# A Comprehensive Review Paper on Air Quality Prediction and Remediation in Using Soft Computing.

[1]Gayatry Sharma, [2]Pooja Tiwari, [3]Sonam Bhandari, [4]Ashutosh Pandey
*gayatrysharma31@gmail.com, poojatiwari9557@gmail.com, sonambhandari@mietkumaon.ac.in, ashupandey030798@gmail.com
(1,2,3,4)Assistant Professor, MIET Kumaon Lamachaur, Haldwani, 263139, India
Department of Computer Applications

**Abstract:**
Air quality prediction and remediation are critical components in addressing environmental pollution and its associated health risks. Predictive modelling and remediation techniques are now crucial for tracking air quality levels and carrying out corrective measures due to the growing global worry about air pollution and how it harms both the environment and people's health. Artificial neural networks (ANN), fuzzy logic, support vector machines (SVM) and genetic algorithms (GA) are applications of soft computing techniques that have demonstrated significant promise in predicting air quality and developing efficient remediation strategies.
This paper explores the approach of soft computing techniques in air quality prediction and remediation and panoramic overview of various method used to predict air pollutants like particulate matter (PM2.5, PM10), nitrogen dioxide (NO2), carbon monoxide (CO), sulphur dioxide (SO2), and ozone (O3). In order to forecast future impurity level and enable alerts and better air quality management, these models make use of past air quality data, meteorological conditions, and other pertinent variables. In order to improve prediction accuracy, the article also addresses the integration of soft computing techniques with real-time data collecting systems and environmental sensors.
Additionally, the remediation component focuses on creating efficient air pollution control strategic by using optimization methods like genetic algorithms and particle swarm optimization. These tactics include deploying pollution-reducing technologies, designing green areas, and placing air purifiers in the best possible locations. Prediction models and remediation techniques work together to provide a comprehensive approach to air quality management, which helps create sustainable urban environments and smart cities.
The present research emphasizes the possibilities of soft computing methods in creating adaptive and intelligent systems for air quality forecasting and remediation.

**Keyword:** *Air quality, AQI, Pollution, Soft computing, Genetic Algorithm, Pollutant*

## Introduction

Air quality plays a vital role to preserve natural balance and safeguard public health.Numerous detrimental health outcomes, including as cardiovascular and respiratory disorders, early death, and shortened life expectancy, are directly associated with poor air quality. According to the WHO, air pollution is the biggest hazard to environmental health worldwide. The World Health Organization (WHO) estimates that it cause about 7 million premature deaths yearly as a result of the combined impacts of household and ambient air pollution. In recent years, protecting the environment and people from negative impacts of air pollution has gained attention. People should to be aware of the air they breathe. Furthermore, deteriorating air quality affects biodiversity, reduces agricultural productivity, and accelerates climate change. The concentration of dangerous particles like $PM_2$, $PM_{10}$, $NO_2$, $SO_2$, CO, and $O_3$ have risen due to urbanization, industry, and vehicle emissions make air quality management a major worldwide concern, making air quality management a critical global challenge.

In traditional method monitoring of air quality is either manual reporting system or ground based sensor technology and the result is find out by this is pollution is gradually increasing on meteorological and anthropogenic factors, which is not easy to predict manually. So the need of an automated intelligence system to predict accurate result with fast response time. This approach help in identifying pollution hotspots, forecasting pollution episodes, and guiding mitigation strategies.

In this, Role of soft computing defines by confine various techniques like ANN, CNN, Fuzzy Logic, evolutionary algorithms and hybrid system which solve complex problem. In context of air quality management, soft computing give solution for problem regarding air pollution level prediction, source attribute and anomaly detection. By adding adaptability, fault tolerance and learning capability, soft computing enables the development of robust models that can operate under noisy, incomplete, or imprecise data condition.

This review paper aims to provide an overview of common air pollutant and their impacts. Comprehensive analysis of existing soft computing techniques applied to air quality prediction and remediation.

## 2. Air Quality and Pollution Overview

### 2.1. Common Air Pollutants and Their Impact

Air pollution consists of a mixture of natural and anthropogenic substances suspended in the atmosphere, which adversely affect people health, ecosystems, and climate.

