

Artificial Intelligence Approaches for Predicting Air Quality Index and Air Quality Grades: A Comparative Study

Dr Bal Krishna Sharma

Professor

Department of Computer Science and Application

Mandsaur University Mandsaur (MP)

drbks7426@gmail.com

Corresponding Author: * Bal Krishna Sharma (drbks7426@gmail.com)

Received: 20 July 2026 | Revised: 25 August, 2026 | Accepted: 3 September, 2026

Abstract—Air pollution has become a critical environmental and public-health concern, creating a need for accurate and timely assessment of air-quality conditions. The Air Quality Index (AQI) provides a standardized numerical representation of atmospheric pollution and converts complex pollutant information into understandable air-quality grades. This review examines Artificial Intelligence (AI) approaches for predicting AQI values and corresponding air-quality grades using diverse environmental datasets. It discusses major pollutants, including PM_{2.5}, PM₁₀, NO₂, SO₂, CO, O₃, and NH₃, along with meteorological parameters such as temperature, relative humidity, wind speed, and wind direction. The review further explores machine learning and deep learning techniques for identifying nonlinear relationships, temporal variations, pollutant patterns, and changing air-quality conditions. Data sources and preprocessing methods, including monitoring stations, low-cost sensors, mobile monitoring, satellite observations, calibration, missing-value treatment, and anomaly detection, are also considered. In addition, the study highlights integrated frameworks combining multi-sensor networks, data fusion, atmospheric information, IoT-enabled monitoring, and AI-based assessment. Challenges related to data quality, nonlinear interactions, spatial-temporal variability, computational complexity, and model reliability are discussed. Overall, AI-driven AQI prediction can effectively support real-time monitoring, health-risk communication, environmental management, pollution-control planning, and timely public-health decision-making.

Keywords—Air Quality Index, Artificial Intelligence, Machine Learning, Deep Learning, Air Quality Prediction, Air Quality Grades, Pollutant Concentration.

I. INTRODUCTION

The major environmental problems facing large cities are air pollution, driven by population growth and anthropogenic fuel consumption in transportation and industrial activities. They are considered a serious environmental threat because they harm humans and the environment. Air pollutants differ in their risk on human according to their concentrations and types, and it is difficult for the public to determine the risk of air pollution from pollutant concentrations in advance, as they differ in their determinants and units [1]. Therefore, a method is needed to determine the combined effect of multiple pollutants and to rank their health risks as a single number. In addition, policy-makers have recommended displaying air quality indices to inform the public and alert them to the risk of pollution. Air quality index (AQI) is a model used to convert the concentrations of air pollutants into a single number representing the level of air quality condition [2]. There are many air pollution indices in the literature with different names like Air Quality Index (AQI), Air Pollution Index (API), and Pollution Standard Index (PSI), there are two types of AQIs according to the principle of their calculation method [3]. The first type considers the highest sub-index as AQI after calculating the sub-indices for all pollutants depending on USEPA or local air pollutant standards [4].

Through this type of AQI, a large amount of information is lost, while public awareness improved.

An air quality grade prediction model based on the KNN algorithm [5]. Firstly, the historical measurement data of air quality is crawled from the relevant weather website. Artificial Intelligence (AI), encompassing both ML and Deep Learning (DL) techniques, has emerged as a powerful solution to improve traditional epidemiological and environmental health models. Furthermore, AI plays a crucial role in the real-time tracking of pollution hotspots, identifying trends in pollution levels, and modelling the impact of meteorological factors [6]. This technology enables researchers and regulators to make sound decisions on air quality management and public health protection. AI provides rapid, precise, and trustworthy answers, helping to make sense of complex data across diverse scientific fields and aiding in data-driven decision-making processes. ML algorithms, such as ANN, RF, SVM, and XGBoost, have proven to be invaluable tools in addressing air pollution and its impact on public health [7]. These algorithms are particularly well-suited for processing large and multidimensional air quality data sets, which can lead to a more in-depth understanding of the complex relationship between air pollutants and their associated health impacts.

A. Structure of the Paper

The paper is organized into the following sections. This section presents an introduction to the study. In Section II, AQI and prediction of air-quality grade are discussed. In Section III, AI-based air-quality assessment is explored. The Air Quality Index is found in Section IV. The paper is completed in Section V with a discussion about future work.

II. AIR QUALITY INDEX AND AIR QUALITY GRADE PREDICTION

A standardized measure of atmospheric pollution and communicates overall air-quality conditions through understandable grades. Accurate AQI forecasting supports timely public health warnings, pollution control planning and environmental management. It also helps authorities identify deteriorating air-quality conditions and implement appropriate mitigation strategies effectively.

A. Advanced Data-Driven Approaches for Air-Quality Assessment

To enhance air-quality assessment by integrating diverse environmental, sensor, and meteorological datasets to identify complex spatial and temporal pollution patterns. These approaches support improved data integration, anomaly identification, exposure estimation, and evidence-based environmental analysis, enabling more comprehensive and reliable assessment of changing air-quality conditions across different locations and time periods.

