ISSN (Online) 2456-1290

International Journal of Engineering Research in Mechanical and Civil Engineering (IJERMCE)

Vol 12 Issue 01 February 2025

Prediction of NO_2, O_3, PM_{10} and $PM_{2.5}$ Concentrations in Ambient Air at Tirumala Using ANN and ANFIS

[1] T. Gireesh, [2] Dr. M. Srimurali, [3] Dr. N. Munilakshmi

[1] M Tech Scholar, Department of Civil Engineering, S.V.U.C.E, Sri Venkateswara University, Tirupati, 517502, Andhra Pradesh, India
[2] Professor, Department of Civil Engineering, S.V.U.C.E, Sri Venkateswara University, Tirupati, 517502, Andhra Pradesh, India
[3] Associate Professor, Department of Civil Engineering, S.V.U.C.E, Sri Venkateswara University, Tirupati, 517502, Andhra Pradesh, India

Abstract— *Air pollution poses a significant threat to global health and the environment. This study investigated air quality in Tirumala, a major pilgrimage destination in India, emphasizing the importance of clean air for the millions of visitors it receives annually and the preservation of its cultural and ecological integrity. Advanced machine learning techniques, Artificial Neural Networks (ANN) and Adaptive Neuro Fuzzy Inference System (ANFIS) were employed, to predict concentrations of key air pollutants: Nitrogen Dioxide (NO_2), Ozone (O_3), Particulate Matter (PM_{10}) and ($PM_{2.5}$). The models were developed and implemented using MATLAB 2023a, with a focus on optimizing model architecture by varying the number of neurons and hidden layers in the ANN and exploring different membership functions within the ANFIS framework. Model performance was rigorously evaluated using statistical metrics, specifically the coefficient of determination (R^2) and the Root Mean Square Error (RMSE), to ensure accuracy and reliability of the predictions. The study compared the predictive capabilities of both ANN and ANFIS, identifying the best-performing model for each pollutant. Results indicated that both ANN and ANFIS successfully predicted pollutant concentrations, with ANFIS demonstrating a slight advantage in predictive accuracy. Crucially, the predicted levels of all measured pollutants remained within the permissible limits set by India's National Ambient Air Quality Standards (NAAQS). ANN achieved R^2 values of 0.9390, 0.9387, 0.9234, 0.9114 and RMSE values of 6.601, 6.741, 5.119,4.724 between observed and predicted values of NO2, O3, PM10 and PM2.5 respectively. ANFIS attained R2 values of 0.9557, 0.9435, 0.9358, 0.9269 and RMSE values of 4.439, 6.024, 4.598, 4.214 between observed and predicted values of NO2, O3, PM10 and PM2.5 respectively.*

Keywords: Artificial Neural Networks (ANN), Adaptive Neuro Fuzzy Inference System (ANFIS), Air Pollution, Pollutant.

I. INTRODUCTION

Industrialization and technological progress, while driving societal advancement, have inadvertently contributed to a growing environmental crisis, with air pollution emerging as a particularly pressing concern. The complex interplay of natural and anthropogenic sources, ranging from volcanic eruptions and wildfires to industrial emissions, vehicular exhaust, and agricultural practices, results in a diverse mix of pollutants that threaten both human health and ecosystem integrity. Exposure to pollutants nitrogen dioxide (NO_2), Ozone (O_3), Particulate Matter (PM_{10} and $PM_{2.5}$) is linked to a spectrum of adverse health outcomes, from respiratory ailments and cardiovascular disease to neurological disorders and cancer. Children, the elderly, and individuals with pre-existing conditions are especially vulnerable to the detrimental effects of poor air quality. Effective air quality management relies heavily on the accurate prediction of pollutant concentrations. Traditional chemical transport and dispersion models, while valuable, often face limitations in computational efficiency and complexity. Consequently, machine learning techniques have emerged as a promising alternative, offering the potential for faster and more accurate predictions. This study investigates the application of two such techniques, Artificial Neural Networks (ANNs) and Adaptive Neuro Fuzzy Inference Systems (ANFIS), for predicting air pollutant concentrations. ANNs, inspired by the structure of the human brain, excel at learning complex non-linear relationships from data. ANFIS, on the other hand, integrates the learning capabilities of neural networks with the interpretability and reasoning power of fuzzy logic, offering a hybrid approach that can capture both data-driven patterns and expert knowledge. By evaluating the performance of ANNs and ANFIS in predicting key air pollutants, this research aims to contribute to the development of more effective air quality monitoring and forecasting tools, ultimately supporting efforts to mitigate air pollution and protect public health.

