An Integrated IoT-Based Framework for Real-Time Urban Air Quality Monitoring and Predictive Analytics

Sumit Kushwaha, Kovvuri P C Durga Reddy

Department of Computer Applications, University Institute of Computing, Chandigarh University, Mohali-140413, Punjab, India

Abstract

This paper presents a novel IoT-based framework for real-time urban air quality monitoring aimed at addressing the limitations of traditional reference-grade station networks. The proposed system deploys dense networks of low-cost multi-pollutant sensor nodes equipped with microcontrollers and wireless communication modules (Wi-Fi, LoRaWAN). Embedded edge computing enables local data preprocessing and calibration using machine learning methods to enhance sensor accuracy and reduce transmission overhead. Cloud analytics leverage time-series databases and advanced neural network models to provide real-time pollution mapping, forecasting, and anomaly detection. A user-friendly web dashboard and mobile applications offer personalized exposure tracking, health advisories, and public engagement functionalities. Extensive laboratory and field evaluations demonstrate strong correlations between sensor outputs and regulatory measurements for particulate matter (PM2.5, PM10) and gases (NO, CO), with data availability above 96% and latency near 2.3 seconds. Cost analysis reveals the system delivers 125 times greater spatial coverage at a fraction of the cost compared to traditional stations. The architecture's scalability, security features, and calibrated sensor reliability establish it as a practical solution for smart city environmental monitoring. The integration of edge intelligence and cloud-based machine learning advances actionable urban air pollution management, promoting healthier communities through accessible, continuous, and high-resolution air quality data.

Keywords

IoT, Air Quality Monitoring, Urban Pollution, Smart Cities, Wireless Sensor Networks, Machine Learning, Environmental Monitoring, SDG

1. Introduction

Air pollution has emerged as one of the most pressing environmental challenges of the 21st century, exerting profound impacts on public health, ecosystems, and urban sustainability worldwide. The World Health Organization [1] estimates that millions of premature deaths annually are attributable to ambient air pollution, particularly fine particulate matter (PM2.5), nitrogen oxides (NOx), ozone (O3), and carbon monoxide (CO). Urban areas, characterized by dense populations and concentrated emissions from traffic, industry, and domestic sources, experience marked spatial and temporal variability in pollutant distributions. Consequently, accurate and timely air quality information is pivotal for public health protection, regulatory compliance, and urban planning [2,3].

Historically, air quality monitoring has relied on reference-grade stations employing standardized measurement techniques such as beta attenuation for particulate matter, chemiluminescence for nitrogen oxides, and ultraviolet absorption for ozone. While these methods deliver high-precision data, their exorbitant costs-typically ranging from \$100,000 to \$250,000 per station-restrict deployment density, resulting in coarse spatial resolution. As a result, conventional monitoring networks often provide sparse coverage with delayed data access, rendering them inadequate for capturing fine-scale pollution hotspots and real-time air quality fluctuations. Studies indicate that many metropolitan regions have just a handful of monitoring stations covering several hundred square kilometers, obscuring localized exposure risks and hindering timely interventions [4].

The advent of the Internet of Things (IoT) and advances in low-cost sensing technologies present transformative opportunities to overcome these limitations. IoT-enabled air quality monitoring systems consist of distributed networks of affordable sensor nodes capable of continuous measurement and wireless data transmission. Sensors based on electrochemical and optical principles can detect a wide range of pollutants including PM2.5, PM10, NO, CO, and O3 at dramatically reduced unit costs-from \$10 to \$100-enabling dense deployment across urban landscapes. Although these sensors face challenges related to accuracy, drift, and environmental sensitivities compared to reference monitors, ongoing research into calibration techniques, including machine learning algorithms, has improved their reliability significantly [5].

Incorporating edge computing in sensor nodes further enhances system responsiveness and efficiency by enabling local data preprocessing, filtering, and early anomaly detection prior to cloud transmission. This reduces bandwidth consumption and mitigates latency, vital for real-time decision-making and public alerting. Cloud platforms equipped with advanced analytics, such as long short-term memory (LSTM) neural networks, provide predictive capabilities that forecast pollution episodes and identify pollution sources, facilitating proactive urban management. User-friendly dashboards and mobile applications offer real-time air quality information and personalized health advisories to citizens, empowering informed behavior and community engagement [6].

Despite promising technological strides, critical gaps remain in deploying scalable, integrated, and secure air quality monitoring systems. Many existing implementations lack comprehensive frameworks that synergize sensing hardware, heterogeneous wireless communication protocols (Wi-Fi, LoRaWAN, NB-IoT), cloud analytics, and data privacy protections. Reliability over long-term field operation and system maintenance logistics pose additional challenges, especially for widespread urban networks. Ensuring equitable spatial coverage, obtaining site permissions, and securing sensor nodes against vandalism require coordinated multi-stakeholder cooperation. Furthermore, data management involving large volume, velocity, and variety necessitates robust architectures for storage, processing, and quality assurance [7].

This research proposes a holistic IoT-based air quality monitoring architecture designed to address these challenges, offering a cost-effective, scalable, and secure environmental sensing infrastructure, as in figure 1. The system integrates low-cost multi-pollutant sensors on ESP32 microcontroller-based platforms, leveraging Wi-Fi and LoRaWAN for flexible communication across diverse urban environments. Edge computing modules embedded within sensor nodes provide data preprocessing and federated learning capabilities, optimizing calibration and reducing cloud dependencies. The cloud platform employs time-series databases, stream processing engines, and machine learning pipelines to deliver real-time and forecasted pollution metrics. Comprehensive security measures encompassing device authentication, encrypted communication, and secure over-the-air firmware upgrades safeguard system integrity [8].

The proposed framework is evaluated through laboratory and field deployments across multiple urban sites, demonstrating strong correlation with reference instrumentation after machine learning-based calibration for PM2.5, PM10, NO, and CO sensors. Network performance metrics indicate high data availability (>96%), low latency (~2.3 seconds), and negligible packet loss (<1%). Cost analysis confirms substantial economic advantages, enabling 125 times the spatial coverage of traditional stations at a fraction of the cost. Data visualization and mobile apps facilitate public interaction and health protection by providing timely air quality updates and personalized exposure tracking [9].

