Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 03, March-2025

An Analysis of Ghaziabad City to Predict the Air Quality Index using Artificial Neural Network (ANN)

Lokesh Kumar
Research Scholar, Department of Mathematics,
NAS College, Meerut, Uttar Pradesh, India 250003

Gaurav Kumar
Professor, Department of Mathematics,
NAS College, Meerut, Uttar Pradesh, India 250003

Corresponding Author: Lokesh Kumar

Abstract

PM10 is a major air pollutant that contributes to air pollution. Using data from 2019 to 2023, this study estimated the contaminant's influence on human health and the atmosphere using ANN, a popular learning approach. Using SPSS programming, the Air Monitoring Center of the Pollution Control Board of Uttar Pradesh (UPPCB) completed the necessary modeling and optimization processes and gathered data on the Ghaziabad industrial center. The resulting air quality assessment findings were subjected to an MLP (multilayer perceptron) process before comparison with the real data. Additionally, especially during times of high output, the province of Ghaziabad's air quality index (AQI) levels have occasionally exceeded the allowable limit.

Keywords: **ANN, Air Pollution, AQI, Multilayer Perceptron.**

1. INTRODUCTION

Air pollutants such as SO2 (sulfur dioxide), CO2 (carbon dioxide), and PM10 (particulate matter) are rising as industrialization advances. Both the environment and human well-being are severely harmed by this rise, which has a detrimental effect on public economies. By doing thus, these air pollutants are tightly regulated, and local and national state-run agencies react suitably. According to these poisons, AQI is the one that is thought to be significant when evaluating the impacts on human health and the environment.

The respiratory system is the primary organ affected by SO2, making it the most hazardous air pollutant for people's health (Boznar et al., 1993 & 2002). Given that it can be transformed into sulfuric acid and sulfate, it may be harmful to human health. The production of SO2 is influenced by both artificial and natural sources. Volcanoes form the base of the vital natural resource. One of the primary activities that people do is burn fuel, especially diesel and coal. Power plants, companies that process and handle metal, and automobiles that operate on these energies generate the majority of SO2.

Both domestically and internationally, these indicators are routinely assessed and closely examined. The contaminants in the air are monitored by the UPPCB's air monitoring facilities in Ghaziabad. These stations monitor chemicals and gases that can affect the quality of the air. The center's webpage is regularly updated with the results

This review's objective is to evaluate the air quality in Ghaziabad, Uttar Pradesh's industrial center, and use a multilayer perceptron (MLP) technique to forecast and depict the condition. In this regard, the UPPCB's 2019–2023 data collection was examined. The projections were then acquired, and the information from 2019 to 2023 was presented using SPSS software. The MLP method was used in this procedure.

2. METHODS

The architecture and functions of actual brain networks serve as the basis for the quantitative and computational model known as an artificial neural network (ANN). Their preparation method, network architecture, relationship example, neural activation technique, and data handling skills are what set them apart. The MLP is the NN model that is most commonly utilized. Because this type of neural network wants an ideal output to train, it is known as a controlled network. Building a model that precisely connects input and output using verifiable data is the aim of this type of network design.

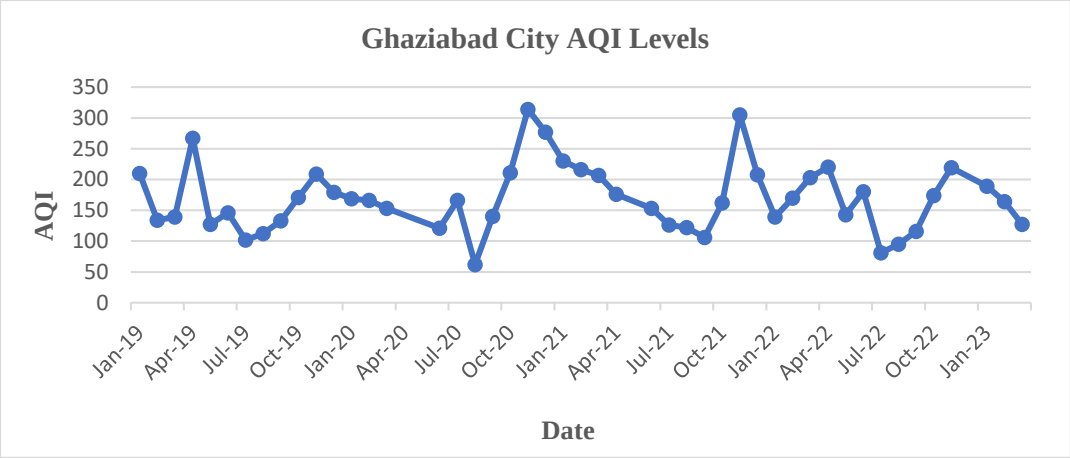
International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 03, March-2025