1) Data Sources and Pre-Processing

Air-quality assessment uses station observations, meteorological measurements, remote sensing products, and geographic datasets. Pre-processing includes data cleaning, calibration, spatial transformation, and geographic correction. Integrating these heterogeneous sources improves data consistency, spatial coverage, and temporal analysis, supporting reliable and comprehensive evaluation of air-quality conditions across diverse regions [8].

2) Changes in Air Quality Data Acquisition

Recent developments in methods for collecting data on air quality have helped to shape the way LUR approaches have developed. Furthermore, consolidate knowledge in this area by delving into the understanding of how monitoring network architecture affects modelling outcomes.

- **Low-cost monitors and sensors:** Recent advances in technologies have enabled expensive and complex stationary equipment for air pollution monitoring to be supplemented with more affordable, portable monitors or sensors Low-cost monitors (LCMs; note that the ongoing debate revolves around what constitutes an LCM [9].
- **Mobile monitoring:** Air quality data collected by mobile monitoring have been widely used to develop LUR models in numerous recent studies Mobile monitoring shows great promise for understanding fine-scale spatial pollutant patterns and identifying emission sources and urban air pollution hotspots efficiently.
- **Satellite observations:** The incorporation of satellite data into LUR models has grown in importance alongside in-situ monitoring. A significant benefit is that satellite data offer extensive and synoptic

coverage, capturing both spatial and temporal variations across vast geographical areas [10].

B. Air Pollution Level Communication

Air pollution level communication translates complex air-quality information into understandable categories, numerical values, and health-related messages. AQI-based communication helps the public recognize pollution severity, understand potential health risks, and make informed decisions about outdoor activities and exposure. It also supports public awareness, community engagement, health advisories, and environmentally responsible behaviour.

1) Accessible Information

The AQI breaks down intricate data on air quality into categories or numerical values that are simple to understand. Through the process of converting technical data into comprehensible representations, such as colour-coded scales or descriptive categories (such as good, moderate, and unhealthy), the public is provided with an understandable overview of the current state of air quality [11].

2) Health Risk Interaction

One of its main functions is to communicate any possible health hazards connected to certain AQI values For example, higher AQI readings indicate more health risks; therefore, people should take precautions, especially those who are sensitive to air pollution [12].

3) Information-Based Empowerment

The AQI gives people the ability to make knowledgeable decisions by giving them access to either periodic or real-time updates on the state of the air [13]. It allows individuals to modify their actions in response to pollution levels, for example, changing their outdoor workout regimens, donning protective gear, or avoiding exposure during times of high pollution.

4) Policy Support and Community Engagement

Making AQI data available to the public encourages community involvement and support for higher air quality standards. It is an essential tool for citizens, environmental organizations, and legislators to promote laws meant to enhance air quality [14][15].

5) Public Health Messaging

When health warnings are issued, AQI data is frequently included to inform susceptible groups such as children, the elderly and people with respiratory conditions about the possible health hazards connected to particular pollution levels [16]. This proactive advertising encourages preventative measures to lessen harmful consequences on health.

6) Long-Term Behavioural Change

Regular exposure to AQI data can lead to long-term behavioural changes in people, both as individuals and as a society [17]. It can promote a change to more environmentally friendly habits, including taking public transportation or endorsing laws that lower emissions from cars and industry Essentially, the AQI acts as a vital conduit for complicated air quality data to the general population, providing them with knowledge they can use to safeguard their health and promote environmental stewardship.

C. Challenges in AQI Prediction

Forecasting AQI is intrinsically problematic owing to the complex and nonlinear dynamics of environmental systems [18]. Key difficulties include:

- **Data Quality and Availability:** Incomplete records, sensor failures, and inconsistent readings might influence the credibility of predictions.
- **Nonlinear Interactions:** Air pollutants interact with meteorological factors (e.g., temperature, humidity, wind speed) in complicated, nonlinear ways.
- **Temporal and Spatial Variability:** Pollutant levels may change greatly over time (diurnal and seasonal patterns) and throughout various metropolitan zones, complicating model applicability.
- **Computing Complexity:** Advanced forecasting models, particularly those based on DL, need large computing resources and careful calibration to prevent overfitting [19]

III. ARTIFICIAL INTELLIGENCE IN AIR-QUALITY ASSESSMENT

Artificial intelligence has transformed air-quality assessment by enabling efficient analysis of complex environmental datasets. Machine learning and deep learning models identify pollutant patterns, estimate AQI, and evaluate meteorological influences. Integration with IoT sensors further supports continuous monitoring, hotspot detection, and data-driven environmental management, improving timely air-quality assessment and public-health protection.

A. AI for Air-Pollution Monitoring

Analysing data from sensors, meteorological observations, and environmental datasets. Machine learning models identify pollutant patterns, detect pollution hotspots, and estimate air-quality conditions [20]. Integration of IoT-enabled sensors with AI enables continuous monitoring, timely pollution assessment, and improved environmental decision-making for effective urban air-quality management.