II. LITERATURE REVIEW

Ahmad Shadab et al. (2023) predicted daily Air Quality Index (AQI) a year in advance using Feed Forward, Collaborative Filtering, and Linear Regression Neural Networks. Their analysis of key pollutants like NO_2, O_3, $PM_{2.5}$, and PM_{10} revealed that the Feed Forward Neural Network (FF-NN) achieved the highest accuracy, with an

RMSE of 26.79. This highlights the potential of FF-NN for long-term AQI forecasting. Aryan Agarwal et al. (2023) explored the use of Artificial Neural Networks (ANN), linear regression, Support Vector Machines (SVM), and Random Forest for predicting air pollution levels in India. Their findings indicated that the linear ANN achieved the highest prediction accuracy of 97%, suggesting its suitability for regional air pollution forecasting. Suri Raunaq et al. (2022) focused on predicting AQI in six Indian cities using ANNs and Gaussian Process Regression (GPR). They emphasized the importance of optimizing ANN architecture based on dataset characteristics and highlighted the impact of data quality and quantity on model performance. The reported correlation coefficients were 0.9843 for ANN and 0.9611 for GPR, demonstrating the effectiveness of both techniques. Ruchita Nehete et al. (2021) compared Linear Regression, Random Forest, Decision Tree, and ANN for $PM_{2.5}$ concentration prediction. While Linear Regression showed limited accuracy at 72.78%, Random Forest, Decision Tree, and ANN achieved high accuracy levels, approaching 99.97%, 99.98%, and 99.81% respectively. This study underscores the potential of ensemble learning methods like Random Forest for capturing complex relationships in $PM_{2.5}$ data. Heydar et al. (2019) investigated the ability of ANNs to predict hourly concentrations of various air pollutants and two air quality indices (AQI and AQHI) in Ahvaz, Iran. Their model achieved an overall correlation coefficient (R^2) of 0.87, demonstrating the capability of ANNs to capture temporal variations in air pollution. The study also noted that the accuracy of the ANN model was influenced by the frequency of high-pollution events. Saxena et al. (2018) compared different neural network topologies (FFNN, LRNN, NARX, RBFN) for predicting Respirable Suspended Particulate Matter (RSPM) levels. Their results indicated that the Radial Basis Function Neural Network (RBFN) provided the most accurate predictions, suggesting its suitability for RSPM forecasting in industrial areas. Srivastava et al. (2018) implemented various classification and regression techniques, including Linear Regression, SDG Regression, Random Forest Regression, Decision Tree Regression, Support Vector Regression (SVR), ANNs, Gradient Boosting Regression, and Adaptive Boosting Regression, to forecast the Air Quality Index (AQI) of major pollutants such as $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, and O_3. They concluded that Support Vector Regression and Artificial Neural Networks were the most effective techniques for predicting air quality in New Delhi. Zhao, X. et al. (2018) explored the use of a deep learning method to predict Air Quality Classification (AQC) in three industrial cities in the United States. For comparison, Support Vector Machine (SVM) and Random Forest (RF) approaches were also applied to AQC forecasting. A Recurrent Neural Network (RNN) was employed as the primary prediction model. The study concluded that RNN has the advantage of sequentially memorizing previous air quality information, providing better predictive capabilities compared to the other machine learning models tested. The RNN model achieved the highest accuracy among the techniques tested, with 76.44% in Los Angeles and 80.27% in Atlanta. However, in Houston, SVM outperformed RNN and RF, achieving an accuracy of 76.71%. Alimissis et al. (2018) investigated the effectiveness of two distinct interpolation methods—Artificial Neural Networks (ANNs) and Multiple Linear Regression (MLR)—for estimating air pollution levels. Using data collected from a network of air quality monitoring stations across the greater Athens metropolitan area, the researchers compared the performance of these techniques. Their findings indicated that Feed Forward Neural Networks (FFNNs) offered a clear advantage over the traditional linear MLR approach. This superior performance was attributed to the FFNN's ability to better capture the intricate and often non-linear spatial variations in air pollution concentrations, a characteristic that simpler linear models struggle to represent accurately. Essentially, the more complex architecture of the FFNN allowed it to learn and model the spatial patterns of pollution more effectively than the MLR method.

III. METHODOLOGY

A. Collection of Data

The concentrations of air pollutants Nitrogen Dioxide (NO_2), Ozone (O_3), Particulate Matter (PM_{10}), ($PM_{2.5}$) and Meteorological parameters Ambient Temperature (AT), Wind Speed (WS), Wind Direction (WD), Relative Humidity (RH), Solar Radiation (SR), and Barometric Pressure (AP) were collected for the study area for the period 2018 2023 from APPCB, Tirupati.

