

Review

Machine Learning-Based Quality Control for Low-Cost Air Quality Monitoring: A Comprehensive Review of the Past Decade

Yong-Hyuk Kim  and Seung-Hyun Moon * 

School of Software, Kwangwoon University, 20 Kwangwoon-ro, Nowon-gu, Seoul 01897, Republic of Korea; yhdfly@kw.ac.kr

* Correspondence: uramoon@kw.ac.kr

Abstract

Air pollution poses major risks to public health, driving the adoption of low-cost sensor (LCS) networks for fine-grained and real-time monitoring. However, the variable accuracy of LCS data compared with reference instruments necessitates robust quality control (QC) frameworks. Over the past decade, machine learning (ML) has emerged as a powerful tool to calibrate sensors, detect anomalies, and mitigate drift in large-scale deployment. This survey reviews advances in three methodological categories: traditional ML models, deep learning architectures, and hybrid or unsupervised methods. We also examine spatiotemporal QC frameworks that exploit redundancies across time and space, as well as real-time implementations based on edge–cloud architectures. Applications include personal exposure monitoring, integration with atmospheric simulations, and support for policy decision making. Despite these achievements, several challenges remain. Traditional models are lightweight but often fail to generalize across contexts, while deep learning models achieve higher accuracy but demand large datasets and remain difficult to interpret. Spatiotemporal approaches improve robustness but face scalability constraints, and real-time systems must balance computational efficiency with accuracy. Broader adoption will also require clear standards, reliable uncertainty quantification, and sustained trust in corrected data. In summary, ML-based QC shows strong potential but is still constrained by data quality, transferability, and governance gaps. Future work should integrate physical knowledge with ML, leverage federated learning for scalability, and establish regulatory benchmarks. Addressing these challenges will enable ML-driven QC to deliver reliable, high-resolution data that directly support science-based policy and public health.

Keywords: air quality monitoring; low-cost sensors; machine learning; quality control; spatiotemporal analysis; real-time systems



Academic Editors: Klaus Schäfer, Nuria Castell and Georgios Tsegas

Received: 4 September 2025

Revised: 23 September 2025

Accepted: 25 September 2025

Published: 27 September 2025

Citation: Kim, Y.-H.; Moon, S.-H. Machine Learning-Based Quality Control for Low-Cost Air Quality Monitoring: A Comprehensive Review of the Past Decade.

Atmosphere **2025**, *16*, 1136. <https://doi.org/10.3390/atmos16101136>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Air pollution poses a significant threat to public health, driving a need for extensive air quality monitoring in both space and time. Traditionally, monitoring has relied on high-grade instruments at fixed stations (e.g., FRM/FEM reference monitors), which provide accurate data but are sparse due to high costs [1,2]. In recent years, low-cost sensor (LCS) networks have emerged as a complementary approach, enabling real-time, high-resolution observation of pollutants through dense deployment at a fraction of the cost [3,4]. These IoT-based sensor networks are being adopted worldwide to fill gaps in coverage and inform

communities about local air quality conditions [5]. However, a major drawback of LCS data is their inconsistent quality when compared with reference instruments. Discrepancies arise from differences in sensing principles and sensitivity to environmental factors, often leading to biased or noisy readings [6,7]. As a result, quality control (QC) measures are essential to detecting and correcting errors in raw sensor observations before the latter can be reliably used [8,9].

Machine learning (ML) has rapidly become a key tool for air quality data QC, enabling automated calibration, anomaly detection, and data correction that outperform traditional linear or rule-based methods [10,11]. Researchers have shown that ML models can substantially improve the agreement of low-cost sensor data with reference measurements, for example, boosting R^2 from 0.4 to 0.99 and reducing errors by orders of magnitude through proper calibration [12]. Beyond calibration, intelligent algorithms can learn normal patterns in sensor data and flag outliers or drifts in real time [13,14]. These advances are critical to smart air quality monitoring, where vast streams of sensor data must be automatically validated and corrected to ensure accuracy for end-users ranging from scientists and policymakers to citizens tracking personal exposure [15].

This survey provides a comprehensive review of the past decade (approximately 2015–2025) of research on ML-based QC for air quality sensor networks, with an emphasis on spatiotemporal techniques and real-time systems. We cover peer-reviewed literature and real-world case studies that illustrate how various ML approaches, spanning classical algorithms to deep learning, are employed to improve data quality in networked air pollution sensors. In particular, we highlight methods that exploit spatial and temporal correlations across sensors, as well as frameworks for on-line or in situ data correction suitable for real-time deployment. Key application domains such as personal exposure monitoring, integration with atmospheric models, and policy decision support are discussed to underscore the impact of these technologies.

To ensure transparency in this review process, we briefly describe the methodology used for literature selection. Publications from 2015 to 2025 were targeted to capture the most recent decade of research. The primary databases searched included Web of Science, Scopus, IEEE Xplore, and Google Scholar, using keyword combinations such as “air quality monitoring,” “low-cost sensors,” “machine learning,” “quality control,” and “calibration.” Studies were included if they were peer-reviewed or presented as well-documented case studies applying ML methods to sensor calibration, anomaly detection, or spatiotemporal QC. Papers focusing solely on hardware development or generic ML methods unrelated to air quality monitoring were excluded. This procedure yielded a representative body of work that forms the basis of the synthesis presented in the following sections.

The remainder of this paper is organized as follows: Section 2 introduces air quality sensor networks and their characteristics and contrasts reference-grade monitors with low-cost sensors. Section 3 catalogs common sources of error in atmospheric sensor observations. Section 4 surveys ML approaches to QC (traditional ML, deep learning, and hybrid/unsupervised methods) and explains how these algorithms are applied in practice. Section 5 focuses on spatiotemporal QC techniques, including inter-sensor correlation methods and representative frameworks. Section 6 reviews real-time and online QC systems, covering streaming data handling, edge deployment, federated/cluster learning, drift detection and re-calibration triggers, and hybrid edge–cloud designs. Section 7 outlines applications and impacts of improved estimation of data-driven quality personal exposure, model/simulation integration, and policy/management relevance. Section 8 discusses remaining challenges and future directions, including standardization, uncertainty quantification, scalability, and benchmarked validations. Finally, Section 9 concludes the paper.

2. Overview of Air Quality Sensor Networks

Air quality sensor networks typically consist of numerous distributed sensing nodes that measure pollutants (e.g., PM_{2.5}, NO₂, and O₃) and environmental parameters (temperature, humidity, etc.) across urban or regional areas. These nodes often utilize low-cost technologies such as electrochemical cells for gases or optical particle counters for particulates, transmitting data via IoT communication protocols to cloud servers for aggregation [1,3]. The appeal of such networks is their ability to provide spatially dense and temporally continuous data in contrast to the sparse coverage of traditional stations [1]. For instance, community sensor networks and crowdsourced platforms (e.g., PurpleAir, as well as the U.S. EPA's (Environmental Protection Agency) AirNow platform, which aggregates corrected sensor data) have deployed thousands of low-cost devices worldwide [4]. Recent evaluations of large-scale deployment, such as in Imperial County (California) and other community-driven monitoring projects, confirm that citizen-led sensor initiatives can generate valuable hyper-local insights, although with varying levels of data quality [16,17]. This high-density monitoring enables detection of neighborhood-level pollution hotspots and short-term pollution episodes that would be missed by coarse networks [5,18]. Ultimately, networked sensors promise to improve public awareness and urban air quality management by offering hyper-local data and trending information [3,19]. To clarify the complementary roles and trade-offs in such networks, we briefly contrast reference-grade monitors and low-cost sensors in Table 1.

Table 1. Reference-grade monitors vs. low-cost sensors (LCSs): typical characteristics and implications for QC.

Dimension	Reference-Grade (FRM/FEM and Regulatory)	Low-Cost Sensors (LCSs)
Traceability and QA/QC	Standardized methods; certified traceability; formal audits	Variable traceability; toolbox/tech specs; site-specific QA/QC essential [20,21]
Accuracy/LOD	Low bias; ppb level (gases) and µg/m ³ level (PM)	Higher bias/variance; cross-sensitivity to T/RH/co-pollutants [3,7,22]
Temporal resolution	Minutes to hourly (method-dependent)	Seconds to minutes; supports real-time QC/alerts
Spatial coverage	Sparse (costly siting and maintenance)	Dense mapping at neighborhood/street scale [1,17]
Drift behavior	Stable with scheduled calibration	Sensor and concept drift common; drift-aware QC required [23]
Calibration workflow	Factory + periodic QC at labs/sites	Co-location with reference; ML-based field calibration increasingly common [9,11,24]
O&M cost/staffing	High capex/opex; trained technicians	Low capex/opex; community participation feasible [17]
Regulatory use	Compliance/official reporting	Supplemental/indicative; growing acceptance with documented QC [21,25]

Abbreviations: FRM = Federal Reference Method; FEM = Federal Equivalent Method; LCS = low-cost sensor; QA/QC = quality assurance/quality control; LOD = limit of detection; O&M = operations and maintenance; T/RH = temperature/relative humidity. Sources: U.S. Environmental Protection Agency, Air Sensor Toolbox [20]; European Committee for Standardization, CEN/TS 17660-1:2021 [21].

