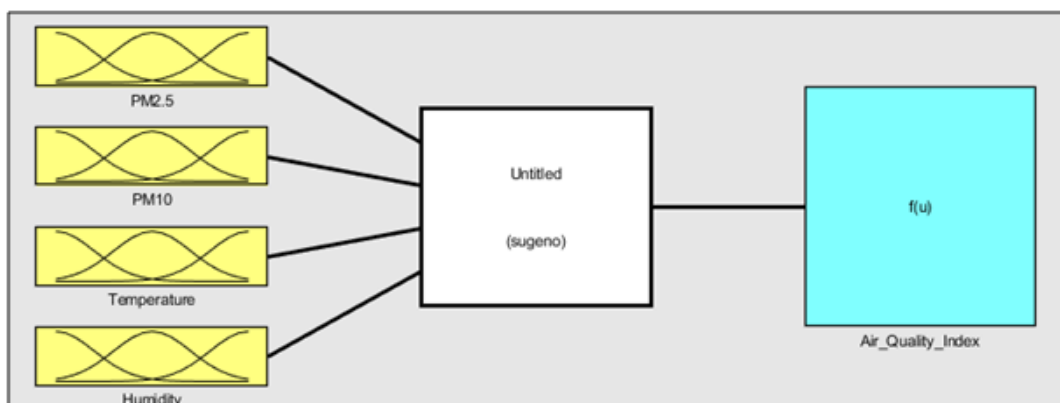


Fig. 1. Sugeno fuzzy model

The representation of rules for Sugeno fuzzy inference system is as follows.

Rule 1: If x is A₁ and y is B₁, then f₁ = p₁x + q₁y + r₁ (1)

Rule 2: If x is A₂ and y is B₂, then f₂ = p₂x + q₂y + r₂ (2)

Where, x and y are the two inputs. A, B are the membership function. p, q, and r linear parameter. 1,2, represent the number of rules [14].

Evaluation of model performances

The performances of models were evaluated based on Root Mean Square Error (RMSE), Regression coefficient (R²) and Average Absolute Relative Deviation (AARD). The calculation for the estimation of above is given in Eqs. 3, 4, and 5 [15].

$$RMSE = \sqrt{\left(\frac{1}{N}\right) \sum_{i=1}^N (X_i^{exp} - X_i^{predicted})^2} \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (X_i^{exp} - X_i^{predicted})^2}{\sum_{i=1}^N (X_i^{exp} - X^{exp})^2} \quad (4)$$

$$AARD = \frac{100}{N} \sum_{i=1}^N \left| \frac{X_i^{exp} - X_i^{predicted}}{X_i^{exp}} \right| \quad (5)$$

Results and discussion

ANFIS modeling

AQI was taken as the output parameter, as the principal focus of this analysis is to forecast the AQI of the various regions in Tamil Nadu, while PM_{2.5}, PM₁₀, temperature and humidity were taken as the input parameters. It is mandatory to fix the number and type of Membership Functions (MF), as well as the number of iterations, in order to employ the ANFIS model. In the present study the Gaussian membership function is employed for the input parameters because of their smoothness, concise notation and better prediction results compared with the other membership functions. The values of the input and the output membership functions are given (see the supporting information file). The ANFIS parameters for this study is shown in Table 2. These parameters are selected based on trial-and-error process [16]. The main advantage of Sugeno Fuzzy Inference System is that they give a more accurate relationship between a larger number of outputs and inputs.

Table 2. Details of the ANFIS for predicting the AQI

Type	Description/Value
Fuzzy structure	Sugeno-type
MF type	Gaussian (gaussmf)
Output MF	Linear
Number of fuzzy rules	81
Number of inputs	4
Number of outputs	1
Training maximum epoch number	1000

The ANFIS model may be effectively trained by normalizing or scaling the values. The boundary is [0 1], and variables are scaled in this boundary via mapping. Min Max Eq. 6 is used to normalise the outputs and inputs:

$$X_n = \frac{X' - X_{min}}{X_{max} - X_{min}} \tag{6}$$

Where X_{max} , X_{min} and X_n are maximum, minimum and normalized data for every parameter.