B. Normalization of Data

After data collection and weather parameter selection, data normalization is essential to avoid convergence issues caused by raw data's varying scales. A consistent input feature range minimizes bias. Normalization also speeds training by placing all features on a similar scale, especially helpful when inputs vary widely. Here, all meteorological data was normalized to a 0-1 range by scaling each value relative to its parameter's minimum and range. The formula used for the normalization of the data is given in equation 1.

$$X = (x_i - x_{min}) / (x_{max} - x_{min}) \qquad \text{eqn. 1}$$

Where, X = Normalized Value, x_i = input value, x_{max} = Maximum value of the input parameters x_{min} = Minimum value of the input parameters.

C. Splitting the Dataset

The dataset was randomly split into training (70%), validation (15%), and testing (15%) sets, following the proportions used by Suri et al. (2022). This partitioning allowed for model training, tuning, and final performance evaluation.

D. Artificial Neural Networks (ANN) Methodology

Flow chart of methodology adopted for the development of ANN Architecture is presented in Figure 1.

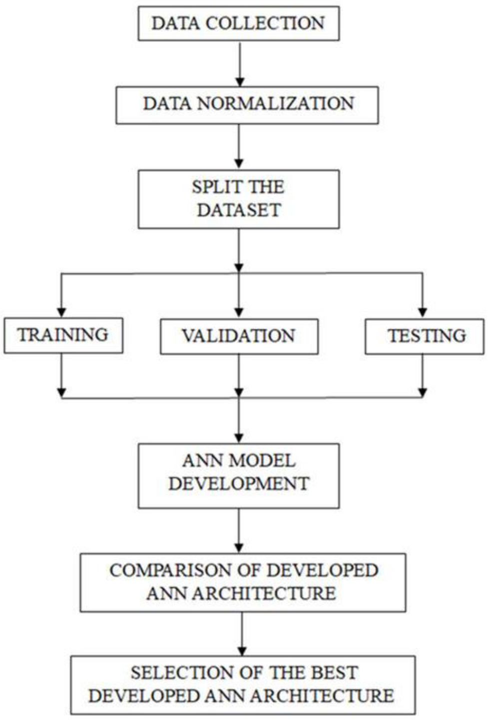

Figure 1. Flow chart of methodology adopted for the development of ANN Architecture

Artificial neural networks (ANNs) are computational systems inspired by the human brain. They function as parallel, distributed processing networks capable of learning and adapting, unlike traditional programmed systems. ANNs are trained, not explicitly programmed, to recognize patterns, form associations, and model relationships within data. This adaptability allows them to be applied across various domains, including prediction. In this study a Feed-Forward Backpropagation Neural Network (FF-BPNN) was implemented in MATLAB 2023a. The FF-BPNN architecture, with its unidirectional flow of data from input to output layers, offers simplicity and ease of understanding. The backpropagation algorithm, used for error correction, allows for efficient training and convergence via gradient descent. The FF-BPNN's ability to model non-linear relationships, coupled with its robustness and relatively low computational demands compared to more complex architectures like Recurrent Neural Networks (RNNs), makes it well-suited for this application. Its scalability also provides flexibility for handling different problem complexities. The FF-BPNN operates in two phases: forward propagation and backpropagation. In forward propagation, input data travels through the network. Each input is multiplied by a weight, summed (often with a bias term), and then passed through an activation function, introducing non-linearity. This process continues through hidden layers to the output layer, generating a prediction. During backpropagation, the difference between the predicted and actual values (the error) is calculated. This error is then propagated back through the network to calculate the gradient of the error with respect to each weight. These weights are adjusted to minimize the error, typically using an optimization algorithm like gradient descent. This iterative process continues until the weights converge, minimizing the overall error. Before application, the network undergoes a training phase. Here, the network's output is compared to the target output, and the resulting error is used to adjust the network's weights and biases via a training algorithm. This training enables the network to learn the underlying data patterns. Neural network architecture, specifically the number of hidden layers and nodes within each layer, plays a crucial role in determining model performance. These architectural parameters define the network's size, which directly influences its learning capacity. A larger network, with more nodes and layers, offers a greater number of free parameters, enabling the model to potentially capture more complex relationships within the data. Selecting an optimal architecture is therefore essential for effective learning and convergence. However, there is no universally applicable formula for determining the ideal number of hidden layers or nodes. The optimal configuration is highly problem-specific, depending on factors such as the nature and complexity of the data, the target variable being predicted, and the underlying relationships being modeled. Consequently, an experimental approach is often adopted. This involves systematically testing various combinations of hidden layers and nodes to identify the configuration that minimizes prediction error while simultaneously maximizing generalization performance. This iterative process allows researchers to empirically determine the most suitable architecture for the specific problem at hand. Performance is evaluated using the Coefficient of Determination (R^2) which provide a comprehensive assessment of the network's predictive capabilities.

