Archives available at journals.mriindia.com

International Journal on Advanced Computer Theory and Engineering

ISSN: 2319 - 2526

Volume 14 Issue 01. 2025

AI-Driven Air Quality: A Path to Sustainable Development

Vijaykumar S. Kumbhar
Associate Professor
Department of Computer Science, Shivaji University, Kolhapur

Peer Review Information

Submission: 08 Nov 2025

Revision: 25 Nov 2025

Acceptance: 10 Dec 2025

Keywords

AI Models, AQI, Deep Learning, SDG

Abstract

Air pollution remains a critical global environmental and public health crisis, severely impacting ecological systems and hindering progress toward the United Nations Sustainable Development Goals (SDGs), particularly SDG 3 (Good Health and Well-being) and SDG 11 (Sustainable Cities and Communities). Accurate, real-time forecasting of air quality is paramount for effective mitigation and policy formulation. This paper investigates the pivotal role of Artificial Intelligence (AI) and its associated models in enhancing the precision, efficiency, and scalability of air pollution predictions. We analyze various AI methodologies—including Machine Learning (ML), Deep Learning (DL), and hybrid models—and their comparative performance against traditional statistical techniques. Furthermore, the study critically evaluates the mechanism through which AI-driven predictive insights translate into actionable policy interventions, thereby accelerating the attainment of sustainable development objectives in urban environments. We identify current challenges, such as data quality and model interpretability, and propose future research directions, particularly the integration of geo-spatial AI and low-cost sensor networks, to foster truly smart and sustainable cities.

Introduction

The escalating rate of urbanization and industrialization has positioned air pollution as one of the most significant environmental risks to human health, contributing to millions of premature deaths annually. Conventional air quality monitoring and prediction methods often rely on sparse sensor networks and linear statistical models (e.g., ARIMA), which struggle to capture the complex, non-linear, and spatio-temporal dynamics of pollutant dispersion (e.g., PM2.5, NO2, O3). The necessity for real-time, granular, and highly accurate air quality intelligence has emerged as a fundamental requirement for informed policy-making and public health protection(Iyer et al., 2022). Consequently, advanced computational approaches, particularly those rooted in Artificial Intelligence, are increasingly being explored to overcome these limitations by providing more robust predictive capabilities (Chadalavada et al., 2024).

The Sustainable Development Imperative

Sustainable Development is defined as meeting the needs of the present without compromising the ability of future generations to meet their own needs. Achieving this requires concerted action across environmental, social, and economic pillars. Air quality management is directly linked to multiple SDGs:

- **SDG 3 (Good Health and Well-being):** Air pollution is a major risk factor for non-communicable diseases. Accurate prediction enables targeted health advisories and preventative measures. AI-driven air quality prediction can significantly contribute to reducing the burden of disease attributable to air pollution, thereby directly supporting SDG 3 by fostering healthier populations(Xu et al., 2021).

- **SDG 11 (Sustainable Cities and Communities):** Air quality is a core indicator of urban sustainability. AI-based prediction supports smart city infrastructure, such as optimized traffic flow and green urban planning(Evangelopoulos et al., 2020).
- **SDG 13 (Climate Action):** Many air pollutants and greenhouse gases share common sources (e.g., fossil fuel combustion). Mitigation strategies informed by AI predictions contribute to climate action. The nexus between air quality and climate change underscores the critical need for integrated strategies, where AI can play a pivotal role in identifying co-beneficial interventions that simultaneously improve air quality and mitigate climate impact(Parmar et al., 2022).

Research Objectives and Scope

This paper aims to:

i) Review and categorize the primary AI models used for air pollution prediction and forecasting.
ii) Evaluate the performance and comparative advantages of AI models over traditional methods in handling spatio-temporal air quality data.
iii) Analyze the pathways through which AI predictions can be integrated into urban policy and sustainable development strategies.
iv) Identify the key challenges and future research gaps in the field.

