E-ISSN:2250-0758
P-ISSN:2394-6962
Research Article
Air Quality Index

International Journal of Engineering and Management Research

2025 Volume 15 Number 1 February

Publisher
Vandana Publications
www.vandanapublications.com

A Prediction of The Air Quality Index: An Analysis of Ghaziabad City

Kumar L[1*], Kumar G[2]

DOI:10.5281/zenodo.14970329

[1*] Lokesh Kumar, Research Scholar, Department of Mathematics, NAS College, Meerut, Uttar Pradesh, India.

[2] Gaurav Kumar, Professor, Department of Mathematics, NAS College, Meerut, Uttar Pradesh, India.

PM10 is one of the main air pollutants that causes air pollution. This study used Artificial Neural Networks (ANN), a common learning technique, to estimate the impact of this contaminant on human health and the environment using data between 2019 and 2023. The Pollution Control Board of Uttar Pradesh (UPPCB)'S air observation center obtained information related to the center of industry of Ghaziabad and finished the simulation and optimization procedures required using SPSS programming. Before being compared with the real data, the obtained air quality estimation results underwent a multilayer perceptron analysis. Moreover, there have been instances where the Ghaziabad province's Air Quality Index (AQI) values have exceeded the allowable limit, especially during times of great output.

Keywords: ANN, Air Pollution, AQI, Multilayer Perceptron

Corresponding Author

Lokesh Kumar, Research Scholar, Department of Mathematics, NAS College, Meerut, Uttar Pradesh, India.
Email: lokesh181889@gmail.com

How to Cite this Article

Kumar L, Kumar G, A Prediction of The Air Quality Index: An Analysis of Ghaziabad City. int. j. eng. mgmt. res.. 2025;15(1):84-88.
Available From
https://ijemr.vandanapublications.com/index.php/j/article/view/1697

To Browse

Manuscript Received	Review Round 1	Review Round 2	Review Round 3	Accepted
2024-12-15	2025-01-10			2025-02-08

Conflict of Interest	Funding	Ethical Approval	Plagiarism X-checker	Note
None	Nil	Yes	5.84	

1. Introduction

As industrialization progresses, air pollutants for example, Particulate particles (PM10), carbon dioxide (CO2), as well as sulfur dioxide (SO2) are increasing (Kumar, 2018). This increase has a severe negative impact on both the environment and human well-being, which negatively impacts public economies. By doing this, these air contaminants are strictly controlled, and state-run administrations, both local and general, respond appropriately. Based on these toxins, AQI is the one that is considered to be important when assessing the effects on ecology and human health (Boznar et al., 1993).

Oneofthe atmospheric pollutants most hazardous to people's wellness is SO2 since It mostly impacts respiration. It may represent an indirect risk to human health due to its conversion into sulfate and sulfuric acid (Reshma, 2020). Both natural and artificial sources create SO2. The cornerstone of the vital natural resource is made up of volcanoes. One of the primary causes that are created by humans is the burning of fuel, especially diesel and coal. Mostly, SO2 is produced by power plants, companies that process and handle metals, and cars that use these fuels.

These indicators are routinely assessed and thoroughly examined both domestically and internationally. These pollutants in Ghaziabad are monitored by the UPPCB's air observation facilities. These stations monitor substances and gases that affect the quality of the air. The outcomes are often updated on the webpage of the center.

This review's objectives are to evaluate the AQI situation in Uttar Pradesh's Ghaziabad City, the state's manufacturing center, and to estimate and depict using a multilayer perceptron (MLP) approach. Data obtained from the UPPCB on the AQI during the years 2019 to 2023 was examined in this respect. After that, projections were acquired, and SPSS software was used to display the data between 2019 to 2023. In this process, the MLP approach was used.

2. Methods

An ANNis a computational as well as quantitative model that is inspired by the composition and functions of actual brain networks. Their neuronal activation method, preparation strategy, network design, relationship example, and data handling capabilities are what primarily define them. The MLP is the NNmodel that is most commonly employed. Because it needs an optimal output to train, this type of neural network is known as a regulated network. Building a model that precisely links both output and input by utilizing variable data is the aim of this type of network architecture. Thus even if the desired result is not evident, the model can nonetheless deliver it.