1) Data Imputation for Handling Missing Values

A number of factors, including malfunctioning sensors, incorrect data transfer, and outside interference, can lead to inaccurate or missing readings in air quality monitoring. For air quality evaluations and analyses that follow to be based on reliable data, it is crucial to handle these outliers correctly [21].

2) Sensor Calibration and Drift Compensation

Data accuracy in air quality monitoring relies on accurate sensor calibration. Sensors are susceptible to drift over time as a result of ageing, environmental changes, or the accumulation of pollutants. Data integrity and accuracy are at risk when this drift is not addressed on a regular basis. ML methods such as ANN, RF, and multiple linear regression have been employed for the calibration process [22]. By making modifications in real-time and adjusting for sensor drift, these methods allow for the optimization and automation of calibration. These techniques make low-cost sensors more effective and increase data quality and dependability, opening the door to their use in environmental monitoring on a grand scale.

3) Anomaly Detection and Outlier Treatment

There are a lot of potential causes for air quality monitoring issues, such as sensor malfunctions,

environmental disturbances, or temporary changes in pollutant levels [23]. In order to keep the data accurate and reliable, it is crucial to detect and handle these outliers. Anomaly detection using ML methods helps differentiate between real environmental abnormalities and sensor mistakes, guaranteeing that analyzes like pollution forecast and cause attribution are grounded on correct data.

4) Short-Term Air Quality Forecasting

Forecasting short-term air quality is crucial for real-time pollution monitoring and control, particularly in metropolitan regions where pollutant concentrations can change quickly. If you have access to historical data, you may use time-series analysis to discover patterns, seasonality, and irregular oscillations in pollution levels [24]. Better datasets are the result of these techniques' ability to detect anomalies, imputation of missing data, and outliers.

5) Air Quality Index (AQI) Estimation

Air quality index (AQI) is a numerical metric utilized by health organizations and government bodies globally to convey air pollution levels. The public can easily grasp the present air quality conditions and the health concerns connected with them since it simplifies complicated data on several contaminants into a single, standardized number [25]. Effective air quality monitoring relies on reliable AQI estimates because they give a common metric for informing the public about possible health hazards.

B. Integrated Air Quality Assessment Frameworks

Air quality assessment lies in integrating the strengths of various techniques to overcome individual limitations. Integrated frameworks combine ground-based monitoring, satellite observations, atmospheric modelling, and ML/AI algorithms to provide a more comprehensive, accurate, and actionable understanding of air quality [26].

1) Multi-Sensor Networks

Deploying dense networks of calibrated low-cost sensors alongside strategic reference-grade stations creates a multi-tiered monitoring system, providing both high-accuracy data at specific points and broad spatial coverage with reasonable accuracy [27].

2) Data Assimilation and Fusion

Advanced data assimilation techniques merge real-time measurements from diverse sources with outputs from chemical transport models, leading to improved air quality forecasts and analyses [28]. ML algorithms are instrumental in this data fusion process, handling the heterogeneity and large volume of data.

3) Citizen Science and Community Engagement

Citizen research in air quality monitoring has been on the rise due to the proliferation of affordable sensors. Community-led initiatives, often supported by accessible sensor technologies and online data platforms, empower citizens to collect local air quality data, raise awareness, and advocate for clean air policies [29]. While data quality from citizen sensors can vary, their sheer number provides valuable localised information and promotes public engagement.

C. Meteorological Parameters and Their Impact on Air Quality

The monitoring of meteorological conditions plays a crucial role in understanding and predicting air quality

patterns [30]. These parameters interact with pollutants in complex ways, significantly influencing their dispersion, concentration, and overall impact on air quality:

- **Temperature:** TEMPERAT °C Temperature variations significantly influence air pollution dynamics through vertical air movement and pollutant dispersion, chemical reaction rates of atmospheric pollutants, formation of ground-level ozone in warmer conditions, thermal inversion effects that can trap pollutants, and the impact on local wind patterns and pollution transport.
- **Relative Humidity (RH%):** Humidity levels affect air quality through several mechanisms: particle size and composition modifications, formation of secondary aerosols, visibility reduction in high humidity conditions, impact on photochemical reaction rates, and the influence on pollutant deposition rates.
- **Wind Speed (WS m/s):** One important aspect of the dispersion of pollutants is the speed of the wind. As a general rule, better air quality is achieved at faster speeds due to more mixing. Low speeds can lead to pollutant accumulation. Noteworthy are the impact on dust suspension and transport, the influence on pollutant concentration patterns, and the effect on atmospheric stability conditions.
- **Wind Direction (WD):** Wind direction determines the pollutant transport patterns from source regions, the impact areas of industrial emissions, the urban pollution distribution patterns, the regional air quality influences, and the seasonal pollution patterns based on prevailing winds.

D. Machine Learning and Deep Learning for Air-Quality Monitoring

The ability of ML and DL to understand complicated correlations from varied meteorological and environmental data has made them crucial tools for air-quality monitoring. These techniques support the analysis of pollutant concentrations, meteorological conditions, and temporal variations while identifying nonlinear patterns that are difficult to capture through conventional methods. Their ability to process large-scale and continuously generated datasets improves air-quality assessment, forecasting, and timely identification of changing pollution conditions [31].