E. Pearson Correlation Coefficient

The Pearson correlation coefficient is a statistical measure that quantifies the strength and direction of the linear relationship between two variables. In essence, Pearson's correlation coefficient quantifies the degree to which two variables change together relative to their individual variability. The formula to find the Pearson's correlation coefficient is given in Equation 2.

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{(n\sum X^2 - (\sum X)^2)(n\sum Y^2 - (\sum Y)^2)}}$$

ISSN (Online) 2456-1290

International Journal of Engineering Research in Mechanical and Civil Engineering (IJERMCE)

Vol 12 Issue 01 February 2025

Eqn. 2

Where:

n is the number of data points,

$\sum X$ is the sum of all values of X,

$\sum Y$ is the sum of all values of Y,

$\sum X \sum Y$ is the sum of the products of corresponding values of X and Y,

$\sum X2$ is the sum of the squares of X,

$\sum Y2$ is the sum of the squares of Y.

Positive correlation coefficient value signifies that as one variable increase so does the other variable. Negative correlation coefficient value signifies that as one variable increases the other variable decreases. Zero correlation coefficient value signifies that movement in one variable cannot be predicted from the other variable. The Pearson correlation coefficient values between meteorological parameters Ambient Temperature (AT), Relative Humidity (RH), Wind Speed (WS), Wind Direction (WD), Solar Radiation (SR), Barometric Pressure (BP) and Pollutants NO_2, O_3, PM_{10} and $PM_{2.5}$ are given in Table 1.

Table 1: Pearson Correlation Coefficient Values

	NO_2	O_3	PM_{10}	$PM_{2.5}$
AT	-0.143	0.193	-0.130	-0.321
RH	-0.380	-0.263	-0.257	-0.148
WS	0.181	0.261	0.145	0.140
WD	0.116	0.125	0.106	0.180
SR	-0.125	0.191	-0.126	-0.177
BP	0.170	0.177	0.186	0.196

Ambient Temperature shows a weak negative correlation with pollutants NO_2, PM_{10} and $PM_{2.5}$ whereas it shows weak positive correlation with O_3. This suggests that as temperature increases, the concentration of pollutants NO_2, PM_{10} and $PM_{2.5}$ tends to decrease and with O_3 it is vice versa. Relative Humidity exhibits a weak negative correlation with NO_2, O_3, PM_{10} and $PM_{2.5}$. This indicates that higher the humidity levels lower the values of pollutants NO_2, O_3, PM_{10} and $PM_{2.5}$. Wind Speed shows a weak positive correlation with O_3, NO_2, PM_{10}, and $PM_{2.5}$. This indicates that as wind speed increases O_3, NO_2, PM_{10}, and $PM_{2.5}$ pollutant concentrations increases. Solar Radiation shows a weak positive correlation with NO_2, PM_{10} and $PM_{2.5}$ and weak negative correlation with O3. From this it can be understood that as solar radiation increases there will be a slight increase in the concentrations of pollutants NO_2, O_3, PM_{10} and $PM_{2.5}$. Wind Direction and Barometric Pressure have weak positive correlation with pollutants NO_2, O_3, PM_{10} and $PM_{2.5}$. This indicates that as Wind Direction and Barometric Pressure increases there will be a slight increase in the concentrations of pollutants NO_2, O_3, PM_{10} and $PM_{2.5}$.

F. Adaptive Neuro Fuzzy Inference System (ANFIS)

Methodology

ANFIS is the combination of the two soft-computing methods: ANN and Fuzzy Inference System, which was first introduced by Jyh-Shing Roger Jang in 1992. It works in Sugeno fuzzy inference system and its structure is similar to the multilayer feedforward neural network structure, except that the links in ANFIS indicate the signals flow direction and there are no associated weights with the links. To simplify the concept, two rules in the method of "If–Then" for the Sugeno model will be considered with *x* and *y* as inputs and *f* as output. ANFIS System incorporates the human like reasoning style of fuzzy systems through the use of fuzzy sets and a linguistic model consisting of a set of IF-THEN fuzzy rules. The main strength of ANFIS System is that they are universal approximators with the ability to solicit interpretable IF-THEN rules. The strength of ANFIS System involves two contradictory requirements in fuzzy modelling: interpretability versus accuracy. The Hybrid models are considered to be more advantageous for air quality prediction. The combination of more than one soft computing techniques forms hybrid soft computing technique. A number of hybrid soft computing techniques applied in assessment of air quality prediction efficiently. A hybrid soft computing technique with the combination of ANN along with Fuzzy Logic can be very effective for air pollution prediction and time series analysis. In the present work, ANN and a hybrid model which is a combination of ANN and Fuzzy Logic was applied. The use of these two soft computing techniques is basically due to following advantages of these techniques. Flowchart of methodology adopted for ANFIS model development is presented in Figure 2.