The number of data sources increases when data is transferred from the input layer to the hidden one due to the connection weights. Following their summarization, a nonlinear function in the concealed layer handles them. If there is more than one hidden layer, the data managed by the connection weights is added, enlarged, and controlled by the next hidden layer, and so on after it leaves the first hidden layer.

Toprovide the neural network's output, the data is finally replicated with accessible weights and handled once again using the layer of output. The neural network must be trained on several input-output mapping tests before it can be used for any particular job. These are the essential details that every trained neural network has to have to provide trustworthy outcomes. Because of this, toinclude all the relevant information, the sample used for training information has to be quite vast and consist of a lot of data from numerous process variables and experimental settings.

Time Series Interpretation

Multi-layer perceptron (MLP) is applied to finish the analysis of time series, and it is defined by:

$$T_{q+1} = f(T_q, T_{q-1}, \ldots, T_1)$$

Where T_1, T_2, T_3, …, T_q are inputs and T_{q+1} is output. MLP is applied to form f.

The output is defined as the fifth information point, following the use of the first four as input.

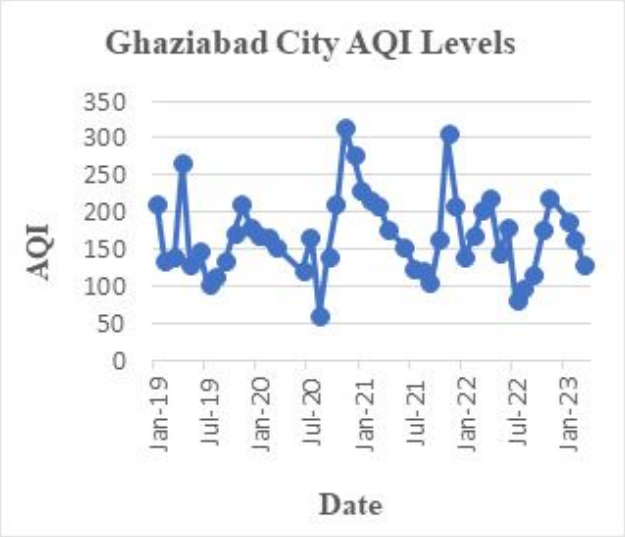
3. Data Analysis

The SO2, PM10, and NO2 air pollution components are the three categories on which the UPPCB screens information for its webpage to compile information about the state's numerous metropolitan districts.

The AQI levels are measured at Ghaziabad City's Khora Street. The stage has proven crucial during the preceding years. The quality of the air in Ghaziabad would deteriorate in tandem with an increase in the AQI.

In this work, AQI data for Ghaziabad's Khora Colony has been gathered betweenJanuary 2019 and June 2023. An overview of the data is shown in Figure 1:

Figure 1: Actual data of AQI



4. Examination of ANN

Using SPSS programming, we have developed an MLPsystem for the town of Ghaziabad. This inquiry uses forty-seven points of information in total. These noteworthy data points are categorized into forty-three groups. For every group, there are five points of information. The result was considered to be the fifthAQI information point, with the starting four points serving just as input. Again, when the first information point is eliminated, the following four points of information are inputs, and the subsequentinformation point within the order will be output, and so on. The four underlying data inputs are AI_A, AI_B, AI_C, and AI_D, whereas the label assigned for the output data point is AI_Output. Below is a time series structure that has been developed. Network information, comprising input,hidden, and output layers, is displayed in Table 1.

Relative error during testing is 2.051 and during training is 0.889, based on the model summary displayed in Table 2. Table 3 displays parameter estimates, while Table 4 discusses the importance of the independent variables.

The specifics of the network's configuration are illustrated in Figure 2. In Figure 3, predicted values are displayed versus output. The normalized importance of the network's independent variables is displayed in Figure 4.