Traditional Air Pollution Prediction Models

Before the advent of AI, air quality prediction was primarily based on:

- **Statistical Time-Series Models:** Such as Auto-Regressive Integrated Moving Average (ARIMA) and variations, which model temporal dependencies but assume linearity. These models, however, often struggle to capture the intricate non-linear relationships and complex interactions present in atmospheric phenomena (Kim et al., 2024).
- **Deterministic (Chemical Transport) Models:** These simulate physical and chemical atmospheric processes but are computationally expensive and highly reliant on accurate emission inventories and boundary conditions. Furthermore, these traditional approaches frequently necessitate extensive meteorological input and detailed topographical data, limiting their applicability in data-scarce regions or for rapid forecasting needs.

The Role of Artificial Intelligence

AI models, particularly those leveraging big data from IoT sensors, satellite imagery, and meteorological stations, offer significant advantages by effectively capturing non-linear relationships and complex interactions between diverse predictor variables. Machine learning, a prominent subset of AI, has shown considerable promise in environmental research, enabling accurate air quality prediction and forecasting through advanced techniques like decision trees, support vector machines, and multilayer perceptrons (Chauhan et al., 2025; Houdou et al., 2023). Deep learning models, a further evolution of machine learning, have advanced these capabilities by adeptly processing large and complex datasets, which is crucial for modeling the intricate dependencies among pollutants for more accurate forecasts (Reddy & Sireesha, 2025).

Machine Learning (ML) Models

- **Support Vector Machines (SVM):** Effective for classification and regression tasks, particularly in smaller datasets. - Random Forest: An ensemble learning method that builds multiple decision trees and merges their predictions to produce a more accurate and stable forecast (Chadalavada et al., 2024; Houdou et al., 2023).
- **Random Forests (RF) and Gradient Boosting Machines (GBM/XGBoost):** Ensemble methods that often provide high accuracy and robust performance, capable of identifying feature importance (e.g., which meteorological factor or pollutant is most influential). - Neural Networks: These models, particularly Artificial Neural Networks, excel at identifying nonlinear relationships and have demonstrated superior performance over conventional approaches in analyzing pollutant data and predicting concentrations over time (Mottahedin et al., 2024).

Deep Learning (DL) Models

- **Artificial Neural Networks (ANN):** Foundational for modeling complex functions. Recurrent Neural Networks and Long Short-Term Memory networks are particularly adept at processing sequential data, making them highly suitable for time-series air quality prediction due to their ability to capture temporal dependencies (Garbagna et al., 2025).

- **Long Short-Term Memory (LSTM) Networks:** Superior for sequence data analysis, making them ideal for capturing long-term temporal dependencies in air quality time series. These deep learning architectures enable the extraction of intricate patterns from vast environmental datasets, leading to more accurate and real-time monitoring capabilities for air pollution levels (García et al., 2025).
- **Convolutional Neural Networks (CNN):** Excellent for spatial feature extraction, particularly when dealing with gridded or image-based pollution data (e.g., satellite aerosol optical depth). This capability is crucial for identifying pollution hotspots and understanding the spatial distribution of pollutants across urban landscapes, thereby aiding in more localized and effective intervention strategies (Mottahedin et al., 2024).
- **Hybrid Models (e.g., CNN-LSTM):** Combining the spatial feature extraction of CNNs with the temporal forecasting power of LSTMs for enhanced spatio-temporal prediction. These integrated architectures often yield superior performance by simultaneously leveraging distinct model strengths to address the multifaceted nature of air quality data (Dhongade et al., 2024).