The table shows about air pollutant and effects:

| Pollutant | Emitted from | Impacts |
|---|---|---|
| **Particulate Matter ($PM_{2.5}$ and $PM_{10}$)** | tiny particles in the air that are not more than 2.5 micrometre in diameter ($PM_{2.5}$) or 10 micrometres ($PM_{10}$) | Cause respiratory and cardiovascular dieses by penetrating deeply into the circulation and respiratory system. |
| **Nitrogen Dioxide ($NO_2$)** | combustion engines | Respiratory issues as well as the production of secondary particles and ground level ozone. |
| **Sulphur Dioxide ($SO_2$)** | coal and oil combustion | Cause irritation to the throat, nose and eyes. They also react with water vapour to create acid rain. |
| **Carbon Monoxide (CO)** | Incomplete burning of fuels based on carbon | Heart attack. |
| **Ozone ($O_3$)** | Chemical processes involving nitrogen oxides, volatile organic compounds (VOCs) and sunlight | impair lung function and aggravate asthma |
| **Lead (Pb)** | industrial activity | cause neurological and developmental damage |

These pollutants impact human health, degrade building materials, reduce agricultural yields, and contribute to climate change and acidification of water bodies.

### 2.2. Air Quality Index (AQI) and Standards

A standardized measure for expressing the degree of air pollution and its possible health effects in Air Quality Index (AQI). The concentrations of major pollutants (usually like $PM_2$, $PM_{10}$, $NO_2$, $SO_2$, CO, and $O_3$) are used to generate AQI values, which are then classed as follows:

| Air Quality Index (AQI) | range | effect |
|---|---|---|
| **Good** | **0–50** | Suitable and presents minimal |
| **Moderate** | **51–100** | Quality of air is acceptable, while few contaminants can cause mild anxiety in sensitive individuals. |
| **Unhealthy for Sensitive Groups** | **101–150** | Sensitive groups can feel their health effects |
| **Unhealthy** | **151–200** | Anybody could start suffering from negative health consequences. |
| **Very Unhealthy** | **201–300** | Health warning: more severe health problems could affect everyone. |
| **Hazardous** | **301–500** | The whole population is more likely to br impacted by emergency situations. |

Every nation and area has its own AQI criteria and standards. To safeguard public health, the World Health organization (WHO), the Central Pollution Control Board (CPCB) Of India , and the U.S. environment protection Agency (EPA) all establish recommendations for acceptable air pollution levels.

**3. Soft Computing Techniques for Air Quality Modelling**

Soft computing is a branch of computational intelligence that deals with unreliability and partial solutions to model and solve complex physical world problems. Unlike traditional computing does not require a precisely defined mathematical model, making it especially suitable for environmental applications like air quality modelling, which involves nonlinear and dynamic systems influenced by multiple interdependent variables. This section presents a detailed overview of the most prominent soft computing techniques applied in air quality prediction and remediation.

---

**3.1. Fuzzy Logic (FL)**

A type of logic known as fuzzy logic deals with approximate reasoning as opposed to fixed and accurate arguments since the truth values of variables can be any real number 0 and 1. It works well for handling ambiguous and imprecise data, which is common in environmental data sets. Fuzzy logic is used in air quality technique to:

- Categorize air contamination levels (e.g., AQI classification).
- Model expert knowledge using fuzzy inference systems.
- Deal with frequent fuzzy associations between meteorological parameters and pollutant values.

One of the suitable methods to simulate the learning process of human brain (cognitive process) is by using Fuzzy logic system which is most appropriate for use as a tool for building the decision support system in environmental management where human's reasoning and decision making is done in uncertain environment.

**3.2. Artificial Neural Networks (ANN)**

The composition and operation of the human brain serve as the model for artificial neural networks. They are made up of linked layers of nodes, or neurons, that use training to discover intricate patterns and connections in data.

In air quality prediction, ANNs are widely used because of their capability to model nonlinear and temporal connection among multiple input variables, such as:

- Historical pollutant concentrations.

- Meteorological factors (temperature, humidity, wind speed).
- Geographic and temporal factors.

ANNs, including feedforward, recurrent (RNN), and convolutional (CNN) architectures, have shown high accuracy in forecasting pollutant levels. However, they require sufficient data and can act as "black boxes" with limited interpretability.

---

### 3.3. Evolutionary Algorithms (GA, PSO)

Evolutionary Algorithms (EAs) are optimization techniques galvanize by natural processes such as evolution and swarm intelligence. The two most commonly used EAs in air quality applications are:

- **Genetic Algorithms (GA):** Mimic natural selection processes to optimize model parameters or feature selection in predictive models.
- **Particle Swarm Optimization (PSO):** Mimics how fish or birds interact with one another to discover the best answers by cooperatively searching a search field.

EAs are often used in combination with other soft computing models to enhance their performance, such as optimizing weights in neural networks or tuning membership functions in fuzzy systems.