A total of 100 datasets is split into training (70), testing (15) and simulation (15). ANFIS parameter assessment for training, testing and checking data sets are shown in the Table 3.

Table 3. ANFIS parameters assessment for training, testing and checking data sets for AQI

Parameter	Training dataset	Testing dataset	Simulation dataset
RMSE	0.0138	0.0326	0.0287
R ²	0.9345	0.9263	0.9842
AARD	0.0943	0.0834	0.0726

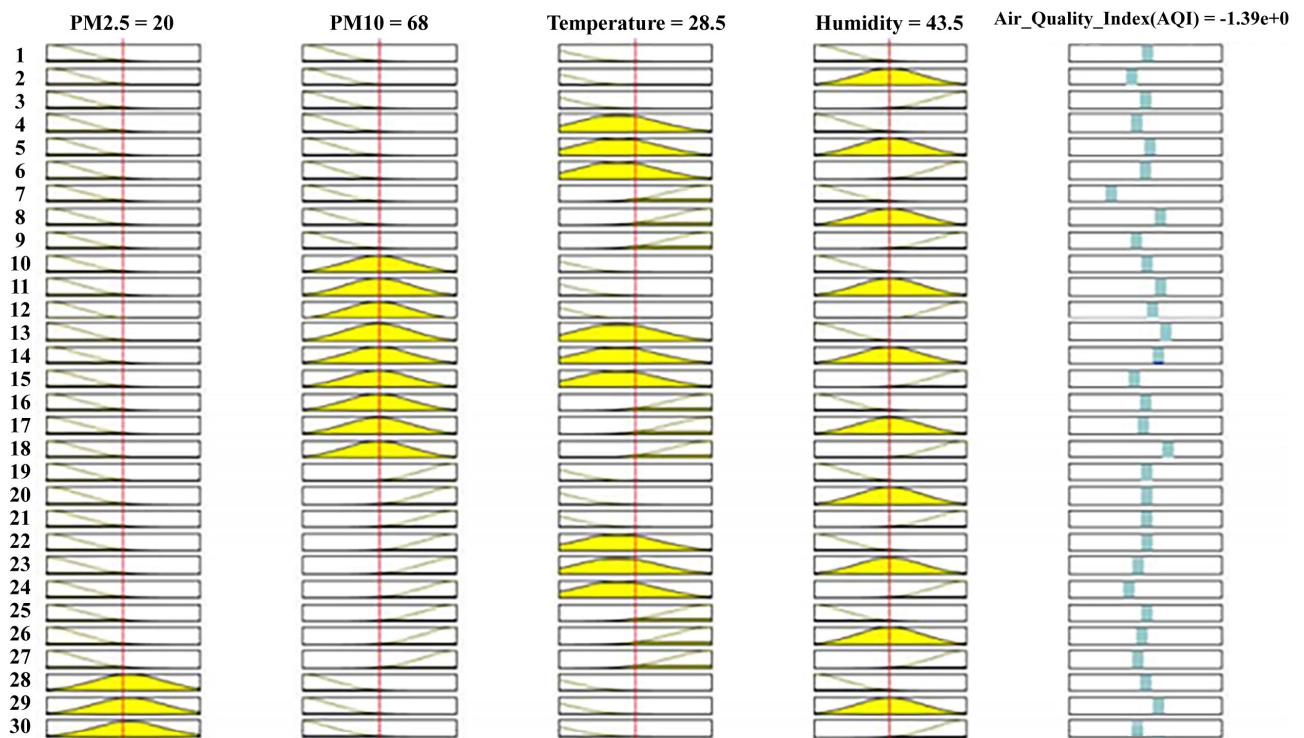


Fig. 2. Rules used in ANFIS for input parameters ($PM_{2.5}$, PM_{10} , temperature and humidity), and output parameter (AQI)