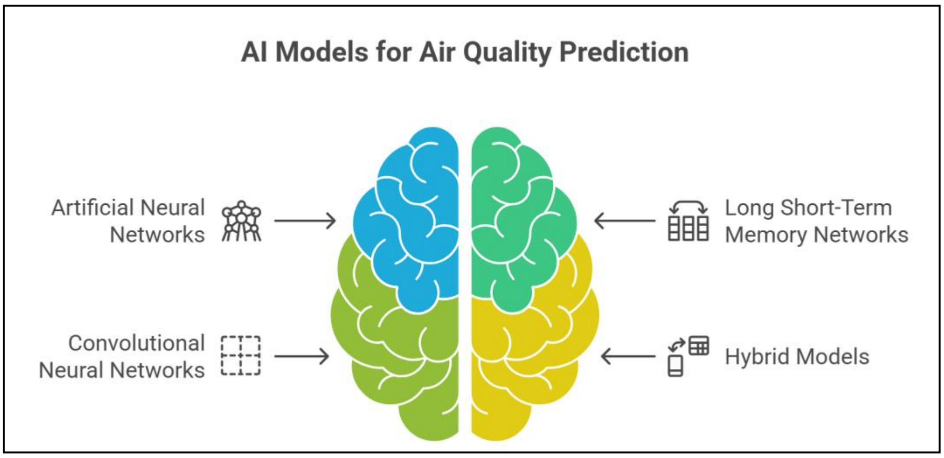

Fig. 1: AI Models for Air Quality Prediction

Case Studies in Urban Environments

Recent studies consistently demonstrate the superior performance of DL and hybrid models over traditional statistical approaches.

- **LSTM Networks:** Have shown R^2 values (coefficient of determination) as high as 0.96 for \text{PM}_{2.5} prediction, significantly outperforming linear models in multi-day forecasts. In air quality monitoring, deep learning techniques have been instrumental in addressing challenges such as missing data, noise, and the need for accurate predictions (García et al., 2025).
- **Hybrid CNN-LSTM Architectures:** Prove highly effective in capturing both localized pollution effects (spatial) and trend persistence (temporal), crucial for fine-grained urban air quality mapping. For instance, models integrating Convolutional Neural Networks for spatial feature extraction with Long Short-Term Memory networks for temporal dependency capture have demonstrated enhanced accuracy in predicting particulate matter concentrations across urban environments (Guo et al., 2025).

Performance Evaluation Metrics

Model efficacy is typically evaluated using metrics focused on prediction accuracy and error:

- **Root Mean Square Error (RMSE):** Measures the standard deviation of the residuals (prediction errors). Lower is better. - Mean Absolute Error: Provides the average magnitude of the errors, indicating the average absolute difference between predicted and actual values. - R-squared (R^2): Indicates the proportion of variance in the dependent variable that is predictable from the independent variables, with higher values signifying a better fit.
- **Mean Absolute Error (MAE):** Measures the average magnitude of the errors. Lower is better. - Coefficient of

Determination ((R^2)): Quantifies the proportion of variance in the dependent variable predictable from the independent variables, with values closer to 1 indicating better fit (Qamar et al., 2025).

- **Coefficient of Determination (R^$):** Represents the proportion of the variance for a dependent variable that's predictable from an independent variable. Closer to 1 is better. For instance, an LSTM model demonstrated superior performance with an R^2 of 0.701, an RMSE of 0.087, and an MAE of 0.056, highlighting its capability to capture temporal dependencies in complex datasets (Buestán-Andrade et al., 2025).

Enabling Proactive Policy Interventions

The core value of AI prediction lies in its ability to enable *proactive* rather than *reactive* environmental management, directly supporting sustainable outcomes.

- **Targeted Health Advisories:** Highly accurate forecasts allow governments to issue specific, localized health warnings (e.g., restricted outdoor activity) to vulnerable populations (SDG 3). These advisories can mitigate acute health impacts by informing individuals about periods of high pollution, enabling them to take precautionary measures such as reducing outdoor activities or using air purifiers (Rautela & Goyal, 2024).
- **Smart Traffic Management:** Predictive models can anticipate pollution hotspots and automatically trigger smart city infrastructure, such as:
  - Optimizing traffic light timing to reduce congestion and idling emissions.
  - Implementing temporary low-emission zones or dynamic road pricing.
- **Energy and Industrial Regulation:** Forecasting peak pollution events enables temporary, data-driven regulation of industrial emissions or power plant output, ensuring compliance with sustainability standards. This proactive approach, underpinned by AI, allows for adaptive urban planning and industrial policy adjustments, fostering a more resilient and environmentally conscious societal framework (Ravindiran et al., 2025).
- **Urban Planning:** Long-term prediction trends offer vital insights for sustainable urban planning, supporting decisions on the placement of new green spaces (which act as natural air filters) and the routing of transportation corridors. Such AI-driven insights are invaluable for developing carbon-free cities and enhancing infrastructure durability, aligning directly with sustainable urban development goals (Salhi et al., 2025).