---

### 3.4. Support Vector Machines (SVM)

Support Vector Machines use the best hyperplane to divide data into distinct classes to carry out classification and regression task.

In air quality prediction, SVMs are employed for:

- Classifying AQI categories based on pollutant levels.
- Performing regression to predict specific pollutant concentrations.

SVMs offer high generalization capability and perform well on small to medium-sized datasets. Nonlinear interactions between variables can be handled by kernel functions like the Radial Basis Function (RBF).

---

### 3.5. Hybrid and Ensemble Methods

Hybrid and ensemble models combine two or more soft computing techniques to leverage the strengths and compensate for the weaknesses of individual models. Common combinations include:

- **Fuzzy-ANN:** Integrates fuzzy logic with neural networks to improve interpretability and learning capability.
- **GA-ANN or PSO-ANN:** Use evolutionary algorithms to optimize the parameters of neural networks.
- **SVM-ANN:** Combines the regression strength of SVM with the nonlinear learning of ANN.

Ensemble techniques, such as bagging, boosting, and stacking, aggregate the outputs of multiple models to enhance prediction accuracy, robustness, and reliability.

These advanced methods have been dealing with complex and high-dimensional air quality data, improving both forecasting and decision-making capabilities in environmental systems.

**Review of Literature-** In this major issue a lot of research in our country as well as in foreign country has been done. Some of the research work discuss hear:

**S.TikheShruti (2013):** For Pune, Maharashtra, which ranks second on the list of polluted cities in India, the study used soft computing algorithms i.e. Artificial Neural Network (ANN) and Genetic Programming (GP), to predict concentration levels of air pollutants, including oxides of Sulphur($SO_2$), oxides of Nitrogen($NO_2$), and Repairable Suspended Particulate Matter (RSPM), for the years 2005 to 2011. Based on hourly average data values of pollutants concentrations of more than 7 years, they have created a total of six models (3 of each algorithm, ANN and GP).GP algorithms perform better than ANNs out of these two methods.

**Nidhi Sharma (2018)**:It include a critical observation of the air pollution trend in Delhi, India, from 2016 to 2017 as well as a deep analysis of air contaminants from 2009 to 2017.They have forecasted the future patterns of number of pollutants, including carbon monoxide (CO), benzene, suspended particular matter (PM), Sulphur dioxide ($SO_2$), nitrogen dioxide ($NO_2$), and ozone ($O_3$). They have projected the future values of the pollutants previously mentioned based on historical records by employing data analytics and time series regression forecasting. The AnandVihar and Shadipur monitoring stations in Delhi are being examined in accordance with the finding of this study. The results indicate a sharp rise of PM10 concentrations, as well as a noticeable increase in $NO_2$ and PM2.5 which indicating a rise in pollution in Delhi. CO is expected to decrease by 0.169 mg/$m^3$, $NO_2$ concentrations will rise by 16.77 µg/$m^3$, ozone will rise by 6.11 mg/$m^3$, benzoene will decrese by 1.33mg/$m^3$ ,and $SO_2$ will rise by 1.24 µg/$m^3$ in upcoming years.

**Aditya C R (2018)**: Utilized machine algorithms to identify and predict the amount of PM2.5 concentration based on a dataset that included the atmospheric conditions in a particular city. On a specific date, they also forecasted the PM2.5 concentration level. They first use the logistic regression method to determine if the air is contaminated or not, and then they utilized the Auto Regression algorithm to forecast the future PM2.5 value based on historical data.

**Mohamed Shakir and N.Rakesh (2018)** have used the WEKA tool to find the impact of environmental factors like temperature, wind speed, and humidity on the aforementioned air pollutants. They take different air pollutants (NO, $NO_2$, CO, PM10, CO, PM10 AND $SO_2$) to find the relation to the time of day and the day of week. In this, Karnataka pollution control board provided the data. They study's findings, which were obtained by applying the ZeroR algorithum in the WEKA tool, indicate that air pollution concentration levels rise during the workday, particularly during peak hours, and fall during weekends or vacations. The study illustrates the connections or dependencies between air pollutants like NO, $NO_2$, PM10, CO AND $SO_2$ and environmental parameters like temperature, wind speed, and humidity using simple K-means clustering methods.

**Yafouz, Ahmed, Zaini, & El-Shafie, (2021)**: Despite covering other air pollutants, it used mostly on Ozone (O3) and the review paper identified four areas of interest in machines (SVM), decision trees (DT), and hybrid models, even though it covered additional air contaminants. Though limited to a specific sub-domain in air pollution prediction using machine learning, this research yields intriguing results.