Therefore, even if the desired result is not immediately obvious, the model can nonetheless provide it. The connection weights increase the amount of data sources when data is moved from the input layer to the hidden layer. They are put together and then processed using a nonlinear function in the buried layer. If there are several hidden layers, the subsequent one adds, grows larger, and regulates the data treated by the linked weights after exiting the preceding hidden layer.
The output layer is then used to process the data once more and provide the neural network's output after copying it with accessible weights. The neural network has to be trained on a number of input-output mapping tests before it can be used for any particular activity. These are the essential features for every trained neural network to generate trustworthy outcomes. For the training data sample to include all the relevant information, It must contain plenty of data from various process variables and experimental setups and be somewhat huge.

3. DATA ANALYSIS

In order to gather information on the state's many urban areas, the UPPCB filters information for its homepage based on three categories: SO2, PM10, and NO2. Khora Street in Ghaziabad City is where the AQI readings are monitored. The stage has been essential in the years prior. An increase in the AQI would be accompanied by a decline in Ghaziabad's air quality. The Khora Colony in Ghaziabad's AQI data was collected for this study between January 2019 and June 2023. Figure 1 presents a summary of the data:

Figure 1. Actual data of AQI



4. ANALYSIS OF ANN

We have developed an MLP system for Ghaziabad City by applying SPSS programming. A total of 47 information points are used. One output data point and three input data points are present. PM10, SO2, and NO2 are the inputs, and the AQI is the result. Network details, such as output, hidden, and input layers, are given in Table 1. The buried layer's activation function is utilized as a hyperbolic tangent, as
seen in Table 1.
The model brief in Table 2 shows that the RE (relative error) is 0.038 during testing and 0.013 during training. Table 3 displays the estimated parameter. Figure 2 shows the specifics of the network design.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 03, March-2025

Table 1.

Network Information			
Input Layer	Covariates	1	PM10
		2	SO2
		3	NO2
	Number of Units[a]		3
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	No. of Hidden Layers		1
	Units taken in Hidden Layer		2
	Function of Activation		Hyperbolic tang
Output Layer	Dependent Component	1	AQI
	Total Units		1
	The method used in Rescaling		Regularized
	Function of Activation		Identity
	Function used for Errors		Sum of Squares

a. Excluding the bias unit

Figure 2. Network Structure

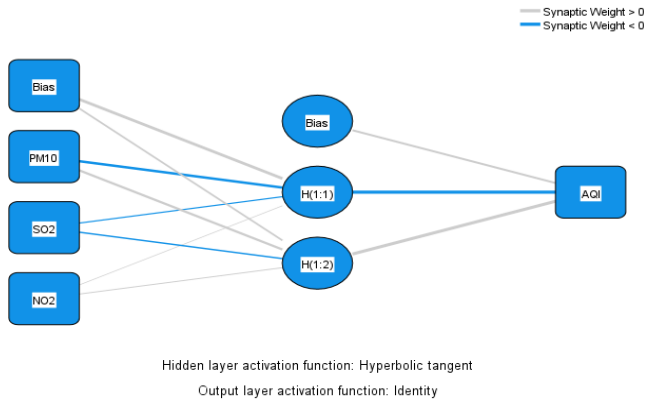


Table 2.

Model Brief		
Training	SSE (Sum of Squares Error)	0.192
	RE (Relative Error)	0.013
	Stopping Rule Used	One step after another with no reduction in errors[a]
	Training Time	0:00:00.00
Testing	SSE	0.279
	RE	0.038

Predicted Variable: Air quality Index

a. Error calculations are performed using the evaluation sample.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 03, March-2025

Table 3.

Parameter Analysis				
		Predicted		
		Hidden Layer 1		Output Layer
Predictor		H(1:1)	H(1:2)	AQI
Input Layer	(Bias)	-.151	-.245	
	PM10	.319	.484	
	SO2	.189	-.262	
	NO2	-.031	.006	
Hidden Layer 1	(Bias)			.517
	H(1:1)			1.732
	H(1:2)			1.239

According to Table 4's model brief, the RE in testing is 0.043 and in training is 0.007 if the sigmoid function is applied as the hidden layer's activation function. Table 5 displays the estimated parameter. Figure 3 shows the specifics of the network design.