Alignment with SDG Targets

AI predictions provide measurable indicators that help track progress toward specific SDG targets. For instance, the reduction in days exceeding unhealthy Air Quality Index (AQI) levels serves as a direct, AI-informed measure of progress toward SDG 11.6 (reducing the environmental impact of cities). Furthermore, the integration of AI-driven air quality intelligence platforms within smart city ecosystems offers a holistic approach to climate change mitigation and sustainable urban development by providing actionable insights for environmental monitoring and governance (Neo et al., 2023).

Current Challenges and Research Gaps

- **Data Quality and Heterogeneity:** The reliance on diverse data sources (official monitors, low-cost sensors, satellite) introduces challenges related to calibration, data fusion, and handling missing values. Low-cost sensors, while providing high-density data, often lack the precision of regulatory monitors. Moreover, effectively integrating these disparate data streams into a cohesive, reliable input for AI models remains a significant hurdle (Hameed et al., 2023).
- **Model Interpretability (The "Black Box" Problem):** Complex deep learning models (e.g., LSTMs) are often difficult to interpret, making it challenging for policy makers to understand *why* a specific prediction was made, which can hinder the adoption of AI-driven policy.
- **Ethical and Environmental Footprint of AI:** The high computational power required for training and running complex AI models can conflict with sustainability goals, due to significant energy consumption (SDG 13).

Future Research Trajectories

i) **Spatio-Temporal Fusion Models:** Further development of Graph Neural Networks (GNNs) and federated learning to model complex inter-station pollution flows across large urban areas in real-time.

ii) **Integration of Low-Cost Sensor Networks (LCSN):** Research is needed to develop sophisticated AI algorithms that can effectively calibrate and fuse the less reliable data from widespread, affordable sensors with high-quality reference data.

iii) **Climate Change Feedback Loops:** Developing models that predict the long-term impact of climate change (e.g., changes in wind patterns, increased heatwaves) on urban air quality and vice versa.

iv) **AI for Policy Simulation:** Creating AI-based simulation environments to model the precise air quality impact of proposed policy changes (e.g., electric vehicle adoption, industrial shutdown) *before* implementation.

