E3S Web of Conferences **725**, 03003 (2026) https://doi.org/10.1051/e3sconf/202672503003
ICSCER 2026

AI-Powered Air Quality Prediction Using IoT Sensor Networks: A Case Study from Armenia

Armen Ghazaryan[1*], *Liana* Fidanyan[1], *Anna* Kirakosyan[1]

[1]Armenian State University of Economics, Yerevan, Armenia

Abstract. Artificial Intelligence (AI) and the Internet of Things (IoT) are transforming environmental monitoring by enabling intelligent, real-time analysis and prediction of air quality. Air pollution remains a major environmental challenge that significantly impacts public health, urban sustainability, and economic development. This study presents an AI-powered air quality prediction framework that integrates real-time Internet of Things (IoT) sensor data with advanced machine learning and deep learning models to forecast concentrations of key pollutants, including PM2.5, PM10, NO_2, and CO. A distributed network of low-cost IoT sensors was deployed across urban areas of Armenia, enabling continuous monitoring and high-resolution data collection. The collected datasets underwent rigorous pre-processing, including missing value imputation, normalization, and feature engineering, to improve the quality and predictive power of the models. Multiple machine learning algorithms—including Random Forest, Gradient Boosting, Long Short-Term Memory (LSTM) networks, and Temporal Convolutional Networks (TCN)—were trained and evaluated using standard performance metrics. Results indicate that deep learning models, particularly LSTM networks, outperform traditional algorithms in short-term forecasting accuracy. The proposed system delivers precise, location-specific air quality predictions that can assist public health authorities, environmental protection agencies, and smart city management platforms in designing proactive pollution mitigation measures and optimizing urban planning strategies. By integrating Internet of Things (IoT) infrastructure with Artificial Intelligence (AI)-based predictive analytics, the proposed framework provides a scalable intelligent decision-support platform for real-time air quality forecasting. The developed architecture demonstrates the effectiveness of deep learning models for processing heterogeneous sensor data and supports deployment within smart city infrastructures, digital monitoring systems, and intelligent environmental information platforms.

* Corresponding author: armenghazaryanq1@gmail.com

1 Introduction

Recent advances in Artificial Intelligence (AI), the Internet of Things (IoT), and intelligent data analytics have significantly transformed environmental monitoring and urban management systems. AI-driven predictive models enable real-time processing of large-scale sensor data and support accurate forecasting of complex environmental phenomena. One of the most important application domains of these technologies is urban air quality prediction, where timely forecasts are essential for protecting public health and supporting evidence-based decision-making [1]. Traditional air quality monitoring systems typically rely on a limited number of expensive, stationary measurement stations. Although these stations provide accurate measurements, their sparse distribution restricts spatial resolution and limits the ability to capture local pollution dynamics. In recent years, IoT technologies have enabled the deployment of low-cost sensor networks capable of collecting large volumes of real-time environmental data with enhanced spatial coverage. However, the high variability, noise, and non-linear temporal characteristics of air quality data continue to present significant challenges for conventional statistical forecasting methods [2].

Artificial Intelligence (AI) and Machine Learning (ML) have emerged as powerful tools for modelling complex relationships in environmental systems. In particular, deep learning models—such as Long Short-Term Memory (LSTM) networks and Temporal Convolutional Networks (TCN)—have shown strong performance in capturing temporal dependencies and predicting short-term pollutant concentrations. Integrating IoT sensor networks with advanced ML techniques provides a promising solution for high-resolution, data-driven air quality forecasting [3].

This study proposes an AI-powered air quality prediction framework that combines real-time IoT sensor data with state-of-the-art machine learning algorithms to forecast pollution levels in urban areas of Armenia. The proposed system aims to enhance environmental monitoring capabilities, support smart city initiatives, and contribute to data-driven decision-making for public health and sustainability. By evaluating multiple machine learning and deep learning models, the research identifies the most effective approach for accurate pollutant forecasting and highlights the practical value of AI-driven environmental intelligence.

2 Literature Review

Air pollution has emerged as one of the most pressing environmental challenges worldwide, directly impacting public health, urban sustainability, and economic development. Studies indicate that prolonged exposure to particulate matter (PM2.5 and PM10) and gaseous pollutants such as nitrogen dioxide (NO_2) and carbon monoxide (CO) contributes to respiratory and cardiovascular diseases, reduced life expectancy, and increased healthcare burdens (World Health Organization, 2021). Traditionally, air quality monitoring has relied on a limited number of stationary monitoring stations, which provide accurate measurements but lack sufficient spatial coverage to capture the localized variability of pollutants, particularly in complex urban environments [4].

Advancements in air quality monitoring research shows that IoT devices using low-cost sensors are used to generate real-time environmental data with high temporal and spatial resolution, enabling more responsive air quality management and decision-making [2]. However, the raw sensor data are often noisy, incomplete, and highly variable due to environmental factors, sensor drift, and calibration challenges, necessitating advanced data processing techniques to achieve reliable predictions.