Beyond these contrasts, low-cost sensor networks face important limitations. The data from these sensors are generally less reliable than those from reference stations, necessitating robust QC [6,23]. Many low-cost sensors are prone to measurement errors due to hardware limitations and environmental interferences. For example, low-cost optical PM sensors can overestimate concentrations at high humidity because water droplets scatter

light, unlike federal monitors, which control humidity in the sample inlet [5,22]. Electrochemical gas sensors may drift or suffer from cross-sensitivity, responding to gases other than their target, which can cause spurious readings [6,7]. Recent systematic assessments further reveal that calibration accuracy strongly depends on algorithm choice, input duration, and environmental predictors, indicating that careful preprocessing is as important as the model itself [26,27]. Moreover, manufacturing variability means that each sensor unit may have a unique bias or offset, so one-size calibration does not fit all [9,12]. Network communications can also introduce issues (data dropouts, timestamp misalignment, etc.), leading to missing or inconsistent records that require imputation [14,28]. Some recent studies have demonstrated that incorporating weather covariates or multivariate statistical models can significantly enhance reliability in such cases [28,29].

The reliance on these networks for decision making makes QC critically important. Increasingly, city authorities, researchers, and even individual citizens incorporate sensor data into health alerts, policy formulation, or personal exposure tracking [15]. For instance, smartphone apps and wearable air quality monitors use IoT sensor readings to advise users of high pollution exposure in real time [3,30]. Ensuring that such data are accurate and trustworthy is a major challenge. Basic quality assurance steps—like filtering out implausible values or applying sensor factory calibrations—are usually insufficient on their own [8]. Therefore, advanced AI and data analytics techniques are being deployed on top of sensor networks to perform dynamic calibration, drift correction, anomaly detection, and data reconciliation across the network [13,31,32]. Recent frameworks, such as HypeAIR and AIRsense, demonstrate that combining multiple anomaly detectors with calibration models can significantly improve the usability of live data streams in smart cities [14,30]. These approaches highlight a shift from passive sensing toward adaptive and self-correcting network architectures. These architectures continuously refine their outputs as conditions evolve, and Section 4 details the ML-based QC methods that enable these capabilities.

3. Sources of Error in Atmospheric Observations

Air quality observations from sensors can be corrupted by various sources of error, which QC algorithms must identify and correct. Instrument bias and calibration error are fundamental issues: low-cost sensors often have systematic biases relative to true concentration due to manufacturing differences or simplistic factory calibration [6,9,12]. Each device may read consistently high or low, requiring an offset or scaling correction. Studies have shown that even when calibration models are applied, unit-to-unit variability remains a significant barrier, underscoring the need for generalized or transferable calibration frameworks [7,11]. Calibration obtained under one set of conditions (e.g., laboratory or co-location testing) may not hold as conditions change, and sensors are also subject to gradual drift over time due to aging, fouling, or material degradation [10,23]. These issues highlight the need for adaptive re-calibration strategies and transferable models, which are further discussed in Section 8.

Another major source of error is environmental interference. Unlike reference instruments that control sampling conditions, low-cost sensors are directly exposed to ambient environmental variability. Humidity is a well-known interferent, especially for optical PM sensors: high relative humidity can cause hygroscopic growth of particles and fogging, leading to overestimation of particle mass by the sensor [3,22]. In contrast, reference PM_{2.5} monitors often include heaters or dryers to maintain constant humidity in the sample stream, thus avoiding this issue [5]. Temperature variations can similarly affect sensor baseline signals and amplifiers [6]. Many gas sensors have internal temperature compensation, but extreme temperatures or rapid changes can still introduce noise. Additionally, cross-sensitivity to non-target species (and other pollutants) plagues low-cost gas sensors;

for example, a NO₂ electrochemical sensor might respond to ozone or to strong changes in humidity, confounding its readings [8,33]. Metal-oxide sensors (MOS) for gases are especially sensitive to temperature and humidity and also require a burn-in period; their resistance measurements can drift or be disrupted by the presence of other volatile compounds [12,27]. In some cases, global data scaling and environmental differentials have been employed to partially mitigate these issues, but they remain an active challenge [32].

Physical malfunctions and outliers also occur. Sensors can suffer from malfunctions like saturation (e.g., a sudden very high reading when the sensor's range is exceeded or a voltage spike occurs) or clipping at zero. Power or circuit issues may introduce spikes or dropouts in the data [14]. Network and data handling errors can produce gaps or duplicate timestamps, which complicate downstream analysis. Some field studies have shown that wireless community networks are particularly vulnerable to such artifacts due to heterogeneous hardware and intermittent connectivity [16,17]. All of these manifest as anomalies in the time series that need to be detected. Standard meteorological QC practices often include rules for physical range checks (discarding values outside plausible bounds), time consistency checks (limiting the rate of change between readings), and persistency checks (ensuring a minimum variability) [5]. For example, Table 1 in the work by Kim et al. [5] defines realistic min/max limits and maximum rates of change for temperature, humidity, PM_{2.5}, wind, etc., based on sensor specs and physical expectations. Values violating these thresholds are flagged as errors by basic QC. These rule-based filters catch gross errors, but more subtle issues (e.g., a sensor slowly drifting or slightly biased readings under certain weather conditions) require more sophisticated approaches [13,28]. Advanced statistical models, such as multivariate Tobit regression or Bayesian neural networks, have recently been applied to improve robustness under such conditions [28,34].

In summary, the primary error sources to address in atmospheric sensor QC include sensor bias, long-term drift, environmental cross-effects (humidity, temperature, etc.), interference from other pollutants, and random anomalies or data dropouts [8,23]. Addressing these errors is the foundation upon which ML techniques are built, and the following sections describe how modern QC frameworks target each of these challenges to maintain data quality in smart monitoring networks.

4. Machine Learning Approaches to Quality Control

Machine learning (ML) provides a powerful arsenal of techniques to perform QC on air quality data. Broadly, ML-based QC methods can be categorized into traditional ML models (often supervised regression or classification algorithms), deep learning approaches (using neural network architectures to capture complex patterns), and hybrid or unsupervised methods (combining multiple algorithms or using data-driven discovery without labeled training data) [8,10]. These approaches are often complementary—for example, a pipeline may use unsupervised outlier detection followed by supervised calibration. This section surveys each category with representative examples from the literature.

Calibration (supervised regression): In LCS networks, supervised calibration models are typically trained on short co-location campaigns with reference monitors, using environmental covariates (temperature, relative humidity, co-pollutants, etc.) to correct cross-sensitivities and nonlinear biases. Tree ensembles (e.g., random forest and gradient boosting) and kernel/linear baselines (e.g., SVR and ridge) remain strong general-purpose choices for PM_{2.5} and NO₂ [9,11,24]. Recent studies report near-reference performance when appropriate predictors and window lengths are used, with R^2 frequently exceeding 0.9 for well-instrumented deployment [11,27,33]. Neural approaches further improve accuracy when interactions are complex or inputs are high-dimensional, including mixed scaling and extended inputs for particulate sensors [12,32].

Anomaly detection and data repair: Because LCS streams can contain spikes, dropouts, and device faults, unsupervised or semi-supervised detectors are layered on top of calibration. Deep sequence models (e.g., LSTM autoencoders and variational autoencoders) detect distributional shifts and recurrent artifacts, enabling automatic flagging and imputation [13,31]. Operational frameworks combine multiple detectors with repair modules so that downstream calibration is stabilized in real time (e.g., HypeAIR and AirSense) [14,30].

Spatiotemporal consistency and network-level QC: Dense networks permit cross-sensor consistency checks, neighborhood-based filtering, and spatially informed regression. Studies have leveraged spatial correlations to interpolate and validate block-level exposure maps while correcting local biases via ML [5,10,18,35]. Best-practice summaries highlight the importance of choosing predictors, durations, and validation splits appropriate to the climatology and source mix [7].

Online (real-time) vs. offline (post hoc) QC: Here, we use online QC to denote streaming corrections and anomaly handling performed on edge devices or in low-latency cloud services and offline QC to denote retrospective batch processing. Real-world systems increasingly adopt hybrid edge–cloud pipelines with drift-aware scheduling and automated re-calibration triggers tied to model diagnostics [14,23,30]. Guidance from the U.S. EPA Air Sensor Toolbox and technical standards (e.g., CEN/TS 17660-1) provide procedures to document these steps and communicate uncertainty [20,21].

4.1. Traditional Machine Learning Methods

Traditional ML approaches to sensor QC typically involve supervised learning algorithms that learn a mapping from sensor inputs to a corrected output (or an error flag) based on reference data or historical patterns. A common application is sensor calibration via regression. Here, an ML regression model is trained on co-location datasets where low-cost sensors and reference instruments perform measurements side by side, so the model can learn to predict the reference-quality concentration from the raw sensor outputs and possibly additional features (temperature, humidity, etc.) [9,12]. Researchers have tried a wide range of algorithms for this task, from simple linear and multilinear regression to more flexible nonlinear models. Comparative studies show that ensemble methods like random forest and gradient boosting tend to outperform linear calibration, especially under variable environmental conditions [6,7]. For instance, Ravindra et al. [11] calibrated low-cost PM_{2.5} sensors (PurpleAir and Atmos) using multiple ML models; the best model (a decision tree) raised R^2 values from ~ 0.40 to ~ 0.99 and reduced RMSE from tens of $\mu\text{g}/\text{m}^3$ to $<1 \mu\text{g}/\text{m}^3$ —a substantial improvement in data quality. Similarly, Koziel and colleagues demonstrated that statistical preprocessing combined with regression can substantially reduce calibration error for NO₂ sensors [27], while their later works highlight how additive/multiplicative scaling and extended calibration inputs further improve robustness across sensor units [32]. Such results underscore that traditional ML is not limited to “basic regression” but can also incorporate data transformation and feature engineering steps to handle sensor-specific variability. However, these approaches still depend heavily on the quality of reference co-location data and may fail to generalize across sensor types or environmental conditions, limiting their transferability beyond the calibration site.