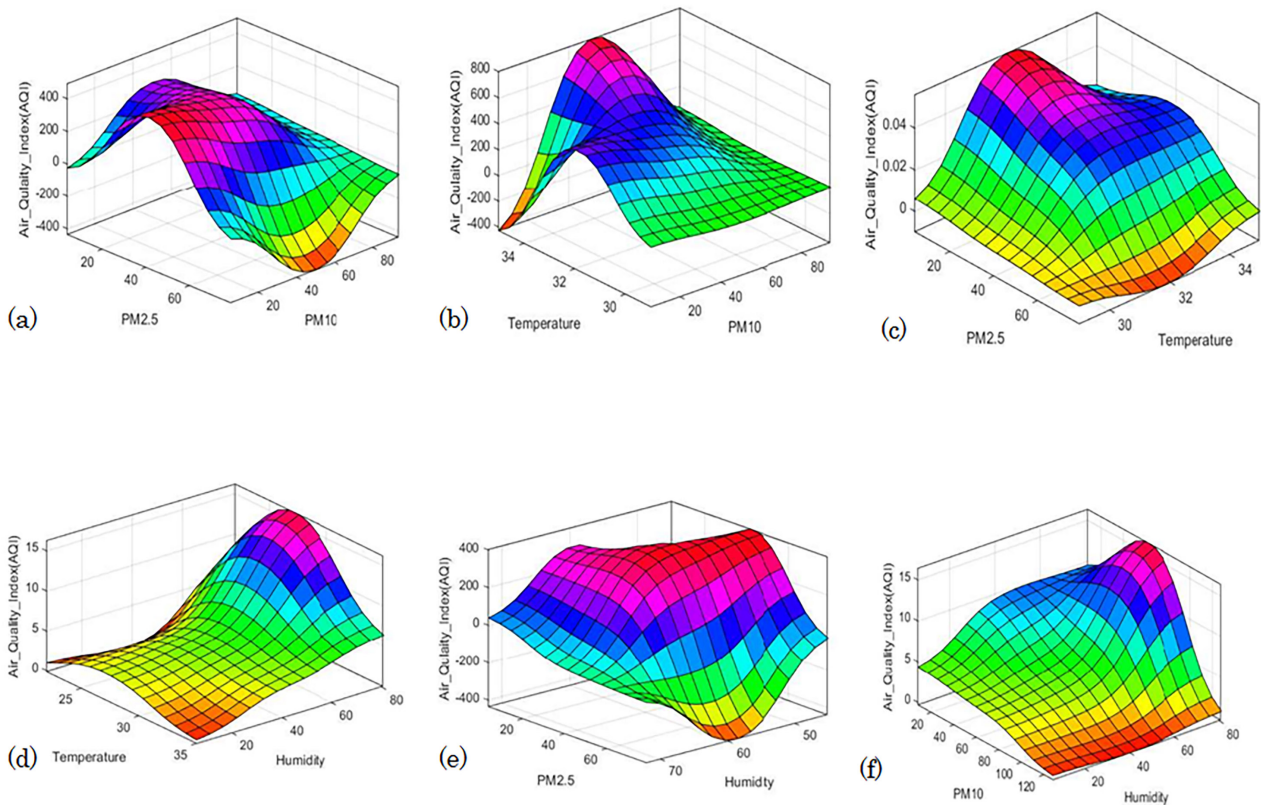


Fig. 3. Surface plot of the AQI with respect to $PM_{2.5}$, PM_{10} , temperature and humidity

Simulation using ANFIS model

ANFIS simulation rules are added in the rule editor using the “IF”, “THEN” rules. The rules are formed by the input pollutants and air quality is obtained by various stages of air pollution is shown in Fig. 2 [17, 18]. For example

1. IF $PM_{2.5}$ is LOW and PM_{10} is LOW and Temperature is MEDIUM and Humidity is LOW, THEN the air quality index is GOOD.
2. IF $PM_{2.5}$ is MEDIUM and PM_{10} is HIGH and Temperature is MEDIUM and Humidity is HIGH, THEN the air quality index is POOR.

Similarly the other rules are formed and included in the ANFIS.