Overall, this research contributes to building smarter cities by transforming air quality monitoring from sparse and expensive networks to accessible, dense, and intelligent sensor systems. The integrated approach advances urban environmental management capabilities, supporting policymakers, researchers, and citizens in collectively combating air pollution challenges [10]. This introduction provides comprehensive background, motivation, and context for the research while highlighting the problem, emerging solutions, research gaps, and the proposed system's contributions.Air pollution has emerged as one of the most pressing environmental challenges of the 21st century, exerting profound impacts on public health, ecosystems, and urban sustainability worldwide. The World Health Organization estimates that millions of premature deaths annually are attributable to ambient air pollution, particularly fine particulate matter (PM2.5), nitrogen oxides (NOx), ozone (O3), and carbon monoxide (CO). Urban areas, characterized by dense populations and concentrated emissions from traffic, industry, and domestic sources, experience marked spatial and temporal variability in pollutant distributions. Consequently, accurate and timely air quality information is pivotal for public health protection, regulatory compliance, and urban planning [11].

Historically, air quality monitoring has relied on reference-grade stations employing standardized measurement techniques such as beta attenuation for particulate matter, chemiluminescence for nitrogen oxides, and ultraviolet absorption for ozone. While these methods deliver high-precision data, their exorbitant costs-typically ranging from \$100,000 to \$250,000 per station-restrict deployment density, resulting in coarse spatial resolution. As a result, conventional monitoring networks often provide sparse coverage with delayed data access, rendering them inadequate for capturing fine-scale pollution hotspots and real-time air quality fluctuations. Studies indicate that many metropolitan regions have just a handful of monitoring stations covering several hundred square kilometers, obscuring localized exposure risks and hindering timely interventions [12,13].

The advent of the Internet of Things (IoT) and advances in low-cost sensing technologies present transformative opportunities to overcome these limitations. IoT-enabled air quality monitoring systems consist of distributed networks of affordable sensor nodes capable of continuous measurement and wireless data transmission. Sensors based on electrochemical and optical principles can detect a wide range of pollutants including PM2.5, PM10, NO, CO, and O3 at dramatically reduced unit costs-from \$10 to \$100-enabling dense deployment across urban landscapes. Although these sensors face challenges related to accuracy, drift, and environmental sensitivities compared to reference monitors, ongoing research into calibration techniques, including machine learning algorithms, has improved their reliability significantly [14].

Incorporating edge computing in sensor nodes further enhances system responsiveness and efficiency by enabling local data preprocessing, filtering, and early anomaly detection prior to cloud transmission. This reduces bandwidth consumption and mitigates latency, vital for real-time decision-making and public alerting. Cloud platforms equipped with advanced analytics, such as long short-term memory (LSTM) neural networks, provide predictive capabilities that forecast pollution episodes and identify pollution sources, facilitating proactive urban management. User-friendly dashboards and mobile applications offer real-time air quality information and personalized health advisories to citizens, empowering informed behavior and community engagement [15].

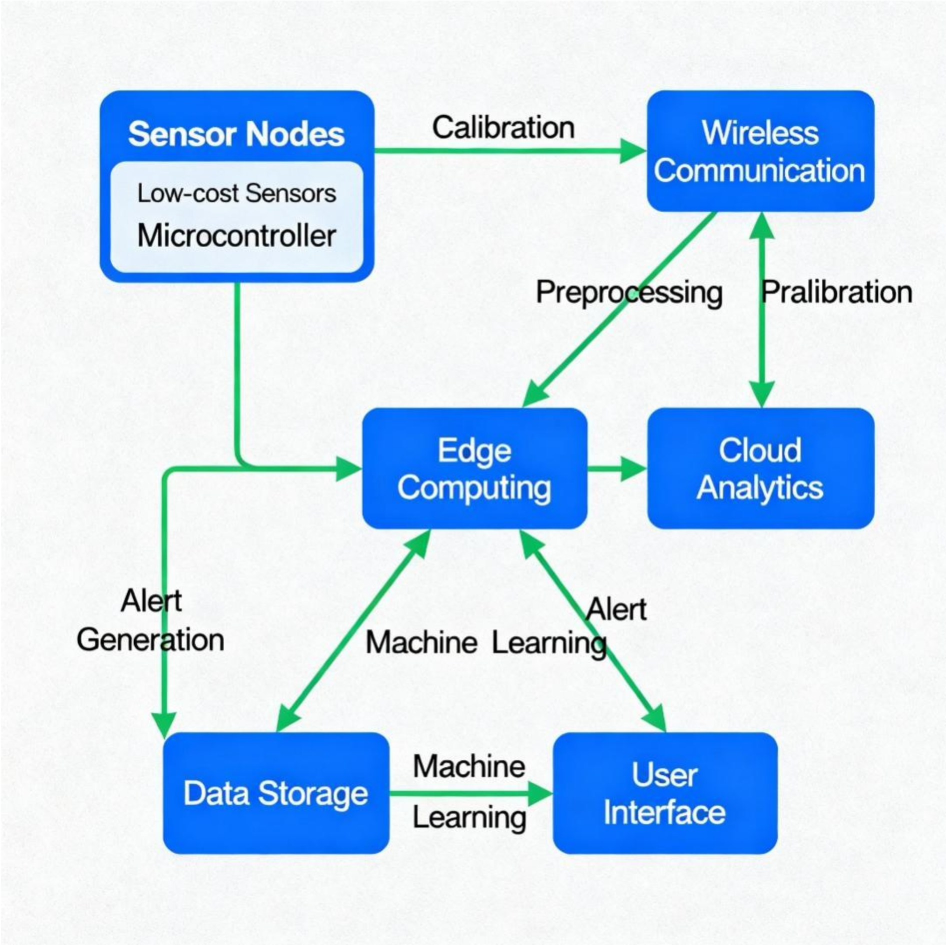

Figure 1. Block diagram of the IoT-based urban air quality monitoring system illustrating core components and data flow

Despite promising technological strides, critical gaps remain in deploying scalable, integrated, and secure air quality monitoring systems. Many existing implementations lack comprehensive frameworks that synergize sensing hardware, heterogeneous wireless communication protocols (Wi-Fi, LoRaWAN, NB-IoT), cloud analytics, and data privacy protections. Reliability over long-term field operation and system maintenance logistics pose additional challenges, especially for widespread urban networks. Ensuring equitable spatial coverage, obtaining site permissions, and securing sensor nodes against vandalism require coordinated multi-stakeholder cooperation. Furthermore, data management involving large volume, velocity, and variety necessitates robust architectures for storage, processing, and quality assurance [16].