**A.Samad , S.Garuda, U. Vogt, B. Yan(2023)** This paper give 3 outcome after testing the model in 4 different scenario. It demonstrated that models such as Ridge Regression, SVR, Random Forest and XGBoost are capable of accurately predicting key pollutants like $PM_{2.5}$, $PM_{10}$, and $NO_2$. The most accurate predictions were achieved when the models used data from nearby physical monitoring stations, showing the spatial relevance of air quality measurements. The approach was also successfully tested in a different city (Karlsruhe), suggesting that the virtual monitoring station framework is transferable and applicable in other regions.

**G. Ravindiran, G. Hayder, K. Kanagarathinam, A. Alagumalai, and C. Sonne (2023)** The forecast of AQI of Visakhapatnam for the observational period from2017 and 2022. The AQI concentrations became higher in the periods of 2017 and2019, and then a dip in 2020 because of the countrywide shutdown due to Covid-19. But the AQI readings persisted grown up. PM2. 5 and PM10 were identified as the major pollutants that participated in calculating the AQI, and the metrological properties were minimal impact. Rate between 68% and 75% of their predictions, which is lower than human respondents. Result with the predicted AQI, Random Forest and Catboost model accurately predicted the AQI and the various decisions based on the AQI of the citizens of Delhi, which exhibited a maximum curve correlation at 0.9998 and 0.9936, respectively. Study told that **meteorological parameters** have very less effect on the AQI predictions—most of the predictive power came from particulate and gaseous pollutant concentrations.

**Conclusion:**

Air pollution is the biggest threats to the environment and public health in the globe today. Accurate prediction and effective control strategies are essential for minimizing its impact. This review highlights the vital role of **soft computing techniques** like fuzzy logic, artificial neural networks, evolutionary algorithms, and support vector machines—in enhancing air quality modelling, forecasting, and decision-making.

Soft computing methods offer significant advantages over traditional models by handling complex, nonlinear, and uncertain environmental data. Numerous case studies have shown their successful application in predicting air quality across different geographic regions and supporting intelligent control systems. In addition, the integration of soft computing with emerging technologies such as IoT, cloud computing, and smart city infrastructure provides a promising path toward real-time, scalable air quality management.

However, challenges such as data quality, model interpretability, and computational complexity still need to be solved. Advance research need to prioritise to improve model transparency, utilizing cross-domain data fusion, and building adaptive, real-time systems to meet the demands of dynamic urban environments.

In conclusion, soft computing holds great potential in advancing air quality prediction and remediation. With continued innovation and interdisciplinary collaboration, it can play a key role in developing sustainable and intelligent solutions for cleaner air and healthier communities.

**References**

1. Y. Zhang, M. Bocquet, V. Mallet, C. Seigneur, and A. Baklanov, “Real-time air quality forecasting, part I: History, techniques, and current status,” *Atmospheric Environment*, vol. 60, pp. 632–655, 2012.
2. T. Cheng and J. Wang, “A hybrid deep learning model for air quality prediction: Integrating LSTM and GA-optimized neural networks,” *Environmental Science and Pollution Research*, vol. 27, no. 25, pp. 31616–31631, 2020.
3. A. Singh and A. Sharma, “Air quality prediction using soft computing techniques: A review,” *Int. J. Environ. Sci. Technol.*, vol. 14, no. 11, pp. 2495–2510, 2017.
4. P. Kumar and P. Goyal, “Application of soft computing methods for air quality prediction and management,” *Sustainable Cities and Society*, vol. 51, p. 101703, 2019.
5. S. Jain and R. Kumar, “Fuzzy logic applications in air pollution control: A review,” *J. Cleaner Prod.*, vol. 172, pp. 2575–2590, 2018.
6. C. Huang and Y. Sun, “IoT-based air pollution monitoring system with soft computing techniques,” *Sensors*, vol. 21, no. 3, p. 1020, 2021.
7. S. Gupta and R. Prakash, “Optimization algorithms in emission control: A review,” *Environ. Technol. Innov.*, vol. 20, p. 101113, 2020.
8. J. R. Arnold et al., "Testing CMAQ chemistry sensitivities in base case and emissions control runs at SEARCH and SOS99 surface sites in the southeastern US," *Atmos. Environ.*, 2006.
9. F. Li et al., "Air pollution prediction based on ensemble learning method with feature selection," *Environmental Science and Pollution Research*, 2021.
10. R. Das and S. Ghosh, "Hybrid machine learning model for air quality forecasting," *Environmental Monitoring and Assessment*, 2020.