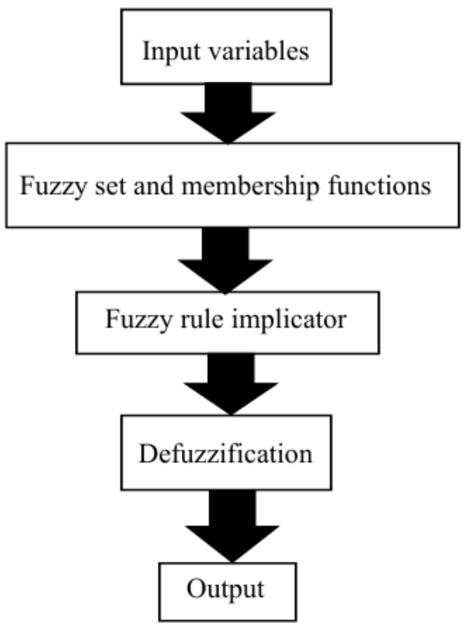

Figure 2. Flowchart of methodology adopted for development of ANFIS model

ANFIS System can be viewed as a 3-layer feed forward

IFERP connecting engineers...developing research

ISSN (Online) 2456-1290

International Journal of Engineering Research in Mechanical and Civil Engineering (IJERMCE)

Vol 12 Issue 01 February 2025

neural network. The first layer represents input variables, the middle (hidden) layer represents fuzzy rules and the third layer represents output variables. Fuzzy sets are encoded as (fuzzy) connection weights. It is not necessary to represent a fuzzy system like this to apply a learning algorithm to it. However, it can be convenient, because it represents the data flow of input processing and learning within the model. The ANFIS system development in MATLAB proceeded in three distinct phases. First, input variables were defined, with meteorological parameters—ambient temperature, wind speed, wind direction, relative humidity, barometric pressure, and solar radiation—selected as predictors for pollutant concentration. Second, fuzzy rules linking these meteorological inputs to the output pollutant concentration were formulated within the rule editor. Finally, membership functions for each input and output variable were defined, including assigning names and values. This quantification of linguistic variables enabled the ANFIS model to process the meteorological data and generate pollutant concentration predictions. The performance of the membership functions is evaluated by the performance evaluation metric R^2 by using the formula given in Eqn 3.

$$R^2 = \left(\frac{\sum_1^n \left[\left(\hat{Y}_i - \bar{\hat{Y}}_i \right) * \left(Y_i - \bar{Y}_i \right) \right]}{\left(\sum_1^n \left[\left(\hat{Y}_i - \bar{\hat{Y}}_i \right)^2 * \left(Y_i - \bar{Y}_i \right) \right] \right)^{0.5}} \right)^2$$

Eqn 3

where y_i and $\bar{y_i}$ are the measured concentrations and average of measured concentrations for a pollutant, respectively. $\hat{y_i}$ and $\bar{\hat{y_i}}$ are predicted concentrations and average of predicted concentrations for a pollutant, respectively.

IV. RESULTS AND DISCUSSION

A. Results of Artificial Neural Networks (ANN)

Artificial Neural Network (ANN) models are developed with different architectures containing different number of neurons and hidden layers. The models with best performance in different phases of model development such as training, validation, testing and overall are selected and presented for prediction of pollutants concentrations NO_2, O_3, PM_{10} and $PM_{2.5}$ are presented in Table 2.

Table 2: Proposed best network models for the prediction of pollutants NO_2, O_3, PM_{10} and $PM_{2.5}$ using ANN

S. No	Pollutant	Network	R^2			
			Training	Validation	Testing	Overall
1	NO_2	6-6-32-1	0.9293	0.9273	0.9260	0.9249
2	O_3	6-10-50-1	0.9153	0.9063	0.9158	0.9088
3	PM_{10}	6-10-38-1	0.9250	0.9266	0.9180	0.9258
4	$PM_{2.5}$	6-10-43-1	0.9063	0.9086	0.9057	0.9044

B. Results of Adaptive Neuro Fuzzy Inference System (ANFIS)

Developed ANFIS models that achieved better performance metrics in training, validation, testing and overall phases are selected predict NO_2, O_3, PM_{10} and $PM_{2.5}$ pollutant concentrations in ambient air. The membership functions that achieved best evaluation metric R^2 for prediction of pollutants concentrations NO_2, O_3, PM_{10} and $PM_{2.5}$ using ANFIS are presented in Table 3.