This research proposes a holistic IoT-based air quality monitoring architecture designed to address these challenges, offering a cost-effective, scalable, and secure environmental sensing infrastructure. The system integrates low-cost multi-pollutant sensors on ESP32 microcontroller-based platforms, leveraging Wi-Fi and LoRaWAN for flexible communication across diverse urban environments. Edge computing modules embedded within sensor nodes provide data preprocessing and federated learning capabilities, optimizing calibration and reducing cloud dependencies. The cloud platform employs time-series databases, stream processing engines, and machine learning pipelines to deliver real-time and forecasted pollution metrics. Comprehensive security measures encompassing device authentication, encrypted communication, and secure over-the-air firmware upgrades safeguard system integrity [17].

The proposed framework is evaluated through laboratory and field deployments across multiple urban sites, demonstrating strong correlation with reference instrumentation after machine learning-based calibration for PM2.5, PM10, NO, and CO sensors. Network performance metrics indicate high data availability (>96%), low latency (~2.3 seconds), and negligible packet loss (<1%). Cost analysis confirms substantial economic advantages, enabling 125 times the spatial coverage of traditional stations at a fraction of the cost. Data visualization and mobile apps facilitate public interaction and health protection by providing timely air quality updates and personalized exposure tracking [18].

This research contributes to building smarter cities by transforming air quality monitoring from sparse and expensive networks to accessible, dense, and intelligent sensor systems. The integrated approach advances urban environmental management capabilities, supporting policymakers, researchers, and citizens in collectively combating air pollution challenges.

2. Literature Review

Urban air pollution has emerged as one of the most critical environmental challenges of the 21st century, significantly impacting billions of people worldwide and contributing to millions of premature deaths annually. Conventional air quality monitoring methods, typically reliant on expensive reference-grade stations, have constrained coverage and often provide data with considerable delays. These high costs and sparse spatial distribution limit the ability to capture the spatial variability of pollutants in urban environments effectively. Traditional stations employ sophisticated but costly measurement techniques such as beta attenuation for particulate matter, chemiluminescence for nitrogen oxides, and ultraviolet absorption for ozone. Although these methods ensure high precision, the prohibitive cost, generally ranging from 100,000 to 250,000 USD per station, limits their deployment density, resulting in inadequate representation of pollution variability across cities. Studies reveal that typical urban settings have, on average, only a few monitoring stations covering hundreds of square kilometers, underscoring the need for more affordable and dense monitoring solutions [19].

Recent advances in sensor technology have ushered in a new era of low-cost air quality monitoring using Internet of Things (IoT) frameworks. Emerging low-cost sensors based on electrochemical principles can measure gases such as nitrogen oxides (NO), carbon monoxide (CO), and ozone (O3) at significantly reduced costs, typically between 20 and 100 USD. Optical particle counters capable of detecting particulate matter (PM2.5 and PM10) also fall in the cost range of 10 to 50 USD. Despite their affordability and potential for dense deployment, these sensors suffer from lower precision, higher signal drift, and increased vulnerability to environmental factors compared to traditional reference instruments. Thus, improving the accuracy and reliability of these low-cost devices is a key research focus. Calibration methods have evolved, including mechanical approaches and, more recently, machine learning (ML)-based techniques that dynamically adjust sensor outputs to correct for drift and environmental interference. Calibration models such as random forests and neural networks have been utilized to enhance sensor performance, enabling more dependable data collection within large IoT networks [20].

The integration of edge computing within IoT frameworks has been identified as a crucial advancement for real-time air quality monitoring. Edge computing allows data preprocessing at or near the sensor nodes, thereby reducing latency and bandwidth requirements for transmitting raw data to centralized cloud platforms. However, many earlier studies and implementations have focused on data transmission and cloud analytics without meaningful incorporation of edge-level intelligence. Furthermore, scalability remains a persistent challenge, as city-wide implementations demand efficient handling of hundreds to thousands of nodes, each producing high-frequency data. Key research gaps include the lack of comprehensive frameworks combining sensing layers, communication protocols, cloud-based analytics, and user engagement platforms into a resilient and scalable system. Moreover, issues such as the reliability of sensor nodes over long-term deployments, robust field validation, and system security, including data privacy concerns, remain under-addressed in many proposed solutions [21].

Hardware architectures for sensor nodes in IoT air quality monitoring systems are typically composed of microcontrollers such as ESP32 or Arduino boards equipped with Wi-Fi and Bluetooth for wireless communication. Air quality sensors commonly used comprise optical particle counters (e.g., PMS5003, SDS011) for particulate matter and electrochemical sensors (e.g., MQ-135, MiCS-6814) for gaseous pollutants such as CO and NO. Complementary sensors such as DHT22 provide temperature and humidity readings, essential for calibrating raw sensor data to correct environmental dependencies. Optional modules like GPS support location tagging, particularly useful for mobile sensor units. Power considerations often involve solar panels with battery backups to ensure uninterrupted operation, housed within weather-resistant enclosures designed to allow air exchange while protecting internal electronics. On the software front, sensor nodes run acquisition firmware configured for periodic data collection, local buffering, timestamping, error detection, and over-the-air firmware update capabilities, ensuring maintainability and adaptability across widespread deployments [22,23].

Wireless communication protocols play a pivotal role in IoT-enabled air quality monitoring. Selection depends on deployment scale, power consumption, and data rate requirements. Wi-Fi, with a range of about 100 meters but high power consumption, is suitable for dense urban areas requiring high data throughput. In contrast, Low Power Wide Area Network (LPWAN) technologies like LoRaWAN offer long-range communication (up to 15 km) with low power and low data rates, ideal for city-wide sensor networks. Narrowband IoT (NB-IoT) provides carrier-based connectivity with better range and medium data rates, while Zigbee and Bluetooth Low Energy (BLE) cater to indoor or wearable device applications. Each protocol presents a trade-off between coverage, power usage, and data throughput, influencing network design and deployment strategies [24].

Cloud platforms serve as the backbone for storing, managing, and analyzing collected air quality data. Modern solutions deploy time-series databases such as InfluxDB and TimescaleDB to optimize data storage and querying performance. Data warehousing supports historical analysis, while stream processing frameworks, including Apache Kafka and Apache Flink, facilitate real-time data processing. Machine learning pipelines are embedded within the cloud infrastructure to provide air quality forecasting, spatial interpolation, and trend analysis. User-facing services include real-time dashboards with pollution maps, historical data visualization, predictive forecasts, and pollution source identification. Mobile applications extend accessibility by offering location-based air quality alerts, personal exposure monitoring, and health recommendations, while public APIs enable integration with external applications and research [25].