The dependence of the input variables on the output response (AQI) is shown in Fig. 3. In Fig. 3a, the surface plot depicts that the AQI less than 8 can be obtained when the particulate matter $PM_{2.5}$ lies between 7 to 15 and PM_{10} lies from 8 to 20. It is evident that the AQI increases considerably with the increasing value of $PM_{2.5}$ and PM_{10} . Fig. 3b shows the surface plot for the interaction effect of PM_{10} and temperature on AQI. The surface plots showed that the AQI less than 11 can be obtained when PM_{10} lies from 7 to 12 and temperature lies from 27 °C to 30 °C. This shows that AQI increases with increasing PM_{10} . The combined effect of $PM_{2.5}$ and temperature is shown in Fig. 3c. It is observed that the increase in $PM_{2.5}$, the AQI increases and then remains constant. Fig. 3d

shows the interaction effect between temperature and humidity. It is noticed from the figure, at lower temperature, the AQI almost remains constant, further increase in the temperature causes the AQI to increase. Similar observation was also noticed from Fig. 3e which shows the interaction effect of $PM_{2.5}$ and humidity. The combined effect of PM_{10} and humidity was shown in Fig 3f. It is perceived that as shown in Fig. 3e and Fig. 3f with the increase

in PM_{10} the AQI remains almost constant and then increases to the maximum with further increase in PM_{10} and humidity [19, 20]. Table 4 depicts the Input datasets taken for the simulation along with the actual and prediction results. Fig. 4 depicts the regression analysis of the Actual vs Predicted AQI data. It is very clear that Actual and Predicted AQI data are in good correlation with each other having $R^2=0.9842$.

Table 4. ANFIS simulation predicted results

S.No	$PM_{2.5}$	PM_{10}	Temperature (°C)	Humidity (%)	AQI (actual)	AQI (predicted)	AQI (predicted range)
1	7	7	34	55	43	42.24	Good
2	17	21	34	55	64	63.17	Moderate
3	5	8	30	72	32	33.24	Good
4	5	18	33	56	37	37.73	Good
5	5	7	31	64	39	42.72	Good
6	11	17	35	45	67	69.28	Moderate
7	59	71	34	50	134	131.89	Poor
8	13	25	34	55	73	74.28	Moderate
9	10	18	34	53	42	43.75	Good
10	79	94	29	69	233	233.24	Unhealthy
11	5	7	33	52	24	25.34	Good
12	10	14	32	54	45	44.76	Good
13	17	24	33	53	123	122.86	Moderate
14	57	46	32	61	57	58.24	Poor
15	4	6	31	64	36	37.23	Good