IV. AIR QUALITY INDEX

The Air Quality Index (AQI) has the ability to simplify the complexity of the information on air pollution into an easily understood numeric value, which makes it easier for both the general public and policymakers to comprehend the air quality [32]. The Air Quality Index (AQI) measures levels of pollutant concentrations beyond those given in WHO criteria in order to determine air quality categorization, which is indicated by a range from very excellent to extremely dangerous.

A. Pollutants Considered in AQI

Air Quality Index (AQI) evaluation takes into account several major air pollutants that directly affect environmental quality and human health. The pollutants that are monitored include particulate matter (PM_{2.5} and PM₁₀), nitrogen dioxide (NO₂), sulfur dioxide (SO₂), carbon monoxide (CO), ozone (O₃) and ammonia (NH₃). These pollutants are emitted

from sources including vehicles, industry, construction, biomass burning and natural processes [33]. They are measured and then translated to pollutant-specific sub-indices which are then used to calculate the overall AQI, and the air-quality grade. The pollutants included in the AQI are listed below:

- **Particulate Matter (PM_{2.5} and PM₁₀):** These are very small solid or liquid particles in the air. The particles are categorized into PM_{2.5} (diameter of 2.5 µm or less) and PM₁₀ (diameter of 10 µm or less).
- **Nitrogen dioxide (NO₂):** A significant gas pollutant emitted mainly from vehicle exhaust, industrial combustion and power generation, and other situations involving combustion at high temperatures.
- **Sulfur dioxide (SO₂):** The gaseous pollutant mainly released from fossil-fuel combustion and industrial activities. Can cause irritation to the respiratory system and cause acid rain and particulate pollution.
- **Carbon monoxide (CO):** A colourless and odourless gaseous pollutant, primarily resulting from incomplete combustion of fuels in transportation, industries and domestic appliances. Because CO decreases the blood's capacity to carry oxygen, high levels can lead to headaches, dizziness, shortness of breath and can have serious cardiovascular and neurological effects.
- **Ozone (O₃):** A reactive gas pollutant produced close to the ground by the reaction of nitrogen oxides and volatile organic compounds in the presence of sunlight. High levels of ozone can cause irritation of the lungs, decreased lung function, and worsening asthma and other breathing conditions.

B. AQI Categories and Health Implications

The AQI categories represent the severity of air pollution and its potential impacts on human health. Good air quality means it poses little risk, Satisfactory/Moderate means that it can impact sensitive people. Respiratory discomfort and exacerbation of existing diseases may occur with Poor and Very Poor air quality [34]. Extremely high levels of pollution pose a significant threat to the health of everyone, but notably to the young, the old, and those who already suffer from respiratory or cardiovascular conditions.

- **Good Air Quality:** Air pollution levels are minimal and provide little to no risk to human health when air quality is good. In this kind of atmosphere, people can carry out normal outdoor activities without having to consider particular health issues arising from pollution.
- **Satisfactory air quality:** Satisfactory Air Quality means that overall concentrations of pollutants are generally acceptable and do not pose significant health risks to most people. Sensitive individuals might experience some pain if they are exposed to pollutants for longer, especially if pollutant levels rise briefly.
- **Poor air quality:** High concentrations of air pollutants, which are harmful to human health, are what is known as poor air quality. Those most vulnerable, such as youngsters, the elderly, and those with compromised immune systems, may have more severe side effects from prolonged exposure, including coughing, sore throat, trouble breathing, and impaired lung function.

V. LITERATURE REVIEW

Machine learning, deep learning, IoT sensors, spatio-temporal models, and air quality index (AQI) prediction have all been the subject of recent AI research. These techniques improve prediction accuracy, early warning, and environmental decision-making, while challenges remain in real-time adaptation, cross-region generalization, and heterogeneous sensor integration.

S. Vorapatratorn and N. Thongsibsong (2026) provided care, an IoT-based air quality monitoring system that combines LoRa and wi-fi connection with inexpensive particle and meteorological sensors to gather data in real time. The device measures PM1.0, PM2.5, PM10, temperature, relative humidity, and atmospheric pressure, and is connected to web and mobile applications developed on the Yakkaw platform, which provide intuitive geospatial mapping and access to historical trends. A field deployment in Chiang Rai, Thailand, demonstrated that m-care achieved a communication range of up to 200 m with a packet delivery rate of 97 % under favourable conditions [35].

M. Karmoude et. al. (2026) presented the first use of spatio-temporal modelling for 24-hour forecasting using data from a cost-effective air-quality monitoring system deployed in Africa. Applied spatio-temporal deep learning models to pollutant data collected from several sites monitored by an air system, enabling the extraction of both spatial and temporal patterns relevant for forecasting [36].

S. Varshney, et al. (2025) used machine learning algorithms to predict the taj trapezium zone's air quality index. Such predictions can inform preventive action to mitigate air pollution. The study compares four algorithmic approaches: light gradient boosting machine (lightgbm), categorical boosting (catboost), adaptive boosting (adaboost), and extreme gradient boosting (xgboost). These algorithms' performance is evaluated based on several parameters, including R-squared, mean squared error (mse), root mean squared error (RMSE), and (MAE) [37].