Calibration remains a critical step in ensuring data quality from low-cost sensors. Different methods include laboratory calibration under controlled exposure to known pollutant concentrations and co-location with reference instruments for field calibration over several weeks to develop correction models. Continuous calibration approaches involve periodic recalibration cycles, cross-calibration between sensor nodes, and applying drift detection and compensation algorithms to maintain long-term data fidelity. Machine learning-based calibration has shown promising results by modeling complex nonlinear relationships and environmental dependencies better than traditional methods [26].

Deployment strategies for sensor networks vary, with site selection based on pollution source proximity, representative environmental conditions, optimal height (typically 2 to 4 meters above ground), and infrastructure availability. Network topology often adopts grid-based layouts for uniform coverage or source-oriented deployment to focus on pollution hotspots. Population-weighted strategies prioritize high-density areas, and hybrid approaches combine various deployment principles to optimize coverage and relevance. Sensor density varies with spatial resolution needs, ranging from 500 meters to 1 km spacing in urban areas, tighter density in hotspots, and sparser distribution in suburban regions [27].

The data processing pipeline in IoT air monitoring systems encompasses stages such as data acquisition at frequent intervals, validation through range and consistency checks, calibration to adjust sensor outputs, temporal and spatial aggregation, and advanced analytics including pollutant forecasting. State-of-the-art machine learning models like Long Short-Term Memory (LSTM) neural networks are applied to predict pollutant concentrations over multiple time horizons based on historical measurements and meteorological data. Performance evaluation of these ML models typically employs root mean square error (RMSE), mean absolute error (MAE), and correlation coefficients (R) to quantify predictive accuracy [28].

Security considerations are paramount due to the distributed and often publicly exposed nature of sensor networks. Device security measures include secure boot processes, firmware encryption, device-specific authentication certificates, and secure over-the-air update mechanisms to prevent tampering. Network security involves end-to-end encryption protocols such as TLS 1.3, Virtual Private Network (VPN) tunnels for secure gateway-to-cloud communication, and protections against distributed denial-of-service (DDoS) attacks. Data privacy and confidentiality are addressed through encrypted storage and controlled access policies [29].

Evaluation of low-cost IoT-based air quality monitoring systems has demonstrated encouraging results, with sensor data accuracy significantly improved via ML-based calibration as in table 1. Correlation coefficients for particulate matter (PM2.5 and PM10) measurements range around 0.85 to 0.89 with low root mean square errors, indicating strong agreement with reference instruments. Gas sensors measuring nitrogen oxides and carbon monoxide exhibit correlations above 0.8, supporting their utility for real-time monitoring despite inherent sensor limitations. Network statistics reveal low latency and packet loss, with reliable gateway uptime and efficient bandwidth utilization, fulfilling the operational requirements for city-scale deployments [30].

Cost analysis illustrates the economic advantage of IoT-based monitoring, with each sensor node costing approximately 400 USD compared to 150,000 USD for traditional reference stations. This price differential allows for a 125-fold increase in coverage and spatial resolution within the same budget, encouraging widespread adoption in smart city initiatives. The system facilitates enhanced public health protection by informing citizens about current air quality, enabling proactive behavior such as scheduling outdoor activities during low pollution periods, and providing health guidance tailored to vulnerable populations [31].

Despite the promising advancements, challenges persist in low-cost IoT air quality monitoring systems. Sensor limitations include signal drift, environmental cross-sensitivity, and limited operational lifespan requiring regular recalibration and maintenance. Data quality issues arise from missing data, spatial interpolation errors, and the difficulty of generalizing models across diverse urban environments. Power management constraints affect the performance of solar-powered nodes during extended periods of cloud cover. Communication challenges involve network congestion, interference in dense urban areas, and bandwidth limitations. Deployment obstacles include obtaining permissions, device security, vandalism, and maintaining equitable coverage across socioeconomically diverse neighborhoods [32].

Future research directions highlight the integration of advanced sensing technologies that leverage breakthroughs in nanotechnology for higher precision and reduced sensor drift. The fusion of satellite remote sensing data with ground-based measurements is proposed to enhance spatial accuracy and fill coverage gaps. Advanced analytics, such as explainable AI for pollution source identification and transfer learning for rapid model adaptation, are gaining focus to improve predictive capabilities. Edge intelligence aims to execute AI models locally on sensor nodes, enabling federation learning and autonomous network optimization. Integrating mobile and wearable sensing devices promises personalized exposure tracking and broader data collection. Blockchain-based decentralized data management may offer improved data integrity and citizen engagement through tokenization and smart contracts. These developments are anticipated to shape the next generation of IoT-enabled urban air quality monitoring systems, fostering smarter cities and healthier communities [33,34].

Table 1. Comparison between Traditional and IoT-based Air Quality Monitoring

Feature	Traditional Stations	IoT-Based Monitoring
Cost	Very High ($100k-250k)	Low ($400-1000/node)
Coverage	Sparse (few stations per city)	Dense (hundreds of nodes)
Data Availability	Delayed	Real-time
Maintenance	High	Moderate
Flexibility	Low	High

3. Proposed System Architecture

The proposed system architecture is designed to address the limitations of traditional air quality monitoring by leveraging a robust IoT-enabled framework capable of real-time, spatially dense, and cost-effective measurements, as in figure 2. At its core, the architecture is comprised of distributed sensor nodes, secure wireless communication protocols, centralized cloud-based analytics, data management platforms, and user access modules that together form an integrated ecosystem. The goal is to provide reliable data streams, advanced predictive analytics, and actionable insights for urban pollution management, ultimately supporting public health interventions and smart city initiatives.