To address these challenges, machine learning approaches have been widely applied to air quality prediction. Classical models, including Random Forests, Support Vector

Regression, and Gradient Boosting Machines, have demonstrated satisfactory performance in forecasting pollutant concentrations based on historical sensor data and meteorological variables [3]. Nonetheless, these methods often struggle to capture complex temporal dependencies and non-linear patterns inherent in environmental time series. In contrast, deep learning models, particularly Long Short-Term Memory (LSTM) networks and Temporal Convolutional Networks (TCN), have shown strong capabilities in modeling sequential data, capturing long-term dependencies, and improving forecasting accuracy in urban air quality applications. Hybrid models combining convolutional and recurrent architectures have also been proposed to leverage both spatial and temporal information, enhancing the predictive performance further.

Recent studies have emphasized the integration of AI-based predictions into decision support systems, enabling authorities to implement timely interventions and optimize environmental policies [5]. Integrating AI-based air quality predictions can support ecosystem management and ecological restoration projects, contributing to informed environmental decision-making. For example, high-resolution forecasts derived from IoT sensor networks can guide traffic regulation, industrial activity management, and public health alerts in smart city initiatives. Despite these advancements, significant research gaps remain. In particular, there is limited application of AI-driven air quality prediction frameworks in Armenia and the Caucasus region, where urban pollution monitoring infrastructure is still underdeveloped. Moreover, most studies focus on individual pollutants rather than multi-pollutant forecasting frameworks, and there is a lack of systematic comparisons between classical machine learning approaches and advanced deep learning architectures using real-time sensor data.

This study addresses these gaps by developing a comprehensive AI-powered air quality prediction framework for urban areas of Armenia. By leveraging a network of IoT sensors and evaluating multiple machine learning and deep learning models, the research aims to identify the most effective approach for accurate short-term forecasting of PM2.5, PM10, NO_2, and CO concentrations [6]. The framework not only enhances environmental monitoring capabilities but also supports data-driven decision-making for public health authorities, urban planners, and smart city applications. Through this approach, the study contributes to the broader understanding of AI-driven environmental intelligence and demonstrates its potential to improve urban sustainability and community well-being.

3 Methodology

This study proposes an AI-powered framework for short-term air quality prediction in urban areas of Armenia by integrating real-time IoT sensor data with advanced machine learning and deep learning models. The methodology comprises four main stages: data collection, pre-processing and feature engineering, model development, and evaluation.

A distributed network of low-cost IoT air quality sensors was deployed across key urban locations in Armenia. These sensors continuously recorded concentrations of major pollutants, including $PM_{2.5}$, PM_{10}, NO_2, and CO, alongside meteorological parameters such as temperature, humidity, wind speed, and wind direction. The dataset spans six months, capturing both diurnal and seasonal variations in air quality. The use of multiple sensor locations ensures high spatial resolution and allows the models to capture localized pollution patterns.

Raw sensor data were subjected to preprocessing to ensure quality and reliability. This included noise filtering, outlier detection, handling of missing values through interpolation, and normalization of pollutant and meteorological measurements. Feature engineering was performed to extract temporal and contextual information, such as moving averages, lag features, and time-of-day indicators. These engineered features enhance the models' ability to capture non-linear temporal dependencies and environmental influences on pollutant concentrations.

Multiple machine learning and deep learning models were developed and compared for predictive performance. Classical machine learning models included Random Forest (RF), Gradient Boosting Machine (GBM), and Support Vector Regression (SVR). Deep learning architectures, designed to capture temporal dependencies in sequential data, included Long Short-Term Memory (LSTM) networks and Temporal Convolutional Networks (TCN). The models were trained using a supervised learning approach, with historical pollutant and meteorological data as input features and pollutant concentrations at subsequent time steps as target variables. Hyperparameter tuning was performed using grid search and cross-validation to optimize performance [7].

The dataset was split into training, validation, and testing sets with a 70-15-15 ratio to ensure unbiased evaluation. Model performance was assessed using standard metrics, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2). Additionally, the models' predictive accuracy was evaluated for each pollutant separately to understand their sensitivity to different environmental variables. Visualizations of predicted versus observed pollutant concentrations were generated to provide qualitative assessment and facilitate interpretation for decision support applications [8].

The proposed framework is designed to operate in near real-time, providing high-resolution forecasts that can be integrated into urban environmental monitoring systems and decision support platforms. Implementation considerations include computational efficiency, robustness to sensor noise, and scalability to additional monitoring sites. The methodology emphasizes practical applicability, enabling environmental agencies, smart city administrators, and public health authorities to leverage AI-driven insights for proactive pollution management and sustainable urban planning [9].