11. M. R. Islam et al., "A review of evolutionary algorithms for air pollution optimization," *Journal of Environmental Management*, 2021.
12. X. Wang and Z. Li, "Application of support vector machines in air quality prediction," *Environmental Modelling & Software*, 2019.
13. L. Zhang et al., "Deep learning approaches for PM2.5 forecasting: A review," *Atmospheric Research*, 2020.
14. S. Rao and P. Nair, "Soft computing based traffic emission control in urban cities," *International Journal of Environmental Science and Technology*, 2019.
15. H. Lee and J. Park, "Smart city air quality monitoring and control with IoT and AI," *Sustainable Cities and Society*, 2022.
16. M. T. Nguyen et al., "Fuzzy logic and ANN hybrid models for air pollution prediction," *Environmental Science and Pollution Research*, 2021.
17. V. Kumar and D. Sharma, "Optimization techniques for emission control: PSO and GA applications," *Environmental Technology*, 2020.
18. K. Patel and S. Shah, "Real-time air quality monitoring system using IoT and cloud computing," *Journal of Cleaner Production*, 2021.
19. D. Thomas and R. Bhattacharya, "Data fusion techniques for environmental monitoring," *Environmental Monitoring and Assessment*, 2020.
20. J. Fernandez et al., "Explainable AI for environmental applications: Challenges and solutions," *AI for Earth Science*, 2022.
21. https://www.who.int/teams/environment-climate-change-and-health/air-quality-and-health/health-impacts/types-of-pollutants
22. S.TikheShruti, K.C.Khare, S.N.Londhe, “Forecasting Criteria Air Pollutants Using Data Driven Approaches: An Indian Case Study”, International Journal of Soft Computing 8 (4), 305-312, 2013, ISSN: 1816-9503
23. Nidhi Sharma , ShwetaTaneja , VaishaliSagar , Arshita Bhatt, “Forecasting air pollution load in Delhi using data analysis tools”, ScienceDirect, 132 (2018) 1077– 1085.
24. A. Samad, S. Garuda, U. Vogt, and B. Yang, “Air pollution prediction using machine learning techniques – An approach to replace existing monitoring stations with virtual monitoring stations,” *Atmospheric Environment*, vol. 310, p. 119987, 2023. doi: 10.1016/j.atmosenv.2023.119987.
25. Aditya C R, Chandana R Deshmukh, Nayana D K, Praveen Gandhi Vidyavastu, “Detection and Prediction of Air Pollution using Machine Learning Models”, International Journal of Engineering Trends and Technology (IJETT) – volume 59 Issue 4 – May 2018
26. Mohamed Shakir, N. Rakesh, “Investigation on Air Pollutant Data Sets using Data Mining Tool”, IEEE Xplore Part Number:CFP18OZV-ART; ISBN:978-1- 5386-1442-
27. Yafouz, A., Ahmed, A. N., Zaini, N., Sherif, M., Sefelnasr, A., & El-Shafie, A. (2021). Hybrid deep learning model for ozone concentration prediction: Comprehensive evaluation and comparison with various machine and deep learning algorithms. Engineering Applications of Computational Fluid Mechanics, 15(1), 902–933.
28. G. Ravindiran, G. Hayder, K. Kanagarathinam, A. Alagumalai, and C. Sonne, “Air quality prediction by machine learning models: A predictive study on the Indian coastal city of Visakhapatnam,” *Chemosphere*, vol. 338, p. 139518, 2023. doi: 10.1016/j.chemosphere.2023.139518.
29. R. M. Patil, H. T. Dinde, and S. K. Powar, "A literature review on prediction of air quality index and forecasting ambient air pollutants using machine learning algorithms," *Int. J. Innov. Sci. Res. Technol.*, vol. 5, no. 8, pp. 1148–1152, Aug. 2020.
30. K. S. R. Kumar, H. S., V. Swetha, and V. S. S. Naik, "Prophecy of air quality using KNN-LSTM," *Int. J. Adv. Res. Sci. Commun. Technol. (IJARSCT)*, vol. 2, no. 9, pp. 410–417, Jun. 2022, doi: 10.48175/IJARSCT-5364.
31. S. Saxena, S. Gupta, and S. Gupta, "Vaayu for Aayu (AQI) – A study to assess the air quality," *Int. J. Sci. Res. Eng. Manag. (IJSREM)*, vol. 5, no. 6, pp. 2–7, Jun. 2021.