Beyond calibration, traditional ML has also been applied to anomaly detection and data cleaning. One approach is to train a regression or time-series model on a rolling basis to predict the expected sensor reading and then compare the prediction to the actual observation. If the actual value deviates beyond a certain threshold, it is flagged as an outlier. Kim et al. [5] implemented this by training models on the past 10 min of data and defining an acceptable range as the ML-predicted value $\pm 3\sigma$. Any new measurement falling outside this range is classified as an error and can be replaced or corrected. Lee et al. [36] further

demonstrated the utility of support vector regression (SVR) for anomaly detection in meteorological data, optimizing input variables with a multi-objective genetic algorithm. Their framework reduced RMSE by an average of 45% compared with baseline estimators while maintaining computational efficiency, illustrating that even relatively lightweight ML models can deliver substantial improvements when paired with optimization techniques. Similar approaches have been used in operational deployment; for example, Sousàn et al. [37] reported that combining decision trees with adaptive thresholds improved detection of abnormal particulate readings in field networks. Classification-based methods have also been tested, where models such as decision trees or SVM classifiers are trained on labeled data to distinguish “normal” versus “faulty” observations. Although fault-labeled datasets are scarce, controlled co-location experiments and synthetic anomaly generation have been employed to bootstrap training sets [29]. These examples show that even relatively lightweight models can act as sophisticated real-time validators, extending rule-based checks with data-driven expectations. Nonetheless, the scarcity of representative fault data and the reliance on synthetic anomalies raise concerns about how well these models will perform under unanticipated sensor failures or new environmental conditions.

4.2. Deep Learning Approaches

Deep learning (DL) techniques have increasingly been adopted for air quality data QC, as they can model complex nonlinear relationships and spatiotemporal patterns in large datasets. One area where deep learning shines is in handling time series and sequence data from sensor networks. Recurrent neural networks (RNNs), particularly long short-term memory (LSTM) networks, are well suited to capture temporal dependencies and trends in pollution data. They have been used to predict pollutant levels based on past readings, effectively learning the temporal dynamics [13]. When used for QC, an LSTM model can play a similar role as the aforementioned regression predictor—forecasting the next value and identifying anomalies when the actual value deviates significantly. Unlike simpler models, an LSTM can leverage long-range dependencies and seasonality (diurnal cycles, weekly patterns, etc.) in the data. Convolutional neural networks (CNNs) have also been applied, sometimes by treating time-series segments or even multi-sensor data as “images” or matrices that the CNN can process. More commonly, CNNs appear in hybrid architectures (e.g., as part of a feature extractor before an LSTM network or in 1D form to capture local trends in a sequence). Recent experiments suggest that 1D CNNs coupled with environmental covariates can outperform standard regression under fluctuating weather conditions [32,33].

Deep models can combine spatial and temporal features. For example, a graph neural network or spatiotemporal CNN/LSTM can incorporate data from neighboring sensors and recent time steps to detect anomalies or fill missing values [13]. These DL models effectively learn the expected multi-dimensional structure of the data. In one recent work, Allka et al. [13] proposed a Pattern-Based Attention Recurrent Autoencoder for anomaly detection (PARAAD) in air quality sensor networks. Their model uses a bi-directional LSTM autoencoder with an attention mechanism, applied to blocks of time-series data rather than individual points. PARAAD achieved over 80% detection and localization of anomalous sensors, outperforming baseline models like standard autoencoders and even transformer-based approaches. Similar spatiotemporal DL pipelines have been explored in other urban deployment scenarios, where graph-based layers encode sensor neighborhood structures to stabilize anomaly detection and enable spatial interpolation [18,19]. However, while these architectures achieve state-of-the-art accuracy, they require large labeled datasets and significant computational resources, raising doubts about their practicality for low-power edge devices or for deployment in data-scarce regions.

Another important application of deep learning in QC is data imputation and denoising. Autoencoders (AEs) and variational autoencoders (VAEs) are unsupervised neural networks that learn a compressed representation (latent space) of the input data. They can be trained on historical sensor data so that the network learns the manifold of “normal” sensor behavior; if a new data point cannot be well reconstructed by the autoencoder, it is likely anomalous. Autoencoders and even Generative Adversarial Networks (GANs) have been used to impute missing values or repair faulty readings by essentially predicting what the sensor should have reported [5]. For instance, Bachechi et al. [30] integrated an autoencoder within their HypeAIR framework to perform on-device calibration and real-time anomaly filtering, showing the feasibility of DL at the network edge. Kim et al. [10] also applied an AE to ensure data integrity by filling gaps in an urban air quality dataset. The VAE-based method by Osman et al. [31] combined a VAE with a random forest (RF) classifier to decide if a given segment was anomalous. This hybrid deep-and-ensemble approach proved robust in identifying pollution anomalies across different scenarios without relying on extensive labeled data. Recent reviews highlight that VAE–RF and CNN–LSTM hybrids are among the most promising strategies for handling complex, multivariate air quality datasets [15,38]. Deep learning models have also demonstrated robustness to sensor noise; for instance, Zimmerman et al. [6] found that a neural network model could inherently filter out some noise and improve calibration, while Villarreal-Marines et al. [39] showed that hybrid DL calibration pipelines improved performance of field sensors in industrialized regions. Nevertheless, the complexity and opacity of deep models make them difficult to interpret and validate for regulatory acceptance, suggesting a need for explainable AI tools tailored to QC applications.

4.3. Hybrid or Unsupervised Methods

Given the diverse nature of sensor errors, hybrid approaches that combine multiple techniques often yield the best results. One strategy is to use ensemble anomaly detection, where different algorithms detect outliers from different perspectives, and their results are combined (for example, via voting or aggregation). Rollo et al. [14] introduced AIrSense, a framework that first applies three complementary anomaly detection algorithms to raw sensor signals before calibration. If at least two of the three algorithms agree a data point is anomalous, it is labeled an outlier (majority vote). After detecting anomalies, AIrSense then repairs them; if sufficient recent non-anomalous data exist, a local prediction model is trained on the past readings to estimate the true value, which is used to replace the anomaly. Finally, the cleaned data stream is passed through a calibration model to convert raw sensor units to pollutant concentrations, significantly improving calibration accuracy on real-world datasets [14].

Another type of hybrid method involves combining clustering or other unsupervised learning with regression. Kim et al. [10] used expectation–maximization (EM) clustering on smartphone barometer data to group data by time of day and trained separate regression models (SVR and MLP) on each cluster. This yielded better correction of atmospheric pressure data compared with a one-model-fits-all approach. More recently, Koziel and colleagues demonstrated that combining clustering with statistical preprocessing improved NO₂ calibration robustness under variable meteorological conditions [27]. Such approaches highlight the importance of context-aware calibration rather than relying on static global models. However, clustering-based models assume that regime boundaries are stable and identifiable, which may not hold under highly dynamic urban conditions.

In general, unsupervised QC methods like PCA or one-class SVMs are also used to detect anomalies without needing labeled examples of faults. Bayesian neural networks have also been investigated for calibration under uncertainty, providing probabilistic confi-

dence intervals that can increase user trust in automated QC decisions [34]. These methods are particularly attractive when labeled fault data are scarce or when networks operate in highly dynamic environments. Yet, their effectiveness often hinges on strong prior assumptions or careful tuning, which may reduce generality across deployment scenarios.

A recent example of combining deep unsupervised learning with classical ML is the VAE–RF hybrid by Osman et al. [31]. There, a VAE was trained on historical multivariate data (NO₂, PM_{2.5}, O₃, CO, and SO₂ plus meteorological features) to capture the normal patterns of these correlated variables [31]. The latent representations from the VAE were then input into a random forest classifier that distinguished anomalies from normal conditions. Hybrid approaches like this underscore a trend in recent research: rather than relying on any single algorithm, the best QC systems integrate multiple models and knowledge sources (statistical, physical, and machine-learning-based ones) [14,31]. This can also include hybrid physical–ML models, where known physical relationships (e.g., gas response vs. temperature or humidity correction curves) are embedded or used to inform the ML model [7,37]. For example, HypeAIR [30] illustrates how AEs and physical consistency checks can be combined for real-time QC at the network edge. While these approaches offer strong performance and flexibility, their complexity can make deployment challenging, and interoperability across heterogeneous sensor platforms remains an open issue.

The overall goal of hybrid and unsupervised approaches is to maximize accuracy and reliability by using all available information within a cohesive ML-driven QC framework. These methods are particularly effective when deployed in large heterogeneous sensor networks, where unit variability, environmental effects, and data gaps coexist [16,17]. By unifying statistical models, physical constraints, and ML predictions, hybrid QC frameworks move closer to resilient, trustworthy smart air quality monitoring systems. At the same time, their dependence on computational resources and integration complexity highlights the importance of future work on lightweight, explainable, and standardized solutions.

To synthesize methodological differences, Table 2 compares traditional ML, deep learning, and hybrid/unsupervised approaches in terms of representative models, strengths, limitations, and example studies.

Table 2. Comparison of machine learning approaches for quality control of low-cost air quality sensors.