J. D. Sinha, et al. (2025) united high-grade series temporal AQI data constructed in the form of Markov decision processes (MDP) along with thorough data pre-processing techniques such as min-max normalisation to improve learning stability and preserve causation. The model yielded a final mse of 0.00208 and an R2 score of 1.00000 alongside very low variance and policy inconsistency in received reward suggesting remarkable convergence and consistent decision-making reliability [38].

S. Lawande, et al. (2024) investigated the feasibility of using ML models for AQI prediction. The air quality data of Delhi city throughout the years is being used to construct this model. Application of exploratory data analysis (EDA) is applied to the dataset prior to delivering it as input to the

learning model. The goal is to evaluate XGBoost and Random Forest, two ML models, for their ability to forecast Delhi, India's air quality index. The air quality index (AQI) with precision, allowing for proactive or preventative measures to be implemented to manage air pollution. Ultimately, it is helpful for public health management. Along with this, are bringing attention to future research on the implementation of an air quality monitoring system using recent technologies [39].

S. K. Sunori, et al. (2024) discussed a very critical parameter – air quality index (AQI) which has a serious impact on human health. It is imperative to properly manage AQI in the current environmental discourse, and it is crucial to have reliable and accurate models to predict AQI. This study is making significant contributions in this aspect with the study of two approaches to AQI prediction modelling. One is the classical one which is known for being simple, namely linear regression model. The other one is an advanced technique based on fuzzy logic and neural networks named adaptive neuro-fuzzy inference system (ANFIS) model. The models have been developed and analyzed using MATLAB as one of the main tools [40].

F. S. Pramudya, et al. (2023) collect data on air pollution in Jakarta, Indonesia, and use it to determine the air quality index (AQI) for the city. The AQI for Jakarta is determined by analysing six major pollutants, namely PM10, PM2.5, CO, O3, SO2, and no2. The researchers have utilised geographic information system (GIS) software to visualise the data, which provides a clear and concise representation of Jakarta's air quality index. The visualisation of AQI allows policymakers to assess the severity of air pollution and take the necessary steps to improve the city's air quality [41].

Despite advances in IoT-enabled monitoring and data-driven air-quality prediction, existing studies remain fragmented across sensing, communication, forecasting, spatial analysis, and decision-making. Most approaches focus on monitoring, AQI prediction, or visualization rather than integrating these capabilities into a unified adaptive framework. ML models show promising performance, but their dependence on location-specific datasets limits generalization across cities, pollution patterns, sensor configurations, and environmental conditions. Low-cost IoT systems provide continuous measurements, yet integrating heterogeneous sensor data with spatio-temporal deep learning remains challenging. Existing forecasting approaches capture temporal or spatial dependencies but provide limited support for autonomous real-time decisions and adaptive responses. GIS-based systems mainly emphasize visualization rather than predictive control. Table I highlights these limitations and the need for an integrated framework combining IoT sensing, communication, multimodal data, deep learning, AQI forecasting, and adaptive decision support.

TABLE I. SUMMARY OF STUDIES ON ARTIFICIAL INTELLIGENCE APPROACHES FOR PREDICTING AIR QUALITY INDEX AND AIR QUALITY GRADES

Author & Year	Objective / Approach	Technology / Method	Key Findings	Advantages	Limitations
S. Vorapatratorn and N. Thongsibsong (2026)	Developed an IoT system for real-time air-quality and meteorological monitoring.	Low-cost sensors, LoRa, Wi-Fi, Yakkaw web/mobile platform, geospatial mapping.	Monitored PM1.0, PM2.5, PM10, temperature, humidity, and pressure; achieved 200 m range and 97% packet delivery rate.	Low-cost, real-time monitoring, wireless connectivity, and spatial visualization.	Communication may decline under unfavourable conditions; limited deployment and geographic validation.
M. Karmoude et al. (2026)	Investigated 24-hour AQI forecasting using cost-	Spatio-temporal deep learning, low-cost sensor networks,	Extracted spatial and temporal pollutant patterns	Combines spatial-temporal information;	Depends on sensor coverage and data quality; deep learning