Figure 3.

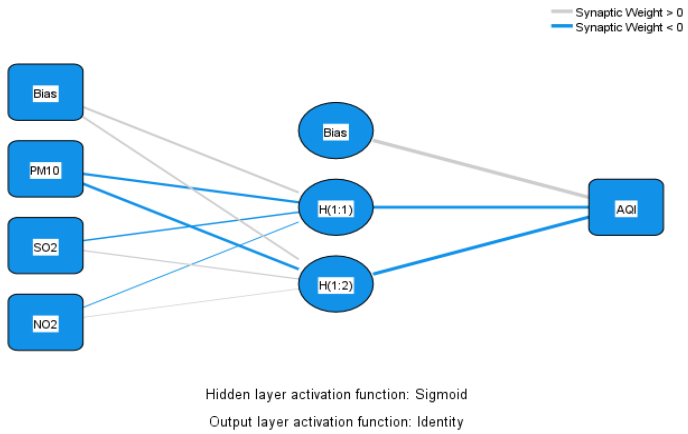


Table 4.

Model Brief		
Training	SSE	.094
	RE	.007
	Stopping Rule Used	One step after another with no reduction in errors[a]
	Training Time	0:00:00.00
Testing	SSE	.896
	RE	.043
Predicted Variable: AQI		
a. Error calculations are performed using the evaluation sample.		

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 03, March-2025

Table 5.

Parameter Analysis				
		Predicted		
		Hidden Layer 1		Output Layer
Predictor		H(1:1)	H(1:2)	AQI
Input Layer	(Bias)	.008	.884	
	PM10	-.785	-.820	
	SO2	.070	-.046	
	NO2	.113	-.071	
Hidden Layer 1	(Bias)			3.733
	H(1:1)			-1.084
	H(1:2)			-4.676

5. CONCLUSION

In this study, the multilayer perceptron approach is applied to calculate the AQI values. We face non-linearity in the analysis of the model. In order to deal with it, we create a model which is based on ANN. We now determine the R square values in order to assess the method's correctness. If the hyperbolic tangent is employed as an activation function in the hidden layer, the R square value is 0.979391; if the sigmoid is employed as an activation function in the hidden layer, the R square value is 0.971944. Since we know that R square falls between 0 and 1, we can conclude that the hyperbolic tangent provides a better approximation than the sigmoid function when used as an activation function in the hidden layer because its value is larger and closer to 1.

REFERENCES

1. Asadollahfardi G., Zangooei H. and Aria S. H. (2016). Predicting PM2.5 Concentrations using Artificial Neural Networks and Markov Chain, a Case Study Keraj City. Asian Journal of Atmospheric Environment, 10(2), 67-79.
2. Bhavsar R. (2019). Air Pollution Monitoring Using Artificial Neural Network. International Journal of Scientific & Engineering Research, 10 (12), 515-519.
3. Boznar M., Lesjak M., and Mlakar P. (1993). A neural network-based method for short-term predictions of ambient So2 concentrations in highly polluted industrial areas of complex terrain. Atmospheric Environment, 27B, 221-230.
4. Boznar M.Z. and Mlakar P. (2002). Use of neural networks in the field of air pollution modelling. Air Pollution Modeling and Its Application XV, 375-383.
5. Cogliani E. (2001). Air pollution forecast in cities by an air pollution index highly correlated with meteorological variables. Atmospheric Environment, 35, 2871- 2877.
6. Comrie A.C. (1997). Comparing Neural Networks and Regression Models for Ozone Forecasting. Air & Waste Management Association, 47, 653- 663.
7. Freeman A. M. III (1974). Air pollution and property values, a further comment. Review of Economics and Statistics, 56, 554– 556.
8. Kumar G., Sharma R.K. (2017). Air Pollution Evaluation Methods. International Journal of Engineering Research and Development, 13 (9), 12-17.
9. Kumar G. (2018). Time series analysis of PM10 for Bulandhshahr Industrial Area in NCR using Multiple Linear Regression. International Journal of Engineering Research and Development, 14 (3), 56-62.
10. Kumar G. (2018). Time series analysis of PM10 for Noida Sector 1 Industrial Area in NCR using Multiple Linear Regression. Bulletin of Pure and Applied Sciences, Section E-Math. & Stat., 37 (2), 273-277.
11. Reshma J. (2020). Analysis and Prediction of Air Quality. International Research Journal of Engineering and Technology, 7 (1), 266 - 270.