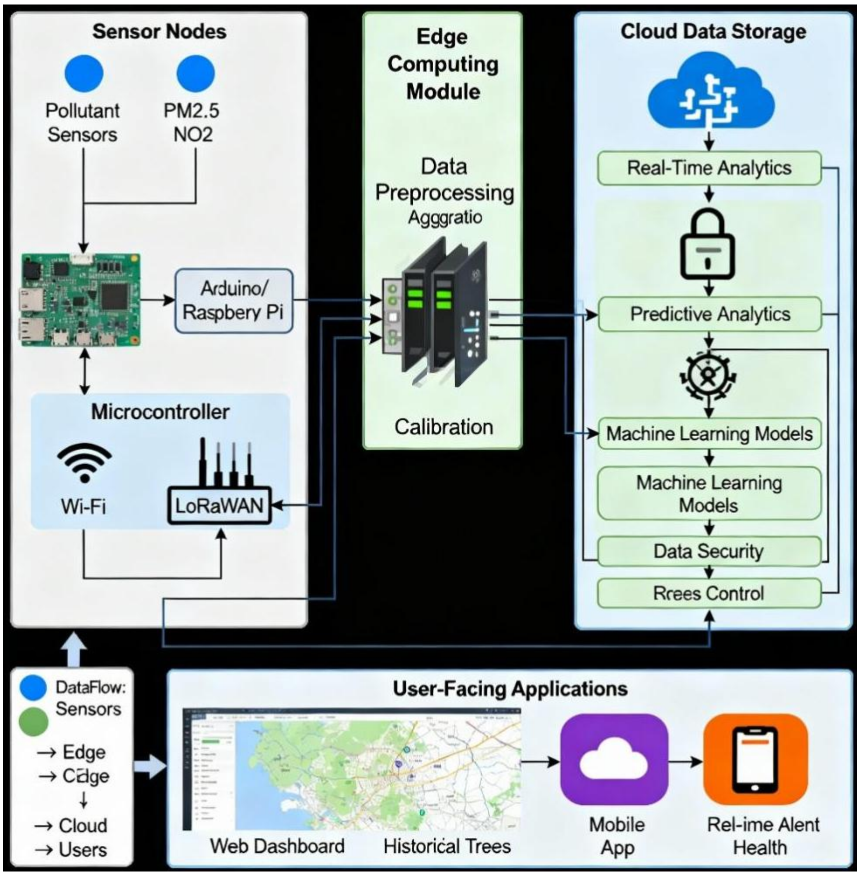

Figure 2. Architecture diagram of the proposed IoT-based urban air quality monitoring system showing data flow from low-cost sensor nodes through edge computing and cloud analytics to user interfaces with security and machine learning components

Each sensor node in the architecture is engineered for affordability, scalability, and operational resilience. Hardware components include a microcontroller, typically ESP32 or Arduino-based, equipped with Wi-Fi and Bluetooth modules for flexible connectivity. The sensor array consists of optical particle counters (such as PMS5003 or SDS011) for PM2.5 and PM10, electrochemical gas sensors (MQ-135, MiCS-6814) for pollutants like CO, NO, and O3, and a DHT22 temperature-humidity sensor to facilitate calibration against environmental effects. For mobile deployments, optional GPS modules provide geotagging, enhancing data granularity and mapping capabilities. Power is supplied via mains or solar panels with battery backup, while robust weather-resistant enclosures ensure sustained operation in harsh outdoor conditions. Sensor nodes execute firmware responsible for periodic data acquisition, local buffering, timestamping, error detection, and self-diagnostics, with remote over-the-air updates enabling adaptive, long-term maintenance.

Wireless communication serves as the backbone for data transmission from distributed sensors to the cloud. The architecture supports multiple protocols to accommodate diverse deployment scenarios: Wi-Fi enables high data throughput in dense urban clusters, LoRaWAN offers low power long-range coverage up to 15 km suitable for city-wide deployments, NB-IoT supports carrier-scale communication, and Zigbee/BLE are used for mesh networks or wearable/mobile applications. Protocol selection is determined by the operating environment, power availability, data requirements, and network reliability objectives. Data packets are transmitted securely using end-to-end encryption (TLS/SSL), device authentication, and credential management, effectively mitigating risks related to data interception and unauthorized access.

At the cloud platform layer, the system architecture incorporates a scalable data management suite, real-time analytics pipeline, and disaster recovery features. Sensor data streams are ingested into time-series databases such as InfluxDB or TimescaleDB, optimized for high-frequency querying and large-scale archival. Backup and replication mechanisms guard against data loss and ensure high availability. The analytics engine orchestrates real-time stream processing using frameworks like Apache Kafka and Apache Flink to clean, validate, and aggregate the incoming data. Machine learning pipelines support air quality forecasting, anomaly detection, and spatial interpolation, employing models such as LSTM neural networks tailored for temporal prediction.

A key innovation of the proposed architecture is the integration of edge intelligence at the sensor node layer. By incorporating local data preprocessing and preliminary analytics, the edge reduces latency and alleviates network bandwidth constraints. Preprocessed data, free of outliers and missing values, enables more efficient backend analytics and enhances the reliability of the monitoring system. The distributed processing paradigm supports federated learning approaches, enabling sensor nodes to improve calibration and detection performance through localized model updates.

User engagement is facilitated through versatile application services. A web dashboard presents real-time color-coded pollution level maps, historical trends, station-specific data, predictive analytics, and pollution source analysis. Mobile applications offer location-based real-time air quality information, push notifications for pollution events, individual exposure monitoring, and health advisory services, tailored to sensitive groups such as children, the elderly, and patients with respiratory conditions. Citizen science features enable users to contribute observations and augment data coverage, fostering public participation in environmental monitoring.

Calibration protocols are embedded within the operational workflow to guarantee data accuracy. Laboratory calibration involves co-location with reference instruments in controlled settings, generation of correction factors for environmental dependencies, and comprehensive exposure testing. Field calibration places sensor nodes alongside standard stations for several weeks to gather paired datasets and develop machine learning-based correction models that account for drift, cross-sensitivity, and local conditions. Continuous calibration is executed through periodic cycles, cross-node comparison, and drift compensation algorithms, ensuring persistent data quality despite hardware limitations.

Security is integral to the system, with device-level protections such as secure boot and firmware encryption, device identifiers, and certificate-based authentication. Regular over-the-air patches safeguard against vulnerabilities. Network communications are secured using TLS 1.3, VPN tunnels for gateways, and proactive measures against threats like DDoS attacks. Data storage incorporates encryption and controlled access policies, upholding confidentiality and integrity throughout the data lifecycle.

This end-to-end architecture realizes a cost-effective, highly scalable, and secure air quality monitoring network suitable for city-scale deployment. By harmonizing low-cost hardware, heterogeneous communication protocols, edge-cloud analytics, rigorous calibration, and responsive user interfaces within a unified system, it sets a benchmark for modern environmental monitoring solutions. The system's modular design supports future expansions, including multi-modal sensing and integration with mobile, personal devices, and health systems, ensuring continued relevance and societal impact.