Approach	Representative Models	Strengths	Limitations/Example Studies
Traditional ML	Linear regression, SVR, decision trees, random forest, and gradient boosting	Lightweight and interpretable; effective for calibration and anomaly detection; suitable for edge devices	Limited generalization in varying environments; requires feature engineering [7,9,11,27]
Deep Learning	CNNs, LSTM, AEs, VAEs, and GNNs	Captures nonlinearities and spatiotemporal dependencies; strong performance in anomaly detection and imputation	Computationally intensive; requires large datasets; black-box nature hinders interpretability [13,30,31,39]
Hybrid/Unsupervised	Clustering + regression, PCA, one-class SVM, VAE–RF hybrids, and physics-informed ML	Combines multiple perspectives; works without labeled data; can integrate physical knowledge	Complex pipelines; performance depends on integration; limited large-scale validation [10,14,16,34]

Table 3 summarizes representative studies on ML-based quality control for low-cost air quality sensors published over the past decade. Calibration-focused works generally report substantial performance gains, with several achieving R^2 values above 0.9 [11,27,33]. In particular, decision tree ensembles and neural networks have been shown to reduce

biases in PM_{2.5} and NO₂ measurements by an order of magnitude, bringing low-cost sensors close to reference-grade accuracy [6,29,37]. Beyond calibration, recent studies highlight the role of anomaly detection frameworks, where advanced deep learning approaches such as autoencoders or variational methods achieve robust error detection and repair in dense sensor networks [13,14,31].

Table 3. Performance of representative ML-based quality control methods for low-cost air quality sensors (2015–2025). Only results reported in the cited studies are shown.

Study	ML Method(s)	Target Pollutant(s)	Performance Summary
Zimmerman et al. (2018) [6]	RF	Low-cost gas sensors (multi-city)	Calibrated network achieved $R^2 \approx 0.80\text{--}0.90$ vs. regulatory monitors.
Ravindra et al. (2024) [11]	DT, SVR, and RF	PM _{2.5} (PurpleAir and ATMOS)	R^2 improved from ~ 0.40 to ~ 0.99 ; RMSE reduced from tens to $<1\text{ }\mu\text{g}/\text{m}^3$.
Sousan et al. (2025) [37]	GB, RF, and ANN	PM _{2.5}	Field deployment showed accuracy improvements (no quantitative metrics reported).
Taştan (2025) [29]	DT, RF, k NN, SVM, and GB	PM _{2.5} /CO ₂ / T/RH	Gradient boosting and k NN identified as most accurate (no quantitative metrics reported).
Koziel et al. (2024a) [27]	ANN surrogate + preprocessing	NO ₂	Achieved $R^2 \approx 0.95$ with low RMSE.
Koziel et al. (2024b) [33]	Differential env. inputs + global scaling	NO ₂	Reported $r > 0.9$ and RMSE $< 3.2\text{ }\mu\text{g}/\text{m}^3$.
Koziel et al. (2025) [32]	ANN with mixed scaling	PM ₁ /PM _{2.5} / PM ₁₀	R^2 up to 0.89 (PM ₁); substantial RMSE reduction.
Allka et al. (2024) [13]	Bi-LSTM AE + attention (PARAAD)	Multi-pollutant network	High detection/localization accuracy (exact % not reported here).
Osman et al. (2024) [31]	VAE + RF	Multi-pollutant	Robust anomaly detection (no quantitative metrics reported).
Rollo et al. (2023) [14]	Ensemble anomaly detection + repair (AirSense)	PM _{2.5}	Data repair significantly improved following calibration accuracy.

Abbreviations: DT = decision tree; RF = random forest; SVR = support vector regression; SVM = support vector machine; GB = gradient boosting; k NN = k -nearest neighbor; ANN = artificial neural network; AE = autoencoder; VAE = variational autoencoder; LSTM = long short-term memory; LCS = low-cost sensor; R^2 = coefficient of determination; r = Pearson correlation coefficient; RMSE = root mean squared error (in $\mu\text{g}/\text{m}^3$ where applicable).

5. Spatiotemporal Quality Control Techniques

Spatiotemporal QC techniques exploit spatial correlations and temporal dynamics to correct biases, enforce cross-sensor consistency, and fill gaps at the network level. These methods are primarily model-centric and can be executed offline or embedded within online systems; here we emphasize the algorithmic aspects (e.g., spatial regression/kriging, graph-based smoothing, spatiotemporal regularization, and cross-sensor reconciliation), while operational concerns such as latency, triggers, and fail-safes are deferred to Section 6.

A distinctive advantage of networked sensors is the ability to leverage spatiotemporal redundancy for QC. In a dense network, neighboring sensors measuring the same pollutant should exhibit similar trends (after accounting for local source differences), and each sensor’s time series typically shows temporal continuity. An anomalous reading can,

therefore, be detected by cross-checking against nearby sensors or by comparing it with the sensor's recent temporal pattern [3,4]. Following prior work by Kim et al. [5,8], we describe three modes of ML-based spatiotemporal QC (MLQC): homogeneous temporal (HT), nonhomogeneous temporal (NT), and spatiotemporal (ST). For brevity, we use HT, NT, and ST hereafter.

In the HT mode, QC uses only the time series of a single sensor/variable. A short-history predictor (e.g., using the last several minutes) forecasts the current value; large deviations from the prediction are flagged as anomalies [8]. This approach is computationally efficient and sensor-local, but it can miss issues that are only evident relative to neighbors, such as a single faulty node during a network-wide episode.

In the NT mode, the sensor's recent history is augmented with other variables (meteorology, co-pollutants, or co-located channels). Incorporating multi-variable covariates often improves robustness to environmental confounding and cross-sensitivities [5,26,27]. However, these models require synchronized multi-sensor or multi-channel data, which may be unavailable or noisy in community and citizen science deployment.

In the ST mode, simultaneous readings from spatial neighbors are combined with the temporal context. Incorporating data from a trusted anchor (e.g., an automatic weather station or a nearby reference monitor) can substantially improve detection and correction. Kim et al. [5] reported that including anchor features reduced RMSE by approximately 17% relative to raw inputs. Beyond linear pooling, spatial machine learning (e.g., Gaussian processes and graph-based models) can explicitly model sensor-to-sensor correlations and stabilize network-level predictions at scale [15,18,19]. The main trade-offs are computational cost and scalability for mega-city networks.

To provide a structured comparison, Table 4 summarizes the three QC frameworks (HT, NT, and ST) in terms of input features, strengths, and limitations. This overview highlights the trade-offs between computational simplicity and robustness when leveraging temporal versus spatiotemporal information.

Correlation-based QC operationalizes these ideas in practice. Inter-sensor correlation is monitored over sliding windows, and alerts are raised when a node decorrelates from its peers beyond expected variability [5,14]. Such methods help distinguish network-wide episodes (all sensors spike) from device-specific faults (one sensor spikes) and have been used in both community networks and industrial regions [28,39]. Still, correlation metrics depend on sensor density and placement: sparse or irregular layouts weaken redundancy and reduce generalizability.

Classical "buddy checks" from meteorology fit naturally into this framework [8]: models learn expected relationships between nearby sites and flag violations. Probabilistic variants, such as Bayesian neural network calibration, attach uncertainty bounds to these relationships, yielding confidence-aware QC adjustments [34]. Spatiotemporal redundancy also enables value recovery: faulty-node estimates can be reconstructed from neighbors, a principle extended by federated and cluster-based frameworks that calibrate within local groups to improve scalability [30,40]. Interoperability across heterogeneous hardware and communication protocols remains an open issue.

Large urban networks often employ hierarchical designs with a few reference-grade "golden nodes" anchoring many low-cost sensors. In Breathe London, reference nodes provided citywide anchors and informed continuous adjustments to low-cost nodes [1]. Similar "virtual calibration" services leverage nearby regulatory monitors to nudge baselines and mitigate drift [3]. The effectiveness of such strategies depends on anchor availability, which is limited in many regions.

Overall, spatiotemporal QC harnesses redundancy in time (self-consistency) and space (cross-sensor consistency) to detect implausible readings and repair data streams

more robustly than single-sensor approaches [5,8]. Remaining challenges (scalability, dependence on anchors, and uneven network density) motivate lightweight, interpretable, and transferable frameworks that can adapt across diverse deployment scenarios. Section 6 discusses how these methods are operationalized in real-time systems with latency, drift, and reliability constraints.

Table 4. Comparison of spatiotemporal QC frameworks for sensor networks (based on Kim et al. [5] and related studies).

Framework	Input Features	Strengths	Limitations/Example Studies
MLQC-HT (Homogeneous Temporal)	Time series of a single variable from one sensor	Computationally lightweight; suitable for real-time anomaly detection; independent from other sensors	Misses cross-sensor/contextual anomalies; accuracy limited under dynamic conditions [8]
MLQC-NT (Nonhomogeneous Temporal)	Time series of multiple variables (e.g., meteorological + pollutant data)	Leverages multivariate correlations; more robust calibration under varying conditions	Requires synchronized datasets; performance depends on covariate selection [5,26,27]
MLQC-ST (Spatiotemporal)	Data from multiple sensor locations + temporal histories	Exploits redundancy across network; detects local anomalies; enables spatial imputation	Higher computational cost; scalability issues in dense networks; depends on availability of anchors [5,15,18,19]

6. Real-Time and Online QC Systems

Following the definitions in Section 4, we use *online QC* to denote streaming calibration, anomaly handling, and reconciliation performed on edge devices or low-latency cloud services, in contrast to offline post hoc processing. This section focuses on operational design: edge–cloud architectures, low-latency detectors and repair, automated re-calibration triggers tied to model diagnostics, and monitoring/machine learning operations (MLOps) practices that keep the spatiotemporal methods in Section 5 reliable at scale.

As air quality sensor networks move toward live data delivery, QC must also operate in (near) real time. This imposes constraints on algorithms: they should be computationally efficient, adaptive to new data, and capable of continuous operation on embedded hardware [41]. Online systems additionally face concept drift (Section 3): models need to update as sensor characteristics and environmental patterns evolve [23]. In what follows, we outline how streaming QC is implemented, including on-device processing, streaming model updates, and end-to-end architectures for live quality control.