	effective monitoring networks across Africa.	multi-site pollutant datasets.	for forecasting and early warning.	supports distributed monitoring and early warning.	may require high computational resources.
S. Varshney et al. (2025)	Predicted AQI in the Taj Trapezium Zone for preventive pollution management.	LightGBM, CatBoost, AdaBoost, XGBoost; R ² , MSE, RMSE, MAE.	Compared boosting-based ML models for AQI prediction using standard metrics.	Systematic model comparison; supports proactive pollution management.	Geographically focused; performance may vary across regions and datasets.
J. D. Sinha et al. (2025)	Developed a real-time decision-support framework for AQI management.	MDP, Min-Max normalization, reinforcement-learning-based decision framework.	Achieved MSE = 0.00208 and R ² = 1.00000, indicating strong convergence and decision reliability.	Supports adaptive decision-making and real-time AQI control.	Very high R ² requires validation on diverse datasets; implementation complexity may limit deployment.
S. Lawande et al. (2024)	Predicted Delhi's AQI for preventive pollution-control actions.	EDA, Random Forest, XGBoost.	Compared RF and XGBoost for AQI prediction and pollution management.	Uses established ML models; supports proactive decision-making.	Delhi-specific historical data may limit generalization; limited model diversity.
S. K. Sunori et al. (2024)	Compared approaches for accurate AQI prediction.	Linear Regression, ANFIS, MATLAB.	Applied linear regression and ANFIS to model AQI relationships.	Compares classical and intelligent approaches; ANFIS captures nonlinear relationships.	Limited to two approaches; ANFIS requires parameter tuning and higher implementation complexity.
F. S. Pramudya et al. (2023)	Assessed Jakarta's AQI and visualized spatial pollution conditions.	GIS, AQI analysis, PM10, PM2.5, CO, O ₃ , SO ₂ , NO ₂ .	GIS visualization represented AQI conditions and pollution severity.	Effective spatial visualization; integrates multiple pollutants.	Emphasizes visualization over prediction; limited forecasting and real-time capabilities.

VI. CONCLUSION AND FUTURE WORK

AI has emerged as a promising approach for improving the prediction and assessment of air quality. This review examined AI-based approaches for predicting Air Quality Index values and corresponding air-quality grades using pollutant, meteorological, sensor, satellite, and monitoring data. Machine learning and deep learning methods can identify complex nonlinear relationships, temporal variations, and pollutant interactions that influence air-quality conditions. The integration of IoT-enabled monitoring, low-cost sensors, mobile platforms, satellite observations, and multi-sensor data fusion further strengthens real-time air-quality assessment. However, prediction reliability remains influenced by data quality, missing observations, sensor calibration, anomaly detection, spatial-temporal variability, and computational requirements. Accurate AQI prediction can support early warnings, health-risk communication, pollution-control planning, and environmental management. Future AI systems should focus on integrated, scalable, and reliable frameworks that combine heterogeneous environmental data with real-time monitoring capabilities. Overall, AI-driven AQI prediction and classification can contribute significantly to timely decision-making, improved pollution management, and effective public-health protection.

Future work should focus on developing hybrid AI models that integrate machine learning, deep learning, IoT sensors, satellite observations, and meteorological data for more accurate real-time AQI prediction. Research should also address sensor reliability, missing data, spatial-temporal variability, model interpretability, and scalable deployment for effective early-warning and environmental management systems.

How to cite this article:

B. K. Sharma, "Artificial Intelligence Approaches for Predicting Air Quality Index and Air Quality Grades: A Comparative Study," *Journal of Artificial Intelligence in Governance and Public Policy* vol. 1, no. 3, pp. 29–35, Jul.–Sep. 2026. Doi-XXXX

REFERENCES

[1] A. S. Shihab, "Assessment of Air Quality through Multiple Air Quality Index Models – A Comparative Study," *J. Ecol. Eng.*, vol. 24, no. 4, pp. 110–116, Apr. 2023, doi: 10.12911/22998993/159398.

[2] D. P. Shah and D. P. Patel, "A comparison between national air quality index, India and composite air quality index for Ahmedabad, India," *Environ. Challenges*, vol. 5, p. 100356, Dec. 2021, doi: 10.1016/j.envc.2021.100356.

[3] K. Chatterjee et al., "AiM: urban air quality forecasting with grid-embedded recurrent MLP model," 2025.

[4] M. C. Mirabelli, S. Ebelt, and S. A. Damon, "Air Quality Index and air quality awareness among adults in the United States.," *Environ. Res.*, vol. 183, p. 109185, Apr. 2020, doi: 10.1016/j.envres.2020.109185.

[5] Y. Gong and P. Zhang, "Research and Realization of Air Quality Grade Prediction Based on KNN," in *2021 3rd International Conference on Artificial Intelligence and Advanced Manufacture (AIAM)*, 2021, pp. 299–304. doi: 10.1109/AIAM54119.2021.00068.

[6] K. Han and Y. Wang, "A review of artificial neural network techniques for environmental issues prediction," *J. Therm. Anal. Calorim.*, vol. 145, no. 4, pp. 2191–2207, 2021, doi: 10.1007/s10973-021-10748-9.

[7] T. Li et al., "Contributions of various driving factors to air pollution events: Interpretability analysis from Machine learning perspective," *Environ. Int.*, vol. 173, p. 107861, Mar. 2023, doi: 10.1016/j.envint.2023.107861.

[8] M. Atreya, V. B. Meruga, R. K. Kanneganti, S. R. Puli, P. Kanungo, and S. Chintala, "Federated AI Orchestration in Heterogeneous Edge Environments," in *2025 IEEE International Conference on Emerging Trends in Computing and Communication (ETCOM)*, Mangalore, India: IEEE, 2026, pp. 1–7, November. doi: 10.1109/ETCOM66606.2025.11436271.

[9] J. Dong et al., "Effect of sample number and location on accuracy of land use regression model in NO2 prediction," *Atmos. Environ.*, vol. 246, p. 118057, Feb. 2021, doi: 10.1016/j.atmosenv.2020.118057.