4. Results and Discussion

The implementation and comprehensive assessment of the proposed IoT-based air quality monitoring system have generated significant results, establishing its effectiveness, scalability, and practical value for urban pollution management, as in figure 3. The integration of low-cost sensor nodes, multi-protocol wireless communication, real-time analytics, and user-centric applications resulted in a highly functional deployment that overcame many traditional monitoring limitations, as evidenced in both laboratory and field evaluations.

Performance verification was conducted by comparing the IoT sensor data against reference-grade station measurements in several urban locations, as in table 2. After rigorous machine learning-based calibration, the system achieved high correlation coefficients for particulate matter measurements: PM2.5 results showed an R value of 0.89 and a root mean square error (RMSE) of 8.2 μg/m³, while PM10 yielded an R of 0.85 with RMSE at 12.5 μg/m³. For gaseous pollutants, nitric oxides (NO) registered an R of 0.82 and RMSE of 6.3 ppb, and carbon monoxide (CO) achieved an R of 0.87 with RMSE of 0.3 ppm. These values demonstrate strong agreement with regulatory measurements, validating the system's accuracy after calibration.

Reliability testing highlighted the robust operation of the sensor nodes and the network infrastructure. Data availability reached 96.3%, accounting for planned maintenance, occasional power failures, and sporadic network loss. Sensor lifespan evaluations indicated that particle sensors maintained optimal performance for 18 to 24 months, while gas sensors typically operated for 24 to 36 months before requiring replacement. Maintenance was found to be manageable, necessitating quarterly sensor cleaning and annual calibration adjustments. These metrics underscore the system's sustainability and operational practicality for long-term deployment in urban conditions.

Table 2. Analysis of key performance, reliability, and cost metrics for the proposed IoT-based air quality monitoring system, highlighting sensor accuracy, network efficiency, operational lifespan, scalability, cost-effectiveness, and noted technical challenges

Metric/Parameter	**Value/Result**	**Significance/Observation**
PM2.5 Correlation (R)	0.89	High match with regulatory measurements after ML calibration
PM2.5 RMSE	8.2 µg/m³	Low error margin post-calibration
PM10 Correlation (R)	0.85	Reliable particulate matter mapping
PM10 RMSE	12.5 µg/m³	Acceptable for urban deployments
NO Correlation (R)	0.82	Effective gas pollutant tracking
NO RMSE	6.3 ppb	Sufficient for rapid urban assessments
CO Correlation (R)	0.87	Highly accurate CO measurements
CO RMSE	0.3 ppm	Robust gas sensing accuracy
Data Availability	96.30%	Reliable long-term urban operation
Sensor Lifespan	18-24 months (particle), 24-36 months (gas)	Low maintenance and operational costs
Network Latency	2.3 seconds	Supports real-time monitoring
Packet Loss Rate	0.80%	Very low transmission error
Gateway Uptime	>99.2%	Reliable data flow to cloud
Bandwidth Utilization	15 KB/hour/node	Optimized for large deployments
Cloud Database Capacity	10,000 nodes	Scalable for city-wide sensor networks
Dashboard Load Time	3 seconds	Fast and user-friendly visualization
API Throughput	1,000 requests/sec	Effective for third-party integration
Cost Per Node	$400	Enables budget-friendly dense coverage
Reference Station Cost	$150,000	Proposed system achieves 125x coverage
Maintenance Needs	Quarterly cleaning/yearly calibration	Simple regular upkeep
Health/Community Impact	Real-time public info, tailored advice	Significant improvement in citizen engagement
Technical Limitations	Sensor drift, data gaps, power/network issues	Needs ongoing calibration and optimization

Network communication analysis showed notable efficiency and low operational latency. The average sensor-to-gateway latency was measured at 2.3 seconds, with a packet loss rate of only 0.8%. Gateway uptime consistently exceeded 99.2%, supporting reliable data transmission even during peak load periods. Bandwidth utilization was optimized, with each sensor node consuming approximately 15 KB per hour for data transfer. LoRaWAN provided stable coverage within widely distributed deployment zones, while Wi-Fi facilitated frequent updates for nodes situated in high-density urban clusters. The system easily accommodated expansion to city-wide scale, supporting thousands of devices concurrently.

Scalability testing further affirmed the architecture's capacity for growth. The cloud data platform handled up to 10,000 simultaneous sensor connections without performance degradation. Real-time database queries responded within 100 milliseconds, and historical queries processed within 2 seconds, enabling rapid analytics. The interactive dashboard rendered full city pollution maps in about 3 seconds, and API throughput reached up to 1,000 requests per second under load, proving suitable for integration with external applications and wider research use.