A core requirement in real time is detecting and correcting anomalies on the fly. Many ML methods from earlier sections can be adapted to streaming [5]. For example, sliding-window regressors can be re-trained on recent data to track shifting behavior, and lightweight models (e.g., small decision trees or compact neural nets) can update incrementally. When abrupt changes occur (e.g., a baseline jump), drift detectors monitoring input statistics or prediction errors can trigger alerts or re-training. D’Elia et al. [23] studied drift mitigation for low-cost NO₂ sensors: upon drift detection, strategies such as weighted incremental updates and ensembles of old/new models extended calibration validity by weeks. Research on field deployment of PM sensors likewise reports that periodic re-training (every 2–4 weeks) is often needed to preserve accuracy under changing

meteorology [29,37]. In practice, brief field co-locations or continuous remote anchoring to references (see U.S. EPA guidelines and public platforms) are used to sustain accuracy in between maintenance cycles [20,25].

Edge computing has become central to online QC [41]. Instead of shipping raw streams to the cloud, first-tier QC runs locally on the sensor or a nearby gateway; outlier filtering, basic calibration, and sanity checks can execute at millisecond-to-second latency, flagging or correcting data before transmission. Advances in microcontrollers and single-board computers make such deployment feasible, and recent frameworks demonstrate real-time autoencoder-based calibration and anomaly screening integrated into smart-city platforms [30]. By performing cleaning at the source, edge processing also reduces bandwidth and provides resilience under intermittent connectivity.

Because edge devices are resource-constrained, many systems adopt a hybrid design. First-tier QC runs at the edge, while second-tier, computationally intensive analysis executes in the cloud, where a global view enables spatial consistency checks and cross-sensor reconciliation [14]. Architectures that summarize locally and centralize only the necessary aggregates have been proposed for resource efficiency [41], and hierarchical QC has been shown to support both indoor and outdoor monitoring at community scale with manageable overhead [16]. Federated or cluster-based learning further reduces bandwidth and can improve privacy: models are trained locally and periodically synchronized, sometimes within sensor clusters to limit communication [15,40]. In a QC context, nodes refine local anomaly detectors and share parameter updates, approaching centralized accuracy while preserving data locality. To clarify these trade-offs, Table 5 summarizes the characteristics of edge, cloud, and hybrid QC architectures, highlighting their main strengths and limitations. This structured view illustrates why many recent deployment instances favor hybrid systems that combine local responsiveness with global analytics.

Table 5. Comparison of real-time QC deployment architectures for sensor networks.

Architecture	Main Features	Strengths	Limitations/Example Studies
Edge Computing	QC performed directly on sensor nodes or local gateways (e.g., microcontrollers and Raspberry Pi)	Low latency; immediate anomaly detection; reduced bandwidth use; resilience to connectivity issues	Limited processing power and memory; unsuitable for complex deep models [30,41]
Cloud Computing	Raw or partially processed data sent to cloud servers for QC	Global network view; scalable analytics; enables spatiotemporal QC across sensors	Higher latency; dependent on reliable connectivity; privacy concerns [6,14]
Hybrid (Edge + Cloud)	First-tier QC (outlier filtering and basic calibration) on edge; second-tier QC (spatiotemporal inference and drift correction) on cloud	Balances latency and computational cost; scalable; integrates local + global patterns	Implementation complexity; requires coordination between layers [15,23]

Operationally, production systems rely on streaming pipelines to organize QC at scale. Message brokers and stream processors route sensor messages through chained operators for decoding, validation, anomaly tagging, and calibration; health metrics and drift diagnostics are logged for automated triggers and human oversight [14,30]. In industrialized

regions, coupling such pipelines with spatiotemporal models has enabled continuous calibration and monitoring under rapidly changing conditions [39].

Demonstrations from the literature illustrate these patterns. D’Elia et al. [23] describe an autonomic calibration loop that adjusted electrochemical sensor baselines in real time using reference feeds, reducing months-long drift. The U.S. EPA’s Fire and Smoke Map integrates corrected low-cost sensor data for public use, supported by Toolbox guidance that documents calibration and QC steps for streaming integration [20,25]. The HypeAIR project showed that edge screening plus cloud reconciliation can be seamlessly embedded into city platforms, stabilizing live data streams for decision support [30].

In summary, online QC systems combine algorithmic techniques with deployment engineering: sliding-window adaptation and drift detection to keep models current; edge computing for low-latency filtering and first-tier calibration; hybrid edge–cloud reconciliation for network-wide consistency; and MLOps practices for monitoring, automated re-calibration, and safe rollbacks [15,41]. The result is an end-to-end pipeline that converts raw sensor readings into quality-assured data products within seconds to minutes, enabling instant alerts that distinguish real events from sensor faults and maintaining accurate streams for public advisories [15,42]. Finally, sustained reliability benefits from drift-aware scheduling and MLOps-style automation for diagnostics and re-calibration, as detailed in Section 8.

7. Applications and Impacts

Improving the quality of data from smart air quality monitoring has broad implications for both research and society. In this section, we highlight some key applications enabled by ML-based QC and their impacts: personal exposure estimation, integration with simulations/models, and policy and management relevance.

7.1. Personal Exposure Estimation

One of the ultimate goals of dense air monitoring is to better understand human exposure to pollutants at an individual level. Low-cost sensors have made it possible to measure pollution on the scale of one’s neighborhood, daily commute, or even directly on the person (via wearables) [1,3]. However, raw data from personal monitors can be quite unreliable without QC; for example, a portable PM sensor might be influenced by how the person wears it or environmental factors like humidity in breath. By applying ML-based calibration and anomaly correction, these sensors can yield more accurate exposure data [12,24]. This has been demonstrated in studies where volunteers carry sensors and the data are later calibrated against reference instruments to quantify true exposure [4]. Some wearable devices now incorporate on-device calibration algorithms to adjust readings in real time. On a community scale, mobile sensing campaigns—like Google Street View cars instrumented with high-grade sensors—have produced high-resolution pollution maps. These maps, when combined with ML, can be interpolated and validated across space to estimate exposure on every city block [3,19]. Recent work has also explored integration with health-related data streams (e.g., heart rate or respiratory monitoring), showing the potential of quality-controlled air quality data to inform personalized health interventions [16,17]. Ultimately, reliable exposure data can feed into epidemiological studies and personal health decisions. Yet, practical adoption of personal exposure monitoring remains limited, as wearable sensors raise issues of comfort, battery life, and privacy, and calibration drift in uncontrolled personal environments is not yet fully resolved.

7.2. Real-World Deployment

Large-scale deployment of low-cost sensor networks provides concrete evidence of their value in practice. In Imperial County, California, a community-led network of more than 40 PM_{2.5} sensors has been operated by residents; with appropriate QC procedures, the resulting dataset has been shown to achieve reliability comparable to that of official monitors [16,17]. In Delhi, India, one of the earliest dense deployment instances of low-cost PM_{2.5} sensors, combined with mobile sensing nodes, demonstrated how simple QC steps such as baseline correction and spike removal enabled the high-resolution mapping of severe urban pollution [4,43]. In Europe, the Breathe London project illustrated how hybrid networks of fixed and mobile low-cost sensors can provide actionable exposure maps for citizens and policymakers, underlining their potential to guide urban planning and public health interventions [1,2].

7.3. Simulation Integration

Quality-controlled sensor data can enhance atmospheric models and simulations. Traditionally, regulatory models (like dispersion models or chemical transport models) have relied on a handful of monitoring stations for input and validation. Now, dense sensor networks provide a wealth of real-time data points that can be assimilated into models to improve their accuracy and resolution [44]. For example, ML-corrected sensor data can be fed into a data assimilation system for an urban air quality forecast model, reducing bias in the initial conditions and thus improving forecasts [44]. Another integration is using sensor data to calibrate emission simulations: a network of sensors might detect hotspot emissions (like an unknown pollution source), and ML algorithms can help back-calculate or adjust the emissions in a simulation model to better fit the observed data [15]. There is also interest in creating digital twins of urban environments for air quality—virtual models that run in parallel to the real world. These require continuous sensor inputs to stay updated; QC is paramount so that the simulation is not driven off-course by faulty data. Villarreal-Marines et al. [39] demonstrated that real-time QC and calibration significantly improved digital twin fidelity in an industrialized region, enabling the more accurate forecasting of short-term PM_{2.5} peaks. Essentially, by increasing the spatial coverage of observations, we constrain simulations better, leading to more reliable scenario analysis [44]. ML also plays a role by serving as surrogate models within simulations, for instance, replacing expensive chemistry modules with fast-trained predictors [38,45]. However, despite these advances, integration is still at a proof-of-concept stage; many studies remain computationally intensive and rarely account for uncertainties in both the ML corrections and the underlying physical models, which may limit regulatory acceptance.

7.4. Policy and Management Relevance

Perhaps the most significant impact of high-quality smart monitoring is on environmental policy and urban management. When data are plentiful and trustworthy, policymakers can make informed decisions in near real time [15]. For example, city authorities can identify neighborhoods with chronically high pollution and target interventions there, such as traffic flow changes, enforcement of industrial regulations, or installation of green infrastructure. ML-calibrated networks have been used to detect illegal pollution sources or excessive emissions by comparing patterns across the network. In one instance, a calibrated sensor network helped a city identify malfunctioning pollution control equipment at a local factory because the anomaly was clearly distinguishable from normal variation. Evidence-based policy is strengthened by having a granular view of air quality: policymakers can evaluate the impact of specific actions by observing changes picked up by the sensor network. The reliability that ML-based QC provides is crucial here; spurious

sensor readings could otherwise trigger false alarms or, conversely, mask real problems. Ravindra et al. [11] showed that after ML calibration, sensor network data became accurate enough to support health risk evaluation and source-oriented policy actions in the Indo-Gangetic Plain, a region with severe pollution challenges. Likewise, high-density sensor networks can help assess compliance with air quality standards at a micro-local scale, revealing disproportionate exposures and thereby guiding more equitable policy interventions [3,37].