Table 1

Network Information			
Input Layer	Covariates	1	AI_A
		2	AI_B
		3	AI_C
		4	AI_D
	Total Unitsa		4
	Rescaling Method		Standardized
Hidden Layer(s)	Hidden Layers		1
	Hidden Layer Units 1a		3
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	AI_Output
	Total Units		1
	Rescaling Method		Standardized
	Activation Function		Identity
	Error Function		Sum of Squares
a. Not including the biased unit			

Figure 2: Network Structure

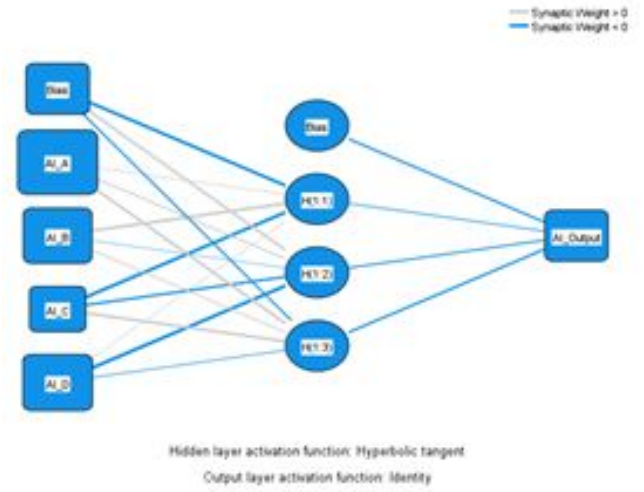


Table 2

Model Brief		
Training	A Sum of Square Error (SEE)	14.661
	Relative Rrror (RE)	0.889
	Stopping Rule	1 Consecutive steps with no decrease in error
	Training Time	0:00:00:00
Testing	SSE	0.806
	RE	2.051
Predicted Variable: AI_Output		
a. Using the assessment sample, error calculations are made		

Table 3

Parameter Estimation				
Predictor		Predicted		
		Hidden Layer 1		Output Layer
		H(1:1)	H(1:2)	AI_Output
Input Layer	(Bias)	-1.046	-.538	
	AI_A	1.819	.804	
	AI_B	.666	-.712	
	AI_C	-1.735	-.327	
	AI_D	-.129	1.447	
Hidden Layer 1	(Bias)			.161
	H(1:1)			-.772
	H(1:2)			.764

Figure 3: Predicted values against air quality index output

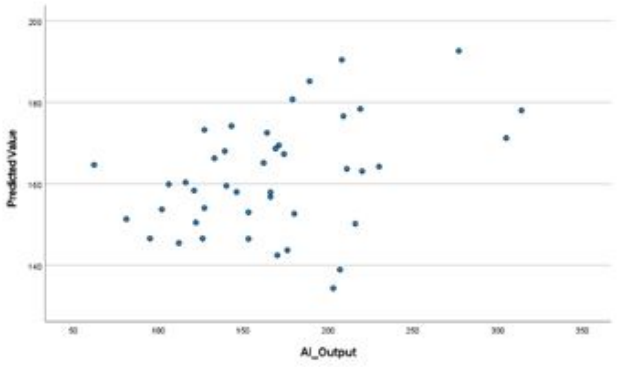
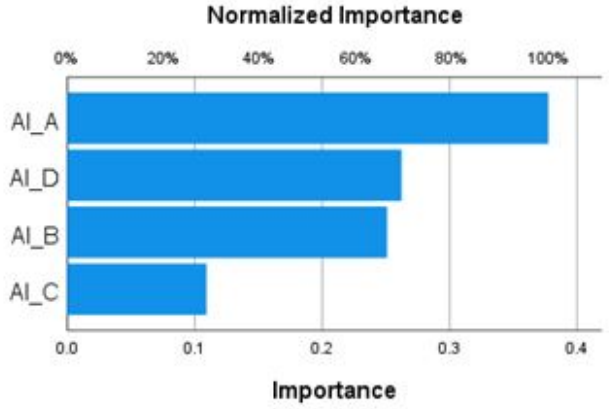


Table 4

Independent Variable Importance		
	Importance	Normalized Importance
AI_A	.377	100.0%
AI_B	.251	66.5%
AI_C	.109	28.9%
AI_D	.262	69.5%

Figure 4: Normalized Importance of Data Inputs



5. Conclusion

Time series analysis is employed to get the AQI levels in this work. The model exhibits non-linearity. A model using ANN is constructed to control non-linearity. The MLP is used in the construction of the model applying time series. The model indicates that the relative error during training is 0.889 and that of during testing is 2.051. This suggests that using past data could be able to predict future AQI levels.

References

[1] Asadollahfardi G., Zangooei H., & Aria S. H. (2016). Predicting5 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., & 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., & 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.