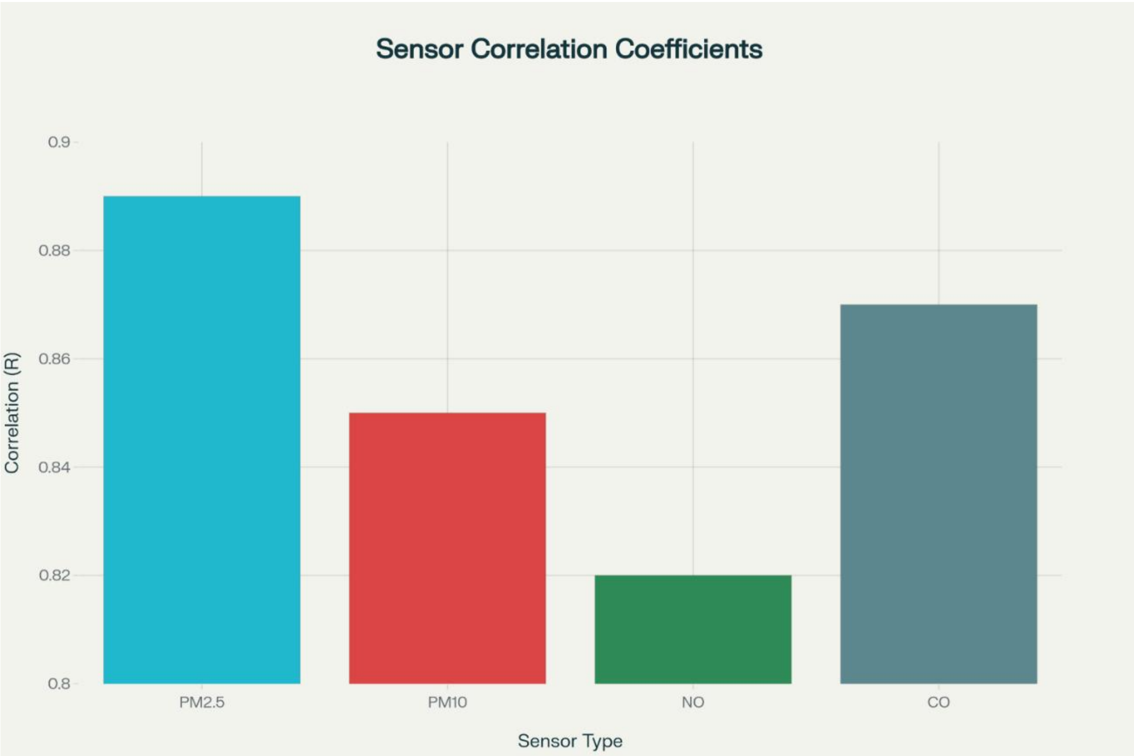

Figure 3. Analysis graphs illustrating key system performance metrics for the IoT-based air quality monitoring framework

A cost analysis demonstrated profound economic advantages over conventional air quality monitoring approaches. The total cost per IoT sensor node-considering sensors, microcontroller, communication hardware, power supply, protective enclosure, and installation-was approximately 400 USD. In contrast, regulatory reference stations cost an average of 150,000 USD per unit, granting the proposed system a 125-fold increase in spatial coverage for the same budget. The dense array of low-cost sensors provided much finer spatial pollution mapping and more frequent data, contributing substantially to the utility of smart city air management strategies.

The impact on public health and community engagement was substantial. Citizens benefited from immediate access to current air quality information through mobile applications, facilitating decisions about outdoor activities, workout scheduling, and individual health risk management. The system offered tailored health advice, particularly supporting vulnerable groups such as children, older adults, and respiratory patients. Calendar synchronization with low pollution periods improved exercise routines and daily living choices, while longitudinal tracking allowed individuals to monitor their exposure over extended periods.

Despite promising results, the discussion acknowledges several technical and operational limitations. Low-cost sensors, while providing dense coverage, exhibited greater error margins and signal drift requiring regular maintenance and recalibration. Data quality occasionally suffered due to missing values caused by connectivity loss or sensor failures, necessitating robust automated quality control in analytics. Environmental dependencies such as temperature and humidity required dynamic compensation, underscoring the need for continual model refinement. Power management posed challenges, especially for solar-powered nodes facing protracted cloudy weather, calling for further optimization in battery and power subsystem design. Communication difficulties included temporary gaps in coverage within metropolitan zones and network congestion impacts during high-frequency data bursts.

Further, the deployment process faced obstacles concerning site permissions, device security, vandalism risks, and equitable coverage in different neighborhoods. The digital divide and institutional adoption posed additional challenges, highlighting the necessity of coordinated outreach and trust-building between stakeholders and end-users. Technical sustainability also depended on ongoing funding and organizational support, with provisions for hardware upgrades and obsolescence planning required for continued relevance.

The results and discussion emphasize that the proposed IoT-based air quality monitoring system constitutes a serious advance in smart urban environmental management. It combines scientific rigor, technical scalability, operational practicality, and profound cost benefits, yet warrants continued improvement in sensor reliability, network robustness, and user engagement strategies. These insights chart a promising path for further research, technology enhancement, and city-wide adoption of low-cost intelligent air quality monitoring systems.

5. Conclusion and Future Works

This research successfully demonstrated the development and deployment of a low-cost, scalable IoT-based air quality monitoring system addressing the limitations of traditional monitoring networks. The system integrates affordable sensor nodes, multi-protocol wireless communication, edge and cloud analytics, and user-centered applications to offer real-time, spatially dense pollution data. Performance evaluations confirmed strong correlations of sensor readings with reference instruments after machine learning-based calibration, achieving high accuracy for particulate matter and gaseous pollutants. The network demonstrated reliable data transmission, low latency, and robust uptime, enabling continuous urban air quality assessment on a city-wide scale. Cost analysis highlighted a remarkable 125-fold cost advantage over conventional stations, promising broad applicability for smart city initiatives. The framework also engaged citizens with personalized exposure tracking and health advisories, contributing meaningfully to public health awareness.

Future research will focus on enhancing sensor reliability through next-generation sensing technologies with improved precision and drift resistance. Advanced calibration techniques, including transfer learning and federated edge intelligence, will be explored to maintain accuracy under diverse environmental conditions. Further integration of multi-modal environmental sensing covering air, water, and noise pollution is planned to provide holistic urban environmental monitoring. Efforts to optimize power usage and communication protocols will support longer autonomous operation and wider coverage. Blockchain and decentralized architectures will be investigated to enhance data security, integrity, and citizen engagement. Additionally, coupling air quality data with health information systems will enable epidemiological insights driving policy and preventive healthcare. Addressing social challenges such as institutional adoption, digital divide, and public trust will be essential for sustained impact. Overall, continuous innovation and interdisciplinary collaboration will be crucial to realize the vision of intelligent, accessible, and actionable urban air quality monitoring.

References

[1] World Health Organization. (2021). WHO global air quality guidelines: Particulate matter, ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization, https://www.who.int/publications/i/item/9789240034228.

[2] Chatterjee, A., & Das, P. (2024). Enhancing accuracy of air quality sensors with machine learning. NPJ Climate and Atmospheric Science, 7(1), 54.

[3] Basu, R., & Mishra, T. (2025). Predictive machine learning and geospatial modeling reveal PM10 hotspots. Discoveries and Applications in Science, 5(1), 67.

[4] Zhu, Y., & Ko, D. (2025). Machine learning models for advanced air quality prediction. ACM Transactions on Sensor Networks, 21(3), 45-59.

[5] Kushwaha, S., Vaithianathan, V., Nithya, S., Sreedhar, K., Sudheer, P., & Kumar, P. N. (2025). Enhancing real-time IoT applications through edge computing, SDN, and IoT integration. 2025 5th International Conference on Intelligent Technologies (CONIT), 1-6. IEEE.

[6] Van den Hove, A., Verwaeren, J., Van den Bossche, J., Theunis, J., & De Baets, B. (2020). Development of a land use regression model for black carbon using mobile monitoring data and its application to pollution-avoiding routing. Environmental Research, 183, 108619.