Another impact area is public engagement and awareness. When people trust sensor data (because they are quality-controlled and perhaps even certified by authorities), they are more likely to use them in their daily decisions—such as choosing a less polluted route for jogging or knowing when to close windows [1]. Projects like AirNow’s Fire and Smoke map incorporate low-cost sensor data (after applying correction algorithms) to supplement official monitors during wildfires, thus providing the public with more information during crises [15]. Recent studies in community-driven monitoring projects (e.g., Imperial County, California) confirm that reliable QC enables citizens to participate more actively in environmental governance and advocacy [16,17]. This data democratization can spur community-level actions and support for clean air policies. Nonetheless, citizen-driven projects face sustainability challenges: data collection can be irregular, maintenance may depend on volunteer capacity, and trust can be undermined if correction algorithms are opaque or inconsistently applied.

To provide a structured overview, Table 6 summarizes the main application domains of ML-based QC, highlighting their benefits, limitations, and representative studies. This comparison underscores both the promise and the practical barriers to adoption across individual, scientific, and policy contexts.

Table 6. Applications of ML-based quality control in air quality monitoring: benefits, limitations, and representative studies.

Application Domain	Benefits	Limitations/Open Issues	Example Studies
Personal Exposure Estimation	Improves reliability of wearable and mobile sensing; enables fine-grained exposure data for individuals and communities	Sensor drift in uncontrolled environments; comfort, battery, and privacy challenges	[4,12,16,17]
Simulation Integration	Enhances accuracy of dispersion/chemical models via ML-calibrated inputs; supports digital twins; identifies hidden emission sources	Computationally intensive; limited treatment of uncertainties; few operational-scale implementations	[15,39,44,45]
Policy and Management Relevance	Enables near-real-time evidence-based decision making; supports regulatory compliance and environmental justice; empowers citizen engagement	Risk of misinforming policy if QC protocols are inconsistent; sustainability of citizen networks uncertain	[3,11,16,37]

In summary, ML-based QC transforms raw sensor feeds into reliable information, which in turn unlocks numerous applications: from personalized health protection to scientific research and from operational decision support to strategic policymaking for long-term air quality improvement [12,15,45]. Yet, across applications, unresolved issues of calibration stability, scalability, and transparency suggest that future work must balance technical advances with practical considerations for adoption.

8. Challenges and Future Directions

Despite considerable progress in ML-based QC for air quality monitoring, there remain numerous challenges and open research directions. Ensuring data quality in ever-expanding sensor networks is a moving target, and both technical and practical hurdles must be overcome to realize the full potential of smart air quality monitoring [15].

Data Quality and Reliability Gaps: A fundamental challenge is that many regions still have sparse monitoring coverage and inconsistent data quality standards [3]. Low-cost sensors are proliferating, but not all implementations follow best practices for calibration or maintenance, leading to highly variable data quality. Comparative studies reveal that sensor drift is one of the most persistent issues: even with initial calibration, long-term deployment suffers from gradual accuracy loss [23,37]. Developing early drift detection methods and efficient re-calibration schemes is, therefore, an active area. Another gap is pollutant coverage: much of the research has focused on PM and a few gases like O₃ or NO₂, but newer sensors for SO₂, VOCs, and ultrafine particles present unique interference patterns and require tailored ML-based QC [7,15]. Case studies such as that by Sayahi et al. [22] on Plantower PMS sensors emphasize how environmental factors like humidity can undermine long-term stability, further motivating robust adaptive methods. Comparative evaluations indicate that QC performance is highly sensitive to predictor selection (e.g., meteorology and co-pollutants), the length of the calibration window, and the algorithm family; the lack of standardized protocols complicates cross-study comparisons [7,26]. Moreover, multi-year, publicly accessible benchmarks with agreed train/validation/test splits and rich metadata remain scarce, limiting reproducibility and the rigorous propagation of QC uncertainties into downstream analyses [7,40]. While these challenges are widely recognized, relatively few studies provide long-term validation data, meaning many proposed solutions remain promising in short trials but untested at scale.

Scalability and System Integration: As networks scale to hundreds or thousands of nodes, scalability of QC algorithms is vital. Techniques that work well for tens of sensors may face bottlenecks at city-wide scale. Graph-based and clustering approaches have been suggested to partition networks for tractable computation [27,40]. System integration challenges include reliable communication, power management, and model updates. Many sensor networks operate on limited power; running complex ML on-device could strain energy resources [41]. As discussed in Section 6, hybrid edge–cloud processing remains a key strategy for balancing local responsiveness with global analytics [14,30]. Operationalizing this at scale will likely require drift-aware scheduling and automated re-calibration triggers tied to model diagnostics, integrated within MLOps-style pipelines for sensor networks [23,40]. Modular frameworks that combine AI, edge computing, and multimodal data (traffic, meteorology, and satellite data) are emerging as a promising paradigm [15,42]. However, large-scale demonstrations are still rare, and questions remain about interoperability across heterogeneous sensor hardware and data platforms.

Adaptability and Transferability: ML models trained in one city or season often do not generalize well to other contexts due to differences in sources, climate, or sensor batches. Several field studies have reported that models calibrated under one set of conditions performed poorly when directly applied elsewhere, underscoring the limits of transferability [44]. Developing transferable calibration models or applying transfer learning could reduce the need to restart training for each deployment. Domain adaptation and federated learning approaches (see Section 6) show promise in this regard [15,42]. Semi-supervised learning is also key, since labeled “ground truth” data are scarce. Approaches like unsupervised anomaly detection, simulation-based anomaly synthesis, or physics-informed ML can reduce dependence on expensive labeled data [5,34]. Few-shot co-location protocols and transfer learning from canonical sites can further reduce the labeled-data

burden while preserving site-specific biases [44]. Explainability remains another challenge: policymakers and scientists may hesitate to trust black-box corrections. Methods like Bayesian neural networks or rule extraction from ensembles provide a way to attach interpretable confidence intervals to predictions [32,34]. Nevertheless, balancing accuracy and interpretability remains unresolved; highly interpretable models may underperform in complex environments, while state-of-the-art black-box models often struggle to gain policy acceptance.

Maintenance and Longevity: In practice, sensors require periodic cleaning, replacement, and re-calibration. QC algorithms could play a predictive role here—for example, identifying gradual baseline shifts as early indicators of sensor aging [23]. Long-term field studies, such as Connolly et al.'s [16] and Villarreal-Marines et al.'s [39], demonstrate the importance of sustained monitoring partnerships to collect multi-year datasets. Such datasets are invaluable for developing next-generation QC models that explicitly account for seasonal cycles, material degradation, and long-term drift. However, these collaborative datasets are still geographically limited, raising concerns about the global representativeness of current QC strategies.

Interdisciplinary Integration: The future of smart monitoring likely involves integration with health data, traffic flows, and citizen engagement platforms. For instance, coupling quality-controlled exposure data with GPS and biometric data (heart rate, respiratory signals, etc.) could support personalized health interventions [3,17]. Ultra-reliable QC is required for critical applications such as issuing public health alerts or managing pollution-sensitive infrastructure [15]. Hybrid deployment (indoor + outdoor) also raises new challenges for consistency across heterogeneous environments [16]. Despite this promise, privacy concerns and governance of cross-domain data remain underexplored, limiting the near-term feasibility of such integrated systems.

Policy and Standardization: On the regulatory side, agencies such as the U.S. EPA and EU are exploring how calibrated sensor data could complement reference monitoring [3]. Real-world policy integration, however, still requires standardized QC protocols and certification benchmarks. Clear thresholds for uncertainty must be defined before regulatory agencies can systematically adopt ML-corrected sensor data. Future frameworks are likely to embed uncertainty quantification. In practice, this means reporting not just corrected values but also confidence intervals [15,33]. For example, probabilistic calibration using Bayesian neural networks or ensemble-based predictive intervals can provide decision-ready uncertainty bounds [7,34]. For instance, the U.S. EPA has piloted the integration of corrected low-cost sensor data into public platforms such as the AirNow Fire and Smoke Map, supported by the Air Sensor Toolbox, which provides calibration and QC guidelines [20,25]. In Europe, the CEN technical committees are drafting performance evaluation standards for low-cost air quality sensors (e.g., CEN/TS 17660-1) to ensure comparability across devices and member states [21]. Similarly, the WMO and UNEP have acknowledged the supplementary role of QC-enhanced LCS networks in expanding monitoring coverage in regions lacking reference stations [46,47]. These ongoing initiatives demonstrate that technical advances in ML-based QC are beginning to converge with institutional and regulatory efforts, laying the groundwork for globally recognized practices. A recent review emphasized that standardization will be critical to scaling community networks into regulatory decision-making processes [38]. Yet, without consensus on benchmarks, there is a risk of fragmented standards across regions, which could undermine global comparability.

Addressing these gaps will hinge on shared benchmarks, standardized QC protocols, and field-scale validations that explicitly account for drift and uncertainty.

To provide a consolidated view, Table 7 summarizes the major challenges in ML-based QC, representative current approaches, and the key limitations that remain unresolved. This

overview highlights both the technical and institutional barriers that must be addressed for widespread adoption.

Table 7. Key challenges in ML-based quality control for air quality monitoring, current approaches, and limitations.