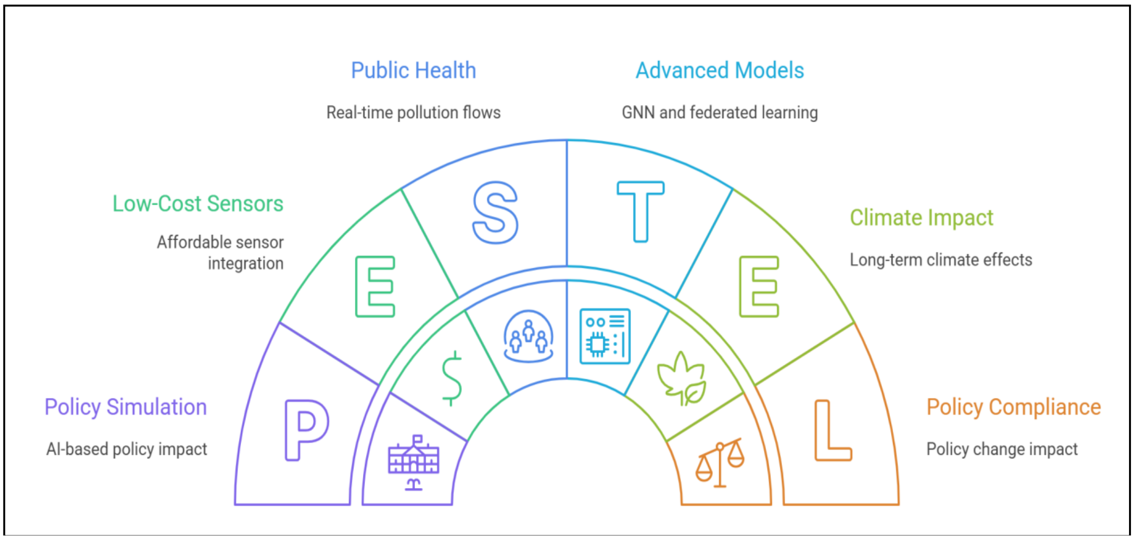

Fig. 2: Current urban Air Quality Challenges

Conclusion

The integration of Artificial Intelligence into air pollution prediction represents a transformative shift from retrospective monitoring to proactive, predictive environmental management. Advanced models, particularly deep learning and hybrid architectures, offer unprecedented accuracy and spatio-temporal resolution, providing the intelligence necessary for targeted and effective mitigation strategies. By providing a clear, data-driven path to cleaner air, AI acts as a potent enabler for key Sustainable Development Goals. While challenges in data quality and model transparency persist, ongoing research into XAI, LCSN integration, and energy-efficient algorithms promises to solidify AI's role as an indispensable tool for building healthier, resilient, and truly sustainable cities.

Acknowledgement

The author gratefully acknowledges Shivaji University, Kolhapur**,** for providing financial support to the project titled ***"Artificial Intelligence Based Air Pollution Predictions in the Wake of Sustainable Development"*** under the Research Strengthening Scheme, which enabled the successful execution of this research

References

Buestán-Andrade, P.-A., Zamora, P. E. C., Lara, Á. Z., & Piedra, J. P. P. (2025). A comprehensive evaluation of ai techniques for air quality index prediction: RNNs and transformers. *Ingenius*, *33*, 60. https://doi.org/10.17163/ings.n33.2025.06

Chadalavada, S., Faust, O., Salvi, M., Seoni, S., Raj, N., Raghavendra, U., Gudigar, A., Barua, P. D., Molinari, F., & Acharya, U. R. (2024). Application of artificial intelligence in air pollution monitoring and forecasting: A systematic review [Review of *Application of artificial intelligence in air pollution monitoring and forecasting: A systematic review*]. *Environmental Modelling & Software*, 106312. Elsevier BV. https://doi.org/10.1016/j.envsoft.2024.106312

Chauhan, B. V. S., Berg, M. J., Smallbone, K., Rautela, I., Ballal, S., & Wyche, K. P. (2025). Machine Learning Driven Prediction and Analysis of NO2 and its Catalyst Based Reduction in Urban Environments. *Topics in Catalysis*. https://doi.org/10.1007/s11244-025-02161-5

Dhongade, S., Tiwaskar, S., Koturwar, A., Motiramani, R., & Zarekar, D. (2024). *Air Predictive Modeling for Air Quality: A Comparative Study of Machine Learning and Deep*

Learning Techniques. 1. https://doi.org/10.1109/sceecs61402.2024.10482082

Evangelopoulos, D., Perez-Velasco, R., Walton, H., Gumy, S., Williams, M. L., Kelly, F. J., & Künzli, N. (2020). The role of burden of disease assessment in tracking progress towards achieving WHO global air quality guidelines [Review of *The role of burden of disease assessment in tracking progress towards achieving WHO global air quality guidelines*]. *International Journal of Public Health*, *65*(8), 1455. Springer Science+Business Media. https://doi.org/10.1007/s00038-020-01479-z

Garbagna, L., Saheer, L. B., & Oghaz, M. M. (2025). AI-driven approaches for air pollution modelling: A comprehensive systematic review [Review of *AI-driven approaches for air pollution modelling: A comprehensive systematic review*]. *Environmental Pollution*, *373*, 125937. Elsevier BV. https://doi.org/10.1016/j.envpol.2025.125937

García, A., Sáez, Y., Harris, I., Huang, X., & Collado, E. (2025). Advancements in air quality monitoring: a systematic review of IoT-based air quality monitoring and AI technologies [Review of *Advancements in air quality monitoring: a systematic review of IoT-based air quality monitoring and AI technologies*]. *Artificial Intelligence Review*, *58*(9). Springer Science+Business Media. https://doi.org/10.1007/s10462-025-11277-9

Guo, Z., Wu, S., Zhu, M., & Guandi, H. (2025). *Air Quality PM2.5 Index Prediction Model Based on CNN-LSTM*. https://doi.org/10.48550/ARXIV.2508.11215

Hameed, S., Islam, A., Ahmad, K., Belhaouari, S. B., Qadir, J., & Al-Fuqaha, A. (2023). Deep learning based multimodal urban air quality prediction and traffic analytics. *Scientific Reports*, *13*(1), 22181. https://doi.org/10.1038/s41598-023-49296-7