Table 3: Proposed best membership functions for the prediction of pollutants NO_2, O_3, PM_{10} and $PM_{2.5}$ using ANFIS

S. No	Pollutant	Membership Function	R^2			
			Training	Validation	Testing	Overall
1	NO_2	trimf	0.9306	0.9443	0.9234	0.9421
2	O_3	trapmf	0.9151	0.9078	0.9221	0.9221
3	PM_{10}	gbellmf	0.9216	0.9341	0.9337	0.9382
4	$PM_{2.5}$	gbellmf	0.9171	0.9125	0.9281	0.9112

The developed ANN and ANFIS models presented in Table 2 and 3 respectively are used to predict pollutants concentrations NO_2, O_3, PM_{10} and $PM_{2.5}$ of one year. The predicted values are compared with the observed values to validate the results of the both ANN and ANFIS models developed. The graphs plotted between predicted and observed values of pollutants concentrations NO_2, O_3, PM_{10} and $PM_{2.5}$ are presented in Figure 3, 4, 5 and 6 respectively.

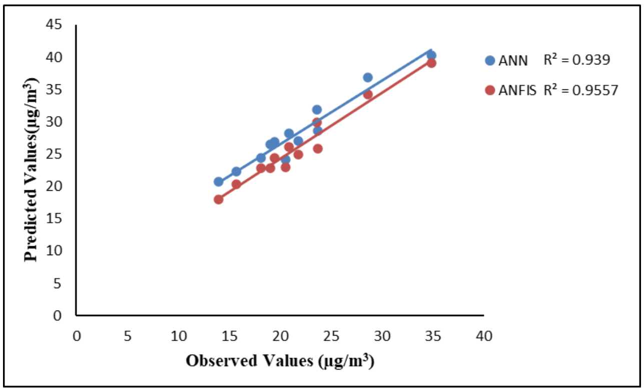

Figure 3. Regression graph plotted between observed and predicted values of NO_2 Using ANN and ANFIS

The R^2 values presented in Figure 3 suggest a marginal performance advantage of the ANFIS model over the ANN model in predicting NO_2 concentrations.

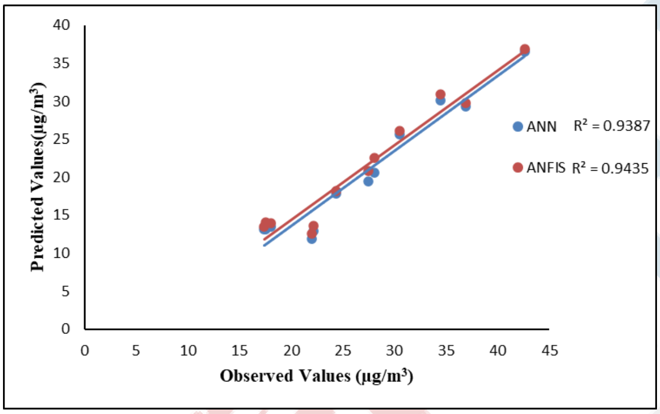

Figure 4. Regression graph plotted between observed and predicted values of O3 Using ANN and ANFIS

From R^2 values in Figure 4 it is evident that ANFIS is marginally better compared to ANN for prediction of O_3 values.

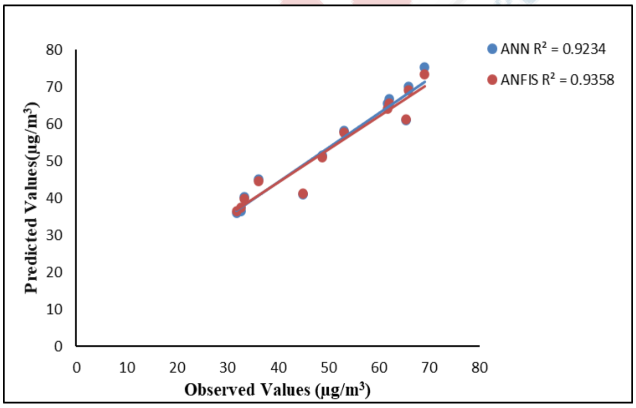

Figure 5. Regression graph plotted between observed and predicted values of PM10 Using ANN and ANFIS

The R^2 values depicted in Figure 5 indicate a marginal difference in prediction accuracy of ANFIS over ANN in predicting PM_{10} concentrations.