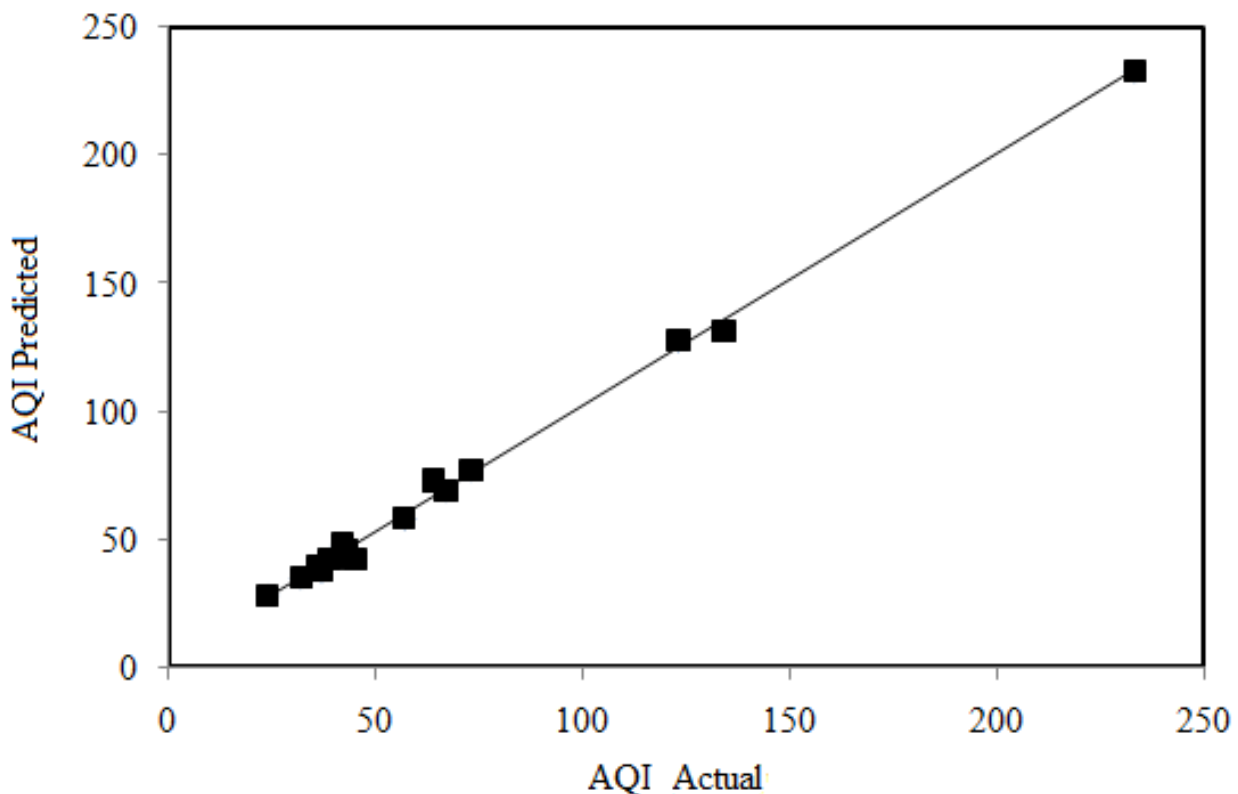


Fig. 4. Prediction of AQI using ANFIS

Conclusion

In this study, ANFIS model was used for forecasting AQI in Tamil Nadu, India. A total of 100 datasets with $PM_{2.5}$, PM_{10} , temperature and humidity was taken as input variables and AQI as output response. Gaussian and Constant membership functions were used as the input and output membership function for 1000 epochs of iterations. The proposed ANFIS model showed good performance for AQI forecasting. The simulation using ANFIS results shows higher correlation coefficient of 0.9842 and an RMSE value of 0.0287. Hence it is concluded from the present study, ANFIS model for forecasting the AQI is good and therefore can be considered, useful, reliable, and comprehensive tool in air quality assessment and management schemes.

Financial supports

This work does not supported by any financial aids.

Competing interests

The authors declare that they have no conflict of interest

Acknowledgements

Authors would like to acknowledge the Department of Chemical Engineering, Anna University for extending the facilities to carry out the work.

Ethical considerations

“Ethical issues (Including plagiarism, Informed Consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