[10] M. Sorek-Hamer, R. Chatfield, and Y. Liu, "Review: Strategies for

- **Funding:** This research received no external funding.
- **Institutional Review Board Statement:** Not applicable.
- **Informed Consent Statement:** Not applicable.
- **Data Availability Statement:** The data supporting the findings of this study are available from the corresponding author upon reasonable request.
- **Conflicts of Interest:** The author declares no conflict of interest.

- using satellite-based products in modelling PM2.5 and short-term pollution episodes,” *Environ. Int.*, vol. 144, p. 106057, Nov. 2020, doi: 10.1016/j.envint.2020.106057.
- [11] H. Xu *et al.*, “Improved risk communications with a Bayesian multipollutant Air Quality Health Index,” *Sci. Total Environ.*, vol. 722, p. 137892, Jun. 2020, doi: 10.1016/j.scitotenv.2020.137892.
- [12] A. Kaginalkar, S. Kumar, P. Gargava, N. Kharkar, and D. Niyogi, “SmartAirQ: A Big Data Governance Framework for Urban Air Quality Management in Smart Cities,” *Front. Environ. Sci.*, vol. 10, Apr. 2022, doi: 10.3389/fenvs.2022.785129.
- [13] V. S. Limaye *et al.*, “Development of Ahmedabad’s Air Information and Response (AIR) Plan to Protect Public Health,” *Int. J. Environ. Res. Public Health*, vol. 15, no. 7, p. 1460, Jul. 2018, doi: 10.3390/ijerph15071460.
- [14] L. A. Gladson, K. R. Cromar, M. Ghazipura, K. E. Knowland, C. A. Keller, and B. Duncan, “Communicating respiratory health risk among children using a global air quality index,” *Environ. Int.*, vol. 159, p. 107023, Jan. 2022, doi: 10.1016/j.envint.2021.107023.
- [15] V. Misal, V. P. Janeja, S. C. Pallaprolu, Y. Yesha, and R. Chintalapati, “Iterative unified clustering in big data,” in *2016 IEEE International Conference on Big Data (Big Data)*, Washington, DC, USA: IEEE, 2016, pp. 3412–3421, December. doi: 10.1109/BigData.2016.7841002.
- [16] M. A. Delmas and A. Kohli, “Can Apps Make Air Pollution Visible? Learning About Health Impacts Through Engagement with Air Quality Information,” *J. Bus. Ethics*, vol. 161, no. 2, pp. 279–302, 2020, doi: 10.1007/s10551-019-04215-7.
- [17] M. Abas *et al.*, “Navigating Air Quality Assessment : A Holistic Review of Air Quality Index Metrics and Assessment Approaches,” vol. 7, no. 2, pp. 184–204, 2024.
- [18] P. Rahi, I. Singh, S. Mahajan, and R. Chauhan, “A Comprehensive Approach to Financial Market Forecasting: Integrating Fundamental and Derivative Analysis and Prediction,” in *2024 International Conference on Advances in Computing, Communication and Materials (ICACCM)*, 2024, pp. 1–6. doi: 10.1109/ICACCM61117.2024.11059083.
- [19] A. Das and H. Gupta, “Air Quality Prediction Model for Monitoring AQI,” *EPJ Web Conf.*, vol. 328, p. 01070, Jun. 2025, doi: 10.1051/epjconf/202532801070.
- [20] A. García, Y. Saez, I. Harris, and X. Huang, “Advancements in air quality monitoring : a systematic review of IoT-based air quality monitoring and AI technologies,” vol. 58, 2025.
- [21] A. Vaswani *et al.*, “Attention is all you need,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, in NIPS’17. Red Hook, NY, USA: Curran Associates Inc., 2017, pp. 6000–6010. doi: 10.48550/arXiv.1706.03762.
- [22] N. Zimmerman *et al.*, “A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring,” *Atmos. Meas. Tech.*, vol. 11, no. 1, pp. 291–313, 2018, doi: 10.5194/amt-11-291-2018.
- [23] M. Mahajan, S. Kumar, B. Pant, and U. K. Tiwari, “Incremental Outlier Detection in Air Quality Data Using Statistical Methods,” in *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)*, 2020, pp. 1–5. doi: 10.1109/ICDABI51230.2020.9325683.
- [24] S. Du, T. Li, Y. Yang, and S.-J. Horng, “Deep Air Quality Forecasting Using Hybrid Deep Learning Framework,” *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 6, pp. 2412–2424, 2021, doi: 10.1109/TKDE.2019.2954510.
- [25] İ. Kök, M. U. Şimşek, and S. Özdemir, “A deep learning model for air quality prediction in smart cities,” in *2017 IEEE International Conference on Big Data (Big Data)*, 2017, pp. 1983–1990. doi: 10.1109/BigData.2017.8258144.