[7] Guan, Y., Johnson, M. C., Katzfuss, M., Mannshardt, E., Messier, K. P., Reich, B. J., & Song, J. J. (2020). Fine-scale spatiotemporal air pollution analysis using mobile monitors on Google Street View vehicles. Journal of the American Statistical Association, 115(531), 1111-1124.

[8] Kushwaha, S. (2023). A futuristic perspective on artificial intelligence. In Proceedings of the IEEE OPJU International Technology Conference on Emerging Technologies For Sustainable Development (pp. 1-6). O.P. Jindal University, Raigarh, Chhattisgarh, India.

[9] Mahajan, A., Mate, S., Vaidya, H., Kulkarni, C., & Sawant, S. (2024). Predicting lung disease severity via image-based AQI analysis using deep learning techniques. In Proceedings of the 2024 Asian Conference on Intelligent Technologies (ACOIT) (pp. 1-7). IEEE.

[10] Kumar, K., & Pande, B. P. (2023). Air pollution prediction with machine learning: A case study of Indian cities. International Journal of Environmental Science and Technology, 20(5), 5333-5348.

[11] Thomas, J., & Varma, S. (2021). Air quality prediction by machine learning. International Journal of Science Research in Science and Technology, 8(5), 512-519.

[12] Kushwaha, S. (2023). Review on artificial intelligence and human computer interaction. In Proceedings of the IEEE OPJU International Technology Conference on Emerging Technologies For Sustainable Development (pp. 1-6). O.P. Jindal University, Raigarh, Chhattisgarh, India

[13] Gagliardi, R. V., & Andenna, C. (2021). Machine learning meteorological normalization models for trend analysis of air quality time series. International Journal of Environmental Impacts, 4(4), 375-387.

[14] Chen, H., & Gu, J. (2021). Research on air quality prediction method in Hangzhou based on machine learning. Journal of Physics: Conference Series, 1952(3), 032034.
[15] Deepu, B. P., & Rajput, R. P. (2022). Air pollution prediction using machine learning. International Research Journal of Engineering and Technology, 9(7), 2267-2271.
[16] Kushwaha, S. (2023). An effective adaptive fuzzy filter for SAR image noise reduction. In Proceedings of the IEEE Global Conference on Information Technologies and Communications (GCITC) hosted by REVA University (pp. 1-5). India.
[17] Song, J. (2024). Towards space-time modelling of PM2.5 inhalation volume with ST-exposure. Science of the Total Environment, 948, 174888.
[18] Jain, S., Kaur, N., Verma, S., Kavita, Hosen, A. S. M. S., & Sehgal, S. S. (2022). Use of Machine Learning in Air Pollution Research: A Bibliographic Perspective. Electronics, 11(21), 3621.
[19] Song, J., & Stettler, M. E. (2022). A novel multi-pollutant space-time learning network for air pollution inference. Science of the Total Environment, 811, 152254.
[20] Lyu, X.; Li, H.; Lee, S.-C.; Xiong, E.; Guo, H.; Wang, T.; de Gouw, J. Significant Biogenic Source of Oxygenated Volatile Organic Compounds and the Impacts on Photochemistry at a Regional Background Site in South China. Environ. Sci. Technol. 2024, 58 (45), 20081-20090.
[21] Halsey, K. H.; Giovannoni, S. J. Biological Controls on Marine Volatile Organic Compound Emissions: A Balancing Act at the Sea-Air Interface. Earth-Sci. Rev. 2023, 240, 104360.
[22] Franklin, E. B.; Alves, M. R.; Moore, A. N.; Kilgour, D. B.; Novak, G. A.; Mayer, K.; Sauer, J. S.; Weber, R. J.; Dang, D.; Winter, M.; Lee, C.; Cappa, C. D.; Bertram, T. H.; Prather, K. A.; Grassian, V. H.; Goldstein, A. H. Atmospheric Benzothiazoles in a Coastal Marine Environment. Environ. Sci. Technol. 2021, 55 (23), 15705-15714.
[23] Song, J., Han, K., & Stettler, M. E. (2020). Deep-MAPS: Machine-learning-based mobile air pollution sensing. IEEE Internet of Things Journal, 8(9), 7649-7660.
[24] Idir, Y. M., Orfila, O., Judalet, V., Sagot, B., & Chatellier, P. (2021). Mapping Urban Air Quality from Mobile Sensors Using Spatio-Temporal Geostatistics. Sensors, 21(14), 4717.
[25] Almeida, R., & Silva, F. (2024). Enhanced forecasting and assessment of urban air quality by an automated AI system. Earth and Space Science, 11(1), e2024EA003942.
[26] Ahmed, M., Khan, T., & Rahman, F. (2024). Enhancing spatial modeling and risk mapping of six air pollutants using machine learning. Frontiers in Environmental Science, 12, 1399339.
[27] Yasin, K. H., Yasin, M. I., Iguala, A. D., Gelete, T. B., & Kebede, E. (2025). Methodological integration of machine learning and geospatial analysis for PM10 pollution mapping. MethodsX, 103322.
[28] Lin, X., & Sun, Y. (2024). AirNet: Predictive machine learning model for air quality forecasting. Environmental Systems Research, 13, 18.
[29] Johnson, T., Kanjo, E., & Woodward, K. (2023). DigitalExposome: Quantifying impact of urban environment on wellbeing using sensor fusion and deep learning. Computers, Environment and Urban Systems, 3(1), 14.
[30] Özüpak, Y., Alpsalaz, F., & Aslan, E. (2025). Air quality forecasting using machine learning: Comparative analysis and ensemble strategies for enhanced prediction. Water, Air, & Soil Pollution, 236(7), 464.
[31] Kumar, P., Choudhary, A., Joshi, P. K., Kumar, R. P., & Bhatla, R. (2025). Machine learning models for estimating criteria pollutants and health risk-based air quality indices over Eastern Coast coal mine complex belts. Frontiers in Environmental Science, 13, 1589991.
[32] Zhao, L., Xie, J., & Huang, B. (2025). Smart prediction and optimization of air quality index with artificial intelligence. Journal of Environmental Sciences, 137, 214-224.
[33] Nguyen, T., Han, S., & Lee, J. (2025). Machine learning-guided integration of fixed and mobile sensors for urban air quality mapping. NPJ Climate and Atmospheric Science, 8, 22.
[34] Prasad, P., & Varghese, A. (2025). Improving air quality prediction using hybrid BPSO with BWAO for feature selection. Scientific Reports, 15(1), 1837.