References

- Cheng WL, Chen YS, Zhang J, Lyons TJ, Pai JL, Chang SH. Comparison of the revised air quality index with the PSI and AQI indices. *Science of the Total Environment*. 2007 Sep 1;382(2-3):191-8. <https://doi.org/10.1016/j.scitotenv.2007.04.036>.
- Li Y, Chiu YH, Lu LC. Energy and AQI performance of 31 cities in China. *Energy Policy*. 2018 Nov 1;122:194-202. <https://doi.org/10.1016/j.enpol.2018.07.037>.
- Kermani M, Bahrami Asl F, Aghaei M, Arfaenia H, Karimzadeh S, Shahsavani A. Comparative investigation of air quality index (AQI) for six industrial cities of Iran. *Studies in Medical Sciences*. 2014 Nov 10;25(9):810-9.
- Ruggieri M, Plaia A. An aggregate AQI: Comparing different standardizations and introducing a variability index. *Science of the Total Environment*. 2012 Mar 15;420:263-72. <https://doi.org/10.1016/j.scitotenv.2011.09.019>.
- Zadeh LA, Klir GJ, Yuan B. *Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers*. World Scientific; 1996. https://doi.org/10.1142/9789814261302_0021.
- Munawar S, Hamid M, Khan MS, Ahmed A, Hameed N. Health Monitoring Considering Air Quality Index Prediction Using Neuro Fuzzy Inference Model: A Case Study of Lahore, Pakistan. *Journal of Basic and Applied Sciences*. 2017 Apr 10;13:123-32.
- Rahman MM, Shafiullah M, Rahman SM, Khondaker AN, Amao A, Zahir M. Soft computing applications in air quality modeling: past, present, and future. *Sustainability*. 2020 Jan;12(10):4045. <https://doi.org/10.3390/su12104045>.
- Jorquera H, Pérez R, Cipriano A, Espejo A, Letelier MV, Acuña G. Forecasting ozone daily maximum levels at Santiago, Chile. *Atmospheric Environment*. 1998 Sep 25;32(20):3415-24. [https://doi.org/10.1016/S1352-2310\(98\)00035-1](https://doi.org/10.1016/S1352-2310(98)00035-1).
- Li Y, Jiang P, She Q, Lin G. Research on air pollutant concentration prediction method based on self-adaptive neuro-fuzzy weighted extreme learning machine. *Environmental Pollution*. 2018 Oct 1;241:1115-27. <https://doi.org/10.1016/j.envpol.2018.05.072>.
- Dursun S, Kunt F, Taylan O. Modelling sulphur dioxide levels of Konya city using artificial intelligent related to ozone, nitrogen dioxide and meteorological factors. *International journal of environmental science and technology*. 2015 Dec;12(12):3915-28. <https://doi.org/10.1007/s13762-015-0821-2>.
- Lei KS, Wan F. Applying ensemble learning techniques to ANFIS for air pollution index prediction in Macau. In *International Symposium on Neural Networks 2012* Jul 11 (pp. 509-516). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-31346-2_57.
- Jang JS. ANFIS: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics*. 1993 May;23(3):665-85. <https://doi.org/10.1109/21.256541>.
- Singh R, Kainthola A, Singh TN. Estimation of elastic constant of rocks using an ANFIS approach. *Applied Soft Computing*. 2012 Jan 1;12(1):40-5. <https://doi.org/10.1016/j.asoc.2011.09.010>.
- Ekici BB, Aksoy UT. Prediction of building energy needs in early stage of design by using ANFIS. *Expert Systems with Applications*. 2011 May 1;38(5):5352-8. <https://doi.org/10.1016/j.eswa.2010.10.021>.
- Wei M, Bai B, Sung AH, Liu Q, Wang J, Cather ME. Predicting injection profiles using ANFIS. *Information Sciences*. 2007 Oct 15;177(20):4445-61. <https://doi.org/10.1016/j.ins.2007.03.021>.
- Lei Y, He Z, Zi Y, Hu Q. Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs. *Mechanical systems and signal processing*. 2007 Jul 1;21(5):2280-94. <https://doi.org/10.1016/j.ymsp.2006.11.003>
- Buragohain M, Mahanta C. A novel

approach for ANFIS modelling based on full factorial design. *Applied soft computing*. 2008 Jan 1;8(1):609-25. <https://doi.org/10.1016/j.asoc.2007.03.010>.

18. Dastorani MT, Moghadamnia A, Piri J, Rico-Ramirez M. Application of ANN and ANFIS models for reconstructing missing flow data. *Environmental monitoring and assessment*. 2010 Jul;166(1):421-34. <https://doi.org/10.1007/s10661-009-1012-8>.

19. Al-Qaness MA, Fan H, Ewees AA, Yousri D, Abd Elaziz M. Improved ANFIS model for forecasting Wuhan City air quality and analysis COVID-19 lockdown impacts on air quality. *Environmental Research*. 2021 Mar 1;194:110607. <https://doi.org/10.1016/j.envres.2020.110607>.

20. Taylan O, Alkabaa AS, Alamoudi M, Basahel A, Balubaid M, Andejany M, Alidrisi H. Air Quality Modeling for Sustainable Clean Environment Using ANFIS and Machine Learning Approaches. *Atmosphere*. 2021 Jun;12(6):713. <https://doi.org/10.3390/atmos12060713>.