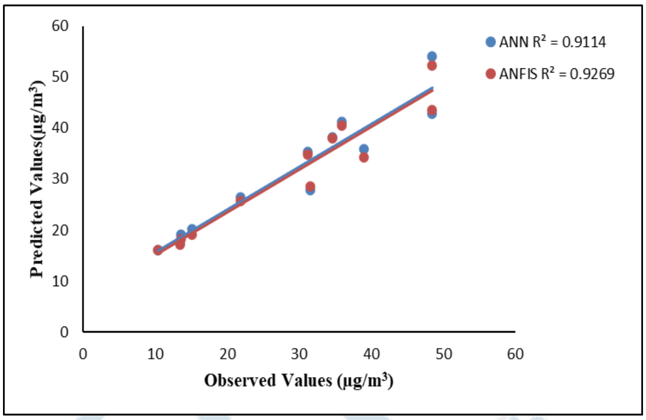

Figure 6. Regression graph plotted between observed and predicted values of PM2.5 Using ANN and ANFIS

Based on the R^2 values presented in Figure 6, it can be concluded that ANFIS exhibits a marginal performance advantage over ANN in the prediction of $PM_{2.5}$ concentrations.

C. Root Mean Square Error (RMSE)

RMSE of the best developed ANN and ANFIS models to predict the concentrations of NO_2, O_3, PM_{10}, and $PM_{2.5}$ which are presented in Tables 2 and 3 respectively are calculated and presented in Table 4. The best model among ANN and ANFIS is decided by comparing the RMSE values, lower the value better the model's performance. RMSE values are calculated using the formula given in equation 4.

$$RMSE = \sqrt{\frac{\sum_1^n (\text{Meas} - \text{Pred})^2}{n}}$$

Eqn 4.

Where,
n is the number of data values,
Meas is the measured data,
Pred is the predicted data.

Table 4: RMSE values of the best selected ANN and ANFIS models used for the prediction of NO_2, O_3, PM_{10} and $PM_{2.5}$ concentrations

Pollutants	Models	Networks (ANN) /Membership Functions (ANFIS)	RMSE
NO_2	ANN	6-6-32-1	6.601
	ANFIS	trimf	4.439
O_3	ANN	6-10-50-1	6.741
	ANFIS	trapmf	6.024

Pollutants	Models	Networks (ANN) /Membership Functions (ANFIS)	RMSE
PM_{10}	ANN	6-10-38-1	5.119
	ANFIS	gbellmf	4.598
$PM_{2.5}$	ANN	6-10-43-1	4.724
	ANFIS	gbellmf	4.214

From Table 4 it can be concluded that ANFIS model with the 'trimf' (triangular membership function) achieved the lowest RMSE, indicating the best prediction accuracy for NO_2. ANFIS with the 'trapmf' (trapezoidal membership function) had marginally lowest RMSE, suggesting its suitability for O_3 prediction. ANFIS with the 'gbellmf' (generalized bell-shaped membership function) outperformed the ANN model in prediction of PM_{10} and $PM_{2.5}$. RMSE values of the ANFIS models used for prediction of NO_2, O_3, PM_{10}, $PM_{2.5}$ concentrations have marginally better accuracy than ANN models. Hence, it can be concluded that ANFIS models are marginally better for prediction of NO_2, O_3, PM_{10}, $PM_{2.5}$ compared to ANN.

V. CONCLUSIONS

The study successfully developed Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) models to predict the concentrations of NO_2, O_3, PM_{10}, and $PM_{2.5}$ pollutants. ANFIS models demonstrated superior accuracy, with the best-performing membership functions being triangular (trimf) for NO_2 with R^2 value of 0.9557 and RMSE value of 4.439, trapezoidal (trapmf) for O_3 with R^2 value of 0.9435 and RMSE value of 6.024, and generalized bell-shaped (gbellmf) for PM_{10} with R^2 value of 0.9358 and RMSE value of 4.598 and generalized bell-shaped (gbellmf) for $PM_{2.5}$ with R^2 value of 0.9269 and RMSE value of 4.214. To significantly reduce air pollution levels in Tirumala, increase the usage of electric vehicles and vehicles that use cleaner fuels such as CNG. Implement traffic restrictions, such as limited entry permits for vehicles during peak seasons, to reduce vehicular emissions. Implement a policy that allows only vehicles with valid Pollution Under Control (PUC) certificates issued by the Regional Transport Office.

REFERENCES

[1] Ahmad, Shadab & Ahmad, Tarique. (2023). AQI prediction using layer recurrent neural network model: a new approach. Environmental monitoring and assessment. 195. 1180. 10.1007/s10661-023-11646-3.

[2] Aryan Agarwal, Pratik Dighole, Abhishek Sabnis, Dhananjay Thosar, Madhuri Mane (2023) Detection and Predicting Air Pollution Level in A Specific City Using Machine Learning Models International Journal of Creative Research Thoughts (IJCRT) Volume 11, Issue 3 March 2023 | Issn: 2320-2882.