Houdou, A., Badisy, I. E., Khomsi, K., Abdala, S. A., Abdulla, F., Najmi, H., Obtel, M., Belyamani, L., Ibrahimi, A., & Khalis, M. (2023). Interpretable Machine Learning Approaches for Forecasting and Predicting Air Pollution: A Systematic Review [Review of *Interpretable Machine Learning Approaches for Forecasting and Predicting Air Pollution: A Systematic Review*]. *Aerosol and Air Quality Research*, *24*(1), 230151. Taiwan Association for Aerosol Research. https://doi.org/10.4209/aaqr.230151

Iyer, S. R., Balashankar, A., Aeberhard, W. H., Bhattacharyya, S., Rusconi, G., Jose, L., Soans, N., Sudarshan, A., Pande, R., & Subramanian, L. (2022). Modeling fine-grained spatio-temporal pollution maps with low-cost sensors. *Npj Climate and Atmospheric Science*, *5*(1), 76. https://doi.org/10.1038/s41612-022-00293-z

Kim, Y., Kim, J. S., Cho, S., Sim, H. S., & Kim, J. Y. (2024). Enhancing Environmental Policy Decisions in Korea and Japan Through AI-Driven Air Pollution Forecast. *Sustainability*, *16*(23), 10436. https://doi.org/10.3390/su162310436

Mottahedin, P., Chahkandi, B., Moezzi, R., Fathollahi-Fard, A. M., Ghandali, M., & Gheibi, M. (2024). Air Quality Prediction and Control Systems Using Machine Learning and Adaptive Neuro-Fuzzy Inference System. *Heliyon*, *10*(21). https://doi.org/10.1016/j.heliyon.2024.e39783

Neo, E. X., Hasikin, K., Lai, K. W., Mokhtar, M. I., Azizan, M. M., Hizaddin, H. F., Razak, S. A., & Yanto, Y. (2023). Artificial intelligence-assisted air quality monitoring for smart city management. *PeerJ Computer Science*, *9*. https://doi.org/10.7717/peerj-cs.1306

Parmar, N., Shah, R., Goswamy, T., Tandon, V., Sahu, R. P., Sutaria, R., Kar, P., & Tripathi, S. N. (2022). DELFI: Deep Mixture Models for Long-term Air Quality Forecasting in the Delhi National Capital Region. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2210.15923

Qamar, M. S., Munir, M. F., & Waseem, A. (2025). AI for Cleaner Air: Predictive Modeling of PM2.5 Using Deep Learning and Traditional Time-Series Approaches. *Computer Modeling in Engineering & Sciences*, *144*(3), 3557. https://doi.org/10.32604/cmes.2025.067447

Rautela, K. S., & Goyal, M. K. (2024). Transforming air pollution management in India with AI and machine learning technologies. *Scientific Reports*, *14*(1), 20412. https://doi.org/10.1038/s41598-024-71269-7

Ravindiran, G., Karthick, K., Rajamanickam, S., Datta, D., Das, B., Guruvare, S., Hayder, G., & Maria, A. (2025). Ensemble stacking of machine learning models for air quality prediction for Hyderabad city in India. *iScience*, *28*(2), 111894. https://doi.org/10.1016/j.isci.2025.111894

Reddy, Y. G., & Sireesha, C. (2025). Improving Air Quality Prediction: A Study on Data-Driven Techniques and Advanced Sensing Technologies. In *Lecture notes on data engineering and* *communications technologies* (p. 275). Springer International Publishing. https://doi.org/10.1007/978-981-96-0451-7_20

Salhi, A., Algarni, F. S., Alshamrani, R., Althbiti, A., Ismail, A., & Hassan, B. M. (2025). Leveraging artificial intelligence to enable sustainable urban development through the creation of smart and environmentally friendly carbon-free cities. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-16801-z

Xu, T., Yan, H., & Bai, Y. (2021). Air Pollutant Analysis and AQI Prediction Based on GRA and Improved SOA-SVR by Considering COVID-19. *Atmosphere*, *12*(3), 336. https://doi.org/10.3390/atmos12030336