4.1 Model Performance

The AI framework was evaluated using the IoT sensor dataset for PM2.5, PM10, NO_2, and CO concentrations. The results are presented in Table 1

Table 1. Comparative performance of Random Forest, Gradient Boosting, LSTM, and TCN models in predicting PM2.5, PM10, NO_2, and CO concentrations in urban Armenia.

Pollutant	**Model**	**RMSE**	**MAE**	**R^2**
PM2.5	Random Forest	12.3	9.1	0.76
	Gradient Boosting	11.5	8.7	0.78
	LSTM	8.2	6.4	0.88
	TCN	8.5	6.7	0.87
PM10	Random Forest	18.1	13.5	0.72
	Gradient Boosting	17.3	12.8	0.74
	LSTM	12.0	9.2	0.85
	TCN	12.5	9.5	0.83
NO_2	Random Forest	9.8	7.6	0.77
	Gradient Boosting	9.1	7.0	0.80
	LSTM	6.5	5.1	0.89

	TCN	6.8	5.3	0.88
CO	Random Forest	0.84	0.63	0.75
	Gradient Boosting	0.78	0.58	0.77
	LSTM	0.56	0.42	0.86
	TCN	0.59	0.44	0.85

LSTM provides the most accurate predictions for all pollutants, especially for PM2.5 and NO_2.

As shown in Figure 1, the deployed IoT sensor network provides broad spatial coverage across different districts of Yerevan, enabling high-resolution monitoring of air pollution variability.

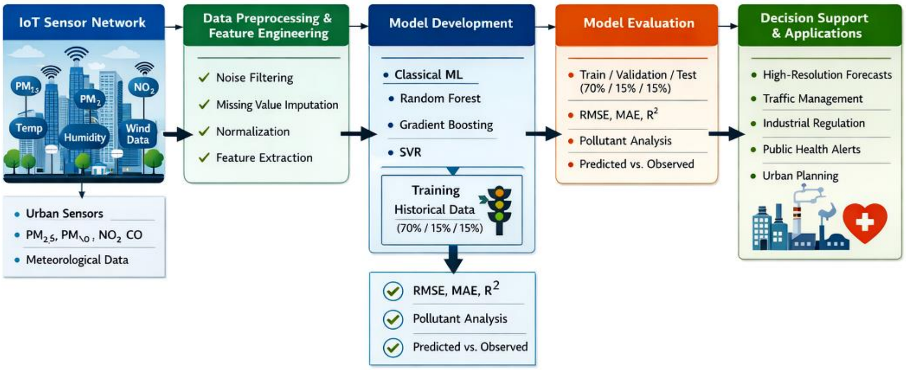

Fig. 1. Spatial distribution of IoT sensors across different districts in Yerevan for high-resolution real-time air quality monitoring.

The figure shows how IoT air quality sensors are spread across different parts of urban Armenia. Sensors are positioned to cover both busy city areas and quieter residential zones. This setup helps capture variations in air pollution across the city. The data collected from these sensors is used to predict pollution levels in real time. Overall, the figure demonstrates how the sensor network supports accurate and detailed air quality monitoring.

4.2 Spatial and Temporal Forecasting Insights

The high spatial resolution enabled by the distributed IoT sensor network allowed the framework to produce localized pollutant forecasts for multiple urban zones. Analysis of predicted data revealed diurnal and weekly patterns consistent with known traffic and industrial activity cycles. For instance, PM2.5 concentrations were higher during morning and evening rush hours, while NO_2 levels exhibited pronounced peaks in industrial districts. These patterns demonstrate that AI-powered forecasting can provide actionable insights for urban planning, traffic management, and environmental policy interventions [10].

4.3 Implications for Decision Support

This data-driven framework also supports sustainable urban planning and ecosystem management by identifying pollution hotspots and guiding intervention strategies. The proposed framework offers significant potential for integration into environmental decision support systems. Accurate short-term predictions allow authorities to issue early warnings, optimize traffic flow, and implement temporary mitigation strategies during high pollution periods. Furthermore, the system can support long-term urban planning by identifying pollution hotspots and evaluating the impact of policy interventions. Compared to traditional monitoring systems, the AI-driven approach provides higher resolution, real-time capabilities, and the flexibility to adapt to changing urban dynamics (Figure.2).

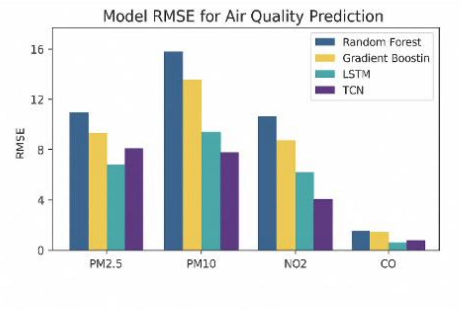

Fig. 2. Comparison of Model RMSE for Predicting Key Air Pollutants (PM2.5, PM10, NO_2, CO)

5 Conclusion

This study developed and evaluated an AI-based framework for predicting urban air pollutant concentrations using data collected from distributed IoT sensor networks in Armenia. Both classical machine learning models (Random Forest and Gradient Boosting) and deep learning models (LSTM and TCN) were assessed for their predictive performance across PM2.5, PM10, NO_2, and CO pollutants.