[3] Suri, Raunaq & Jain, Ajay & Kapoor, Nishant & Kumar, Aman & Arora, Harish & Kumar, Krishna & Jahangir, Hashem. (2023). Air Quality Prediction - A Study Using Neural Network Based Approach. Journal of Soft Computing in Civil Engineering. 7. 99-113. 10.22115/SCCE.2022.352017.1488.

[4] Ruchita Nehete, D. D. Patil (2021). Air Quality Prediction Using Machine Learning International Journal of Creative Research Thoughts (IJCRT) | Volume 9, Issue 6 June 2021 | ISSN: 2320-2882.

[5] Heydar Maleki, & Sorooshian, Armin & Goudarzi, Gholamreza & Baboli, Zeynab & Tahmasebi Birgani, Yaser & Rahmati, Mojtaba. (2019). Air pollution prediction by using an artificial neural network model. Clean Technologies and Environmental Policy. 21. 10.1007/s10098-019-01709-w.

[6] Saxena, Shikha & Mathur, Anil & Dwivedi, A. (2018). Feed-Forward Artificial Neural Network for Air Quality Index Prediction in Kota City. Conference Proceedings: Indian Engineering Congress At: Udaipur

[7] Srivastava, Chavi & Singh, Shyamli & Singh, Amit. (2018). Estimation of Air Pollution in Delhi Using Machine Learning Techniques. 304-309. Conference procedings: 2018 International Conference on Computing, Power and Communication Technologies (GUCON) 10.1109/GUCON.2018.8675022.

[8] Zhao, X. & Zhang, R. & Wu, J.-L & Chang, P.-C. (2018). A deep recurrent neural network for air quality classification. Journal of Information Hiding and Multimedia Signal Processing. 9. 346-354.

[9] A. Alimissis, K. Philippopoulos, C.G. Tzanis, D. Deligiorgi, (2018) Spatial estimation of urban air pollution with the use of artificial neural network models, Atmospheric Environment, Volume 191, 2018, Pages 205-213, ISSN 1352-2310, https://doi.org/10.1016/j.atmosenv.2018.07.058.

[10] Anna Li, A Xiao Xu (2018) A New PM2.5 Air Pollution Forecasting Model Based on Data Mining and BP Neural Network Model, Conference Proceedings: Proceedings of the 2018 3rd International Conference on Communications, Information Management and Network Security (CIMNS 2018). Atlantis Press, 2018/11 Pages 110-113, 2352-538X. https://doi.org/10.2991/cimns-18.2018.25

[11] Fabio Biancofiore, Marco Verdecchia, Piero Di Carlo, Barbara Tomassetti, Eleonora Aruffo, Marcella Busilacchio, Sebastiano Bianco, Sinibaldo Di Tommaso, Carlo Colangeli, (2015) Analysis of surface ozone using a recurrent neural network,Science of The Total Environment, Volume 514, 2015, Pages 379-387, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2015.01.106.

[12] Geetha Mani, Joshi Kumar Viswanadhapalli and Albert Alexander Stonier (2021). Prediction and forecasting of air quality index in Chennai using regression and ARIMA time series models Journal of Engg. Research Vol.10 No. (2A) pp. 179-194. Doi: 10.36909/jer.10253.

[13] Murtaza Roondiwala, Harshal Patel, and Shraddha Varma. 2017. 'Predicting Stock Prices using LSTM'. International Journal of Science and Research 6(4).

[14] Oprea, M., Popescu, M., Mihalache, S. and Dragomir, E. (2017). Data Mining and ANFIS Application to Particulate Matter Air Pollutant Prediction. A Comparative Study. In Proceedings of the 9th International Conference on Agents

and Artificial Intelligence - Volume 1: ICAART; ISBN 978-989-758-220-2; ISSN 2184-433X, SciTePress, pages 551-558. DOI: 10.5220/0006196405510558.

[15] Rahman, M. M., Shafiullah, M., Rahman, S. M., Khondaker, A. N., Amao, A., & Zahir, M. H. (2020). Soft computing applications in air quality modeling: Past, present, and future. Sustainability, 12(10), 4045.

[16] Ritik Sharma, Gaurav Shilimkar, Shivam Pisal (2021) Air Quality Prediction by Machine Learning International Journal of Scientific Research in Science and Technology Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com) Doi: https://doi.org/10.32628/IJSRST218396.

[17] Tanisha Madan, Shrddha Sagar, Deepali Virmani (2020) Air Quality Prediction using Machine Learning Algorithms –A Review 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN) 78-1-7281-8337-4/20/$31.00 2020 IEEE | Doi: 10.1109/ICACCCN51052.2020.9362912.