Challenges	Current Approaches	Limitations/Open Issues
Data Quality and Reliability	Initial calibration and co-location with reference instruments; drift detection methods; anomaly detection	Long-term drift remains difficult to detect early; limited validation datasets; emerging pollutants (e.g., VOCs and ultrafine particles) lack robust QC frameworks
Scalability and Integration	Graph-based QC, clustering, and modular edge–cloud architectures [14,30]	Computational bottlenecks at city-wide scale; interoperability challenges across heterogeneous hardware and data platforms
Adaptability and Transferability	Transfer learning, domain adaptation, and federated learning [15,42]; semi-supervised anomaly detection	Limited cross-city generalization; dependence on scarce labeled data; black-box models hinder trust and regulatory uptake
Maintenance and Longevity	Predictive QC to detect baseline shifts; long-term field campaigns [16,39]	Multi-year datasets are geographically limited; global representativeness and standardized maintenance protocols are lacking
Interdisciplinary Integration	Linking quality-controlled air quality data with health, traffic, and citizen science platforms [17]	Privacy concerns; governance frameworks underdeveloped; indoor–outdoor consistency remains unresolved
Policy and Standardization	Exploratory standards by EPA/EU; uncertainty quantification (e.g., Bayesian methods) [34]	No global consensus; risk of fragmented regional standards; unclear thresholds for regulatory acceptance

In summary, future work must focus on making ML-based QC more automated, scalable, and robust: handling data deluge with cloud–edge hybrids, maintaining accuracy despite drift via online learning, and ensuring that improved data quality leads directly to better decisions. Recent reviews highlighted data quality, scalability, and integration as the most pressing research directions [15,38]. By fostering interdisciplinary collaboration and establishing standards, the next decade should see QC techniques mature into a standard practice for proactive air quality management. However, the success of this vision will depend not only on technical advances but also on trust building, governance, and long-term sustainability of sensor networks.

9. Conclusions

Machine learning-based quality control (ML-based QC) has become a cornerstone of smart air quality monitoring over the past decade. A wide range of techniques—from regression and decision trees to deep neural networks—have been applied to calibrate low-cost sensors, detect anomalies, and leverage spatiotemporal correlations. These approaches have substantially improved the reliability of low-cost sensor data, in many cases approaching the accuracy of reference instruments. Real-time applications, supported by edge computing and online learning, are already enabling more responsive environmental management and more accurate assessments of personal exposure.

Despite these advances, several challenges remain. Generalization across contexts, interpretability of models, scalability to large networks, and the lack of standardized benchmarks continue to limit widespread adoption. Addressing these issues is essential to regulatory acceptance and the long-term sustainability of ML-based QC.

Looking ahead, QC frameworks must become more robust, scalable, and autonomous. Future networks should be able to self-calibrate continuously, detect sensor faults early, and integrate multimodal data sources such as traffic, meteorology, and health. Equally important is the development of explainable and standardized QC procedures so that scientists, policymakers, and regulators can trust ML-driven corrections. Establishing benchmarks and certification schemes will be critical to institutional uptake.

In summary, ML-based QC has shown strong potential but remains constrained by data quality, transferability, and governance gaps. By uniting technical innovation with institutional trust, it can evolve from a promising research field into a reliable foundation for environmental monitoring. Recent initiatives, such as the U.S. EPA AirNow Fire and Smoke Map and the European CEN/TS 17660-1 standard, already demonstrate how ML-corrected sensor data can inform real-world policy. This convergence of IoT sensing, AI-driven QC, and regulatory adoption signals a gradual but significant transformation in the way air quality is monitored and managed.

Author Contributions: Conceptualization, Y.-H.K. and S.-H.M.; validation, S.-H.M.; formal analysis, S.-H.M.; investigation, Y.-H.K.; resources, Y.-H.K.; writing—original draft preparation, Y.-H.K.; writing—review and editing, S.-H.M.; supervision, Y.-H.K.; project administration, Y.-H.K.; funding acquisition, Y.-H.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the project titled “Marine Oil Spill Risk Assessment and Development of Response Support System through Big Data Analysis” of the Korea Institute of Marine Science & Technology Promotion (KIMST), funded by the Korea Coast Guard (grant number KCG-01-2017-05; 20190439; RS-2019-KS191280).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The work reported in this paper was conducted during the sabbatical year of Kwangwoon University in 2025. During the preparation of this manuscript, the authors used ChatGPT (GPT-5, OpenAI, August 2025) for the purposes of language refinement, formatting assistance, and conversion of references into LaTeX/BibTeX style. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AE	Autoencoder
AI	Artificial Intelligence
ANN	Artificial neural network
AWS	Automated Weather Station
CNN	Convolutional neural network
DL	Deep learning
DT	Decision tree
EPA	Environmental Protection Agency
FEM	Federal Equivalent Method
FRM	Federal Reference Method
GAN	Generative Adversarial Network
GB	Gradient boosting
GNN	Graph neural network

GPS	Global Positioning System
IoT	Internet of Things
k NN	k -nearest neighbor
LCS	Low-cost sensor
LOD	Limit of detection
LSTM	Long short-term memory
ML	Machine learning
MLOps	Machine learning operations
MLP	Multi-Layer Perceptron
MLQC	Machine learning-based quality control
MEC	Multi-access Edge Computing
MOS	Metal-oxide sensor
O&M	Operations and maintenance
PARAAD	Pattern-based Attention Recurrent Autoencoder
PCA	Principal Component Analysis
PM	Particulate Matter (e.g., PM ₁ , PM _{2.5} , PM ₁₀)
QA	Quality assurance
QC	Quality control
r	Pearson correlation coefficient
R^2	Coefficient of determination
RF	Random forest
RMSE	Root mean squared error
RNN	Recurrent neural network
SVR	Support vector regression
SVM	Support vector machine
T/RH	Temperature/relative humidity
VAE	Variational autoencoder
VOC(s)	Volatile Organic Compound(s)

References

- Castell, N.; Dauge, F.R.; Schneider, P.; Vogt, M.; Lerner, U.; Fishbain, B.; Broday, D.; Bartonova, A. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? *Environ. Int.* **2017**, *99*, 293–302. [\[CrossRef\]](#)
- Lewis, A.; Edwards, P. Validate personal air-pollution sensors. *Nature* **2016**, *535*, 29–31. [\[CrossRef\]](#)
- Morawska, L.; Thai, P.K.; Liu, X.; Asumadu-Sakyi, A.; Ayoko, G.; Bartonova, A.; Bedini, A.; Chai, F.; Christensen, B.; Dunbabin, M.; et al. Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone? *Environ. Int.* **2018**, *116*, 286–299. [\[CrossRef\]](#) [\[PubMed\]](#)
- Hasenfratz, D.; Saukh, O.; Walser, C.; Hueglin, C.; Fierz, M.; Arn, T.; Beutel, J.; Thiele, L. Deriving high-resolution urban air pollution maps using mobile sensor nodes. *Pervasive Mob. Comput.* **2015**, *16*, 268–285. [\[CrossRef\]](#)
- Kim, H.J.; Park, S.M.; Choi, B.J.; Moon, S.H.; Kim, Y.H. Spatiotemporal approaches for quality control and error correction of atmospheric data through machine learning. *Comput. Intell. Neurosci.* **2020**, *2020*, 7980434. [\[CrossRef\]](#)
- Zimmerman, N.; Presto, A.A.; Kumar, S.P.N.; Gu, J.; Hauryliuk, A.; Robinson, E.S.; Robinson, A.L.; Subramanian, R. A machine learning calibration model for low-cost gas sensors. *Atmos. Meas. Tech.* **2018**, *11*, 291–313. [\[CrossRef\]](#)
- Giordano, M.R.; Malings, C.; Pandis, S.N.; Presto, A.A.; McNeill, V.F.; Westervelt, D.M.; Beekmann, M.; Subramanian, R. From low-cost sensors to high-quality data: A summary of challenges and best practices for effectively calibrating low-cost particulate matter mass sensors. *J. Aerosol Sci.* **2021**, *158*, 105833. [\[CrossRef\]](#)
- Kim, H.J.; Lee, H.S.; Choi, B.J.; Kim, Y.H. Machine learning-based quality control and error correction using homogeneous temporal data collected by IoT sensors. *J. Korea Conver. Soc.* **2019**, *10*, 17–23.
- Malings, C.; Tanzer, R.; Hauryliuk, A.; Kumar, S.P.N.; Zimmerman, N.; Kara, L.B.; Presto, A.A.; Subramanian, R. Development of a general calibration model and long-term performance evaluation of low-cost sensors for air pollutant gas monitoring. *Atmos. Meas. Tech.* **2019**, *12*, 903–920. [\[CrossRef\]](#)
- Kim, Y.H.; Ha, J.H.; Yoon, Y.; Kim, N.Y.; Im, H.H.; Sim, S.; Choi, R.K.Y. Improved correction of atmospheric pressure data obtained by smartphones through machine learning. *Comput. Intell. Neurosci.* **2016**, *2016*, 9467878. [\[CrossRef\]](#)
- Ravindra, K.; Kumar, S.; Kumar, A.; Mor, S. Enhancing accuracy of air quality sensors with machine learning to augment large-scale monitoring networks. *npj Clim. Atmos. Sci.* **2024**, *7*, 326. [\[CrossRef\]](#)