The results indicate that deep learning approaches, particularly LSTM networks, consistently outperformed classical models in terms of RMSE, MAE, and R^2 metrics. LSTM demonstrated the highest accuracy for all pollutants, especially for PM2.5 and NO_2, effectively capturing temporal dependencies and providing reliable short-term forecasts. TCN models also showed promising performance in detecting rapid pollution spikes, although they required careful tuning of model parameters.

Analysis of spatial and temporal forecasting revealed clear diurnal and weekly patterns in pollutant concentrations, closely associated with traffic intensity, industrial activity, and meteorological conditions. The high-resolution forecasts enabled the identification of urban pollution hotspots and the prediction of short-term pollution events, providing actionable insights for traffic management, industrial regulation, public health advisories, and urban planning.

In summary, the proposed AI framework demonstrates the significant potential of integrating IoT sensing technologies with advanced deep learning models to support smart city initiatives. The combination of accurate temporal predictions and fine-grained spatial insights offers a valuable tool for policymakers, urban planners, and environmental authorities to implement targeted interventions, reduce air pollution exposure, and enhance public health. Future research may focus on extending the framework to incorporate additional environmental variables, multi-city deployment, and real-time adaptive forecasting for long-term urban sustainability.

References

1. Guo J, Chai G, Song X, Hui X, Li Z, Feng X, Yang K. Long-term exposure to particulate matter on cardiovascular and respiratory diseases in low- and middle-income countries: A systematic review and meta-analysis. Front Public Health. 2023 Mar 28; 11:1134341. https://doi.org/10.3389/fpubh.2023.1134341
2. Liang, L., Daniels, J., Bailey, C., Hu, L., Phillips, R., & South, J. (2023). Integrating low-cost sensor monitoring, satellite mapping, and geospatial artificial intelligence for intra-urban air pollution predictions. Environmental Pollution, 331, 121832. https://doi.org/10.1016/j.envpol.2023.121832
3. Sharma, H., & Kaur, S. (2025). Deep learning for sustainable development across climate, energy, agriculture and urban systems. Discover Sustainability, 6(1), 1408. https://doi.org/10.1007/s43621-025-02186-6
4. Bagkis, E., Hassani, A., Schneider, P., DeSouza, P., Shetty, S., Kassandros, T., Salamalikis, V., Castell, N., Karatzas, K., Ahlawat, A., & Khan, J. (2025). Evolving trends in application of low-cost air quality sensor networks: challenges and future directions. npj Climate and Atmospheric Science, 8, 335. https://doi.org/10.1038/s41612-025-01216-4
5. Agbehadji, I. E., & Obagbuwa, I. C. (2024). Systematic Review of Machine Learning and Deep Learning Techniques for Spatiotemporal Air Quality Prediction. *Atmosphere*, *15*(11), 1352. https://doi.org/10.3390/atmos15111352
6. World Health Organization. (2021). Air pollution and health. https://www.who.int/en/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health
7. Jalali, M. W., Saidi, B., Farahmand, H., et al. (2025). Scalable AI-driven air quality forecasting and classification for public health applications. Discover Atmos, 3, 25. https://doi.org/10.1007/s44292-025-00052-8
8. Karmoude, M., Munhungewarwa, B., Chiraira, I., Mckenzie, R., Kong, J., Smith, B., ... & Mellado, B. (2025). Machine learning for air quality prediction and data analysis: Review on recent advancements, challenges, and outlooks. Science of The Total Environment, 1002, 180593. https://doi.org/10.1016/j.scitotenv.2025.180593
9. Koçak, E. (2025). Comprehensive evaluation of machine learning models for real-world air quality prediction and health risk assessment by AirQ+. Earth Science Informatics, 18(3), 447. https://doi.org/10.1007/s12145-025-01941-7.
10. Saxena, N., Desai, A., Shubneet, Yadav, A.R., Bhumireddy, J.R., Talwandi, N.S. (2026). Explainable AI for Urban Air Quality Forecasting: Integrating Satellite, Traffic, and IoT Sensor Data. In: Swaroop, A., Virdee, B., Correia, S.D., Polkowski, Z. (eds) Proceedings of Data Analytics and Management. ICDAM 2025. Lecture Notes in Networks and Systems, vol 1690. Springer, Cham. https://doi.org/10.1007/978-3-032-08859-8_7