- [26] Y. Colléaux, C. Willaume, B. Mohandes, J.-C. Nebel, and F. Rahman, “Air Pollution Monitoring Using Cost-Effective Devices Enhanced by Machine Learning,” *Sensors (Basel)*, vol. 25, no. 5, Feb. 2025, doi: 10.3390/s25051423.
- [27] T. Holloway *et al.*, “Satellite Monitoring for Air Quality and Health,” *Annu. Rev. Biomed. Data Sci.*, vol. 4, no. 1, pp. 417–447, Jul. 2021, doi: 10.1146/annurev-biodatasci-110920-093120.
- [28] B. A. Fubara, “Review Paper on Air Quality Assessment Techniques: Advancements and Applications (2015-2025),” *IRE Journals*, vol. 8, no. 11, 2025.
- [29] Y. Kim and S.-H. Moon, “Machine Learning-Based Quality Control for Low-Cost Air Quality Monitoring: A Comprehensive Review of the Past Decade,” *Atmosphere (Basel)*, vol. 16, no. 10, p. 1136, Sep. 2025, doi: 10.3390/atmos16101136.
- [30] H. Alrashidi, F. N. Sibai, A. Abonamah, M. Alrashidi, and A. Alsaber, “PM2.5: Air Quality Index Prediction Using Machine Learning: Evidence from Kuwait’s Air Quality Monitoring Stations,” *Sustainability*, vol. 17, no. 20, p. 9136, Oct. 2025, doi: 10.3390/su17209136.
- [31] G. K. Kang, J. Z. Gao, S. Chiao, S. Lu, and G. Xie, “Air Quality Prediction : Big Data and Machine Learning Approaches,” vol. 9, no. 1, 2018, doi: 10.18178/ijesd.2018.9.1.1066.
- [32] A. Tiwari and A. Kumar, “Air Quality Index Prediction : A Review of Machine Learning , Deep Learning , and Hybrid Approaches,” vol. 15, no. 06, pp. 1–10, 2026, doi: 10.5281/zenodo.20626963.
- [33] S. Sachdeva, R. Kaur, Kimmi, H. Singh, K. Aggarwal, and S. Kharb, “Meteorological AQI and pollutants concentration-based AQI predictor,” *Int. J. Environ. Sci. Technol.*, vol. 21, no. 5, pp. 4979–4996, Mar. 2024, doi: 10.1007/s13762-023-05307-8.
- [34] A. Bhuyan, T. Bordoloi, R. Debnath, A. M. A. Ikbai, B. Debnath, and W. S. Singh, “Assessing AQI of air pollution crisis 2024 in Delhi: its health risks and nationwide impact,” *Discov. Atmos.*, vol. 3, no. 1, p. 13, Jun. 2025, doi: 10.1007/s44292-025-00041-x.
- [35] S. Vorapatratorn and N. Thongsibsong, “Real-Time Air Quality Monitoring Device Using LoRaWAN Technology,” in *2026 18th International Conference on Knowledge and Smart Technology (KST)*, IEEE, Jan. 2026, pp. 484–489. doi: 10.1109/KST67832.2026.11432141.
- [36] M. Karmoude *et al.*, “Spatio-temporal Modelling Using a Cost-Effective System for Air Quality Monitoring,” in *2026 IEEE Conference on Artificial Intelligence (CAI)*, IEEE, May 2026, pp. 1019–1024. doi: 10.1109/CAI68641.2026.11536651.
- [37] S. Varshney, J. N. Shrivastava, and N. Gupta, “Machine Learning Algorithms for Forecasting Air Quality Index: A Predictive Analysis in the Taj Trapezium Zone (TTZ) of Agra,” in *2025 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, IEEE, Mar. 2025, pp. 1–6. doi: 10.1109/IATMSI64286.2025.10985125.
- [38] J. D. Sinha, S. Gupta, S. Samanta, B. B. Dash, U. C. De, and S. S. Patra, “Intelligent Air Quality Control Through Continuous Policy Learning,” in *2025 9th International Conference on Inventive Systems and Control (ICISC)*, IEEE, Aug. 2025, pp. 1282–1288. doi: 10.1109/ICISC65841.2025.11188208.
- [39] S. Lawande, V. N. Gavali, I. Sutar, and R. Nehe, “Comparative Analysis of Machine Learning Models for Prediction of Air Quality Index,” in *2024 International Conference on Intelligent Systems and Advanced Applications (ICISAA)*, IEEE, Oct. 2024, pp. 1–5. doi: 10.1109/ICISAA62385.2024.10829167.
- [40] S. K. Sunori, D. Verma, P. B. Negi, and P. Juneja, “Air Quality Index Prediction using Linear Regression and ANFIS,” in *2024 International Conference on Inventive Computation Technologies (ICICT)*, IEEE, Apr. 2024, pp. 1929–1933. doi: 10.1109/ICICT60155.2024.10544842.
- [41] F. S. Pramudya, F. I. Kurniadi, and A. N. Rakhiemah, “Breathing in Jakarta: Uncovering the Air Quality Index using Data Visualization,” in *2023 7th International Conference on New Media Studies (CONMEDIA)*, IEEE, Dec. 2023, pp. 162–166. doi: 10.1109/CONMEDIA60526.2023.10428349.