12. Yaqoob, I.; Kumar, V.; Chaudhry, S.A. Machine learning calibration of low-cost sensor PM_{2.5} data. In Proceedings of the IEEE International Symposium on Systems Engineering (ISSE), Perugia, Italy, 16–19 October 2024; pp. 1–8. [\[CrossRef\]](#)
13. Allka, X.; Ferrer-Cid, P.; Barcelo-Ordinas, J.M.; Garcia-Vidal, J. Pattern-based attention recurrent autoencoder for anomaly detection in air quality sensor networks. *IEEE Trans. Netw. Sci. Eng.* **2024**, *11*, 6372–6381. [\[CrossRef\]](#)
14. Rollo, F.; Bachechi, C.; Po, L. Anomaly detection and repairing for improving air quality monitoring. *Sensors* **2023**, *23*, 640. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Garcia, A.; Saez, Y.; Harris, I.; Huang, X.; Collado, E. Advancements in air quality monitoring: A systematic review of IoT-based air quality monitoring and AI technologies. *Artif. Intell. Rev.* **2025**, *58*, 275. [\[CrossRef\]](#)
16. Connolly, R.E.; Yu, Q.; Wang, Z.; Chen, Y.H.; Liu, J.Z.; Collier-Oxandale, A.; Papapostolou, V.; Polidori, A.; Zhu, Y. Long-term evaluation of a low-cost air sensor network for monitoring indoor and outdoor air quality at the community scale. *Sci. Total Environ.* **2022**, *807*, 150797. [\[CrossRef\]](#)
17. English, P.; Amato, H.; Bejarano, E.; Carvlin, G.; Lugo, H.; Jerrett, M.; King, G.; Madrigal, D.; Meltzer, D.; Northcross, A.; et al. Performance of a low-cost sensor community air monitoring network in Imperial County, CA. *Sensors* **2020**, *20*, 3031. [\[CrossRef\]](#)
18. Iyer, S.; Balashankar, A.; Aeberhard, W.; Bhattacharyya, S.; Rusconi, G.; Jose, L.; Soans, N.; Sudarshan, A.; Pande, R.; Subramanian, L. Modeling fine-grained spatio-temporal pollution maps with low-cost sensors. *npj Clim. Atmos. Sci.* **2022**, *5*, 76. [\[CrossRef\]](#)
19. Blaga, R.; Gautam, S. Improving PM10 sensor accuracy in urban areas through calibration in Timișoara. *npj Clim. Atmos. Sci.* **2024**, *7*, 268. [\[CrossRef\]](#)
20. US Environmental Protection Agency. Air Sensor Toolbox for Citizen Scientists, Researchers and Developers. 2023. Available online: <https://www.epa.gov/air-sensor-toolbox> (accessed on 28 August 2025).
21. CEN/TS 17660-1:2021; Air Quality—Performance Evaluation of Air Quality Sensors—Part 1: Gaseous Pollutants in Ambient Air. European Committee for Standardization: Brussels, Belgium, 2021. .
22. Sayahi, T.; Butterfield, A.; Kelly, K.E. Long-term field evaluation of the Plantower PMS low-cost particulate matter sensors. *Environ. Pollut.* **2019**, *245*, 932–940. [\[CrossRef\]](#)
23. D'Elia, G.; Ferro, M.; Sommella, P.; Ferlito, S.; De Vito, S.; Di Francia, G. Concept drift mitigation in low-cost air quality monitoring networks. *Sensors* **2024**, *24*, 2786. [\[CrossRef\]](#)
24. Wang, Y.; Du, Y.; Wang, J.; Li, T. Calibration of a low-cost PM_{2.5} monitor using a random forest model. *Environ. Int.* **2019**, *133*, 105161. [\[CrossRef\]](#) [\[PubMed\]](#)
25. US Environmental Protection Agency. Technical Approaches for Sensor Data in the AirNow Fire and Smoke Map. 2023. Available online: <https://www.epa.gov/air-sensor-toolbox/technical-approaches-sensor-data-airnow-fire-and-smoke-map> (accessed on 28 August 2025).
26. Liang, L.; Daniels, J. What influences low-cost sensor data calibration?—A systematic assessment of algorithms, duration, and predictor selection. *Aerosol Air Qual. Res.* **2022**, *22*, 220076. [\[CrossRef\]](#)
27. Koziel, S.; Pietrenko-Dabrowska, A.; Wojcikowski, M.; Pankiewicz, B. Statistical data pre-processing and time series incorporation for high-efficacy calibration of low-cost NO₂ sensor using machine learning. *Sci. Rep.* **2024**, *14*, 9152. [\[CrossRef\]](#)
28. Won, W.S.; Noh, J.; Oh, R.; Lee, W.; Lee, J.-W.; Su, P.-C.; Yoon, Y.-J. Enhancing the reliability of particulate matter sensing by multivariate Tobit model using weather and air quality data. *Sci. Rep.* **2023**, *13*, 13150. [\[CrossRef\]](#)
29. Taştan, M. Machine learning-based calibration and performance evaluation of low-cost Internet of Things air quality sensors. *Sensors* **2025**, *25*, 3183. [\[CrossRef\]](#)
30. Bachechi, C.; Rollo, F.; Po, L. HypeAIR: A novel framework for real-time low-cost sensor calibration for air quality monitoring in smart cities. *Ecol. Informatics* **2024**, *81*, 102568. [\[CrossRef\]](#)
31. Osman, E.; Banerjee, C.; Poonia, A.; Sankarapu, P. Detecting environmental anomalies: Variational autoencoder-based analysis of air quality time series data. *Int. J. Intell. Syst. Appl. Eng.* **2024**, *12*, 3687–3694. .
32. Koziel, S.; Pietrenko-Dabrowska, A.; Wojcikowski, M.; Pankiewicz, B. Efficient field correction of low-cost particulate matter sensors using machine learning, mixed multiplicative/additive scaling and extended calibration inputs. *Sci. Rep.* **2025**, *15*, 18573. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Koziel, S.; Pietrenko-Dabrowska, A.; Wojcikowski, M.; Pankiewicz, B. High-performance machine-learning-based calibration of low-cost nitrogen dioxide sensor using environmental parameter differentials and global data scaling. *Sci. Rep.* **2024**, *14*, 26120. [\[CrossRef\]](#)
34. Taira, G.; Leal, A.; Santos, A.; Park, S.W. Bayesian neural network-based calibration for urban air quality sensors. In *32nd European Symposium on Computer Aided Process Engineering*, Toulouse, France, 12–15 June 2022; Montastruc, L., Negny, S., Eds.; Elsevier: Amsterdam, The Netherlands, 2022; Volume 51, pp. 1549–1554. [\[CrossRef\]](#)
35. Chu, H.J.; Ali, M.Z.; He, Y.C. Spatial calibration and PM_{2.5} mapping of low-cost air quality sensors. *Sci. Rep.* **2020**, *10*, 22079. [\[CrossRef\]](#)
36. Lee, M.K.; Moon, S.H.; Yoon, Y.; Kim, Y.H.; Moon, B.R. Detecting anomalies in meteorological data using support vector regression. *Adv. Meteorol.* **2018**, *2018*, 5439256. [\[CrossRef\]](#)

37. Sousan, S.; Wu, R.; Popoviciu, C.; Fresquez, S.; Park, Y.M. Advancing low-cost air quality monitor calibration with machine learning methods. *Environ. Pollut.* **2025**, 374, 126191. [[CrossRef](#)] [[PubMed](#)]
38. Chadalavada, S.; Faust, O.; Salvi, M.; Seoni, S.; Raj, N.; Raghavendra, U.; Gudigar, A.; Barua, P.D.; Molinari, F.; Acharya, R. Application of artificial intelligence in air pollution monitoring and forecasting: A systematic review. *Environ. Model. Softw.* **2025**, 185, 106312. [[CrossRef](#)]
39. Villarreal-Marines, M.; Pérez-Rodríguez, M.; Mancilla, Y.; Ortiz, G.; Mendoza, A. Field calibration of fine particulate matter low-cost sensors in a highly industrialized semi-arid conurbation. *npj Clim. Atmos. Sci.* **2024**, 7, 293. [[CrossRef](#)]
40. Narayana, M.V.; Jaliha, D.; Nagendra, S.M.S. Establishing a sustainable low-cost air quality monitoring setup: A survey of the state-of-the-art. *Sensors* **2022**, 22, 394. [[CrossRef](#)]
41. Idrees, Z.; Zou, Z.; Zheng, L. Edge computing based IoT architecture for low cost air pollution monitoring systems: A comprehensive system analysis, design considerations, and development. *Sensors* **2018**, 18, 3021. [[CrossRef](#)]
42. Fasano, G.; Deldjoo, Y.; di Noia, T.; Lau, B.; Adham-Khiabani, S.; Morris, E.; O'Faolain, L. Use of air quality sensor network data for real-time pollution-aware POI suggestion. In Proceedings of the ACM on Web Conference Companion, Sydney, Australia, 28 April–2 May 2025; pp. 2827–2830. [[CrossRef](#)]
43. Kumar, A.; Chaudhuri, S. Improving Urban Air Quality Monitoring in Delhi, India: Reflections on Low-Cost Air Quality Sensors (LCAQS) and Participatory Engagement. *Environ. Urban. ASIA* **2022**, 13, 265–283. [[CrossRef](#)]
44. Kagainalkar, A.; Kumar, S.; Gargava, P.; Niyogi, D. Review of urban computing in air quality management as smart city service: An integrated IoT, AI, and cloud technology perspective. *Urban Clim.* **2021**, 39, 100972. [[CrossRef](#)]
45. Fan, S.; Hao, D.; Feng, Y.; Xia, K.; Yang, W. A hybrid model for air quality prediction based on data decomposition. *Information* **2021**, 12, 210. [[CrossRef](#)]
46. World Meteorological Organization. Low-cost Sensors Can Improve Air Quality Monitoring and People's Health. 2022. Available online: <https://wmo.int/media/news/low-cost-sensors-can-improve-air-quality-monitoring-and-peoples-health> (accessed on 28 August 2025).
47. United Nations Environment Programme. Air Quality and Low-Cost Sensors: UNEP Initiatives. 2022. Available online: <https://www.unep.org/> (accessed on 28 August 2025).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.