Proceedings of the 6th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/71/20241635

Evaluating the efficacy of machine learning in calibrating low-cost sensors

Zhihan Wang

Concordia University, Sir George Williams Campus, 1455 De Maisonneuve Blvd. W. Montreal, QC, H3G 1M8, Canada

wzhihan420@outlook.com

Abstract. Ambient air quality monitoring requires low-cost environmental sensor devices that are affordable and feasible for large-scale implementation. However, issues such as sensor drift, environmental sensitivity, and inter-sensor variability affect data accuracy and cannot be adequately addressed by traditional calibration methods. This paper summarizes the use of machine learning techniques for calibrating low-cost sensors. The literature review shows that machine learning models like Random Forest, Support Vector Regression, and Neural Networks significantly improve sensor accuracy and reliability. For instance, Random Forest models reduced the root mean squared error by 30% for $PM_{2.5}$ measurements, while Neural Networks achieved an R^2 value of 0.997 for methane sensors. Integrating machine learning with IoT and mobile technologies enhances real-time monitoring and spatial resolution. Identified gaps include the quality of training datasets, managing environmental variability, and improving model transferability across different contexts. Addressing these gaps through advanced models and real-time calibration methodologies will further enhance sensor performance, ensuring more precise and reliable environmental data.

Keywords: Machine learning, sensor calibration, low-cost sensors, air quality monitoring, IoT.

1. Introduction

Affordable, easy deployment of low-cost sensors has led to revolutionary large-coverage continuous air quality monitoring. However, these sensors suffer from large-scale limitations, such as sensor drift, environmental sensitivity, and inter-sensor variability, which compromise accuracy in data; calibration should hence be performed for reliable data collection of these sensors.

While useful, the traditional calibration methods do not sufficiently address the many and varied challenges that low-cost sensor imaging systems present. Current methods have difficulty taking into account the dynamic and complex nature of factors influencing sensor performance, such as different environmental conditions or slowly drifting outputs due to aging.

Machine learning over the last several years has proven to be a strong instrument for placing such low-cost sensors into proper calibration. As a result, machine learning will find biases or inaccuracies that can be corrected in large datasets and complex patterns traditional methods miss. Previous research has evaluated the improvements in sensor performance that machine learning models can provide; For example, Kumar and Sahu obtained substantial decreases in root mean squared error (RMSE), as well

DOI: 10.54254/2755-2721/71/20241635

as increased values of R^2 for $PM_{2.5}$ Random Forest models with measurements [1]. Similarly, Mitchell et al obtained an R^2 of 0.997 for sensors, modeled as methane with neural network calibration [2].

This review covers literature related to calibrating low-cost sensor systems using machine learning. The paper aims to introduce current low-cost sensor technology, present challenges in calibration of them, and give an overview of different machine-learning techniques. Analysed performance improvements reported by literature are also discussed. The goal is to provide a comprehensive understanding of how machine learning can enhance the reliability of low-cost environmental monitoring systems and identify gaps for future research.

2. Literature review

2.1. Overview of low-cost sensor systems

Low-cost sensors with user-friendliness and affordability make them useful for the increased popularity of low-cost sensors in air quality monitoring, thus possibly enabling large-scale, continuous monitoring projects. The sensors include electrochemical, metal oxide semiconductor (MOS) sensors, and optical particle counters for the measurement of NO_2, CO, O_3, $PM_{2.5}$, and PM_{10}. Electrochemical sensors are sensitive and specific for gases such as NO_2, CO, and O_3. The MOS sensors are sensitive to gases, which are detected by the change in electrical resistance. They are mainly used in the detection of CO and VOC, but there is high environmental sensitivity calling for frequent calibration. Optical particle counters determine the concentrations of particulate matter by detecting light-scattered particles, hence enabling fast and real-time measurements [2,3].

Nonetheless, low-cost sensors have problems associated with sensor drift, environmental sensitivity, and inter-sensor variability that debase their accuracy in data collection. Aging, wear, and environmental conditions are some of the factors that create deviations from initial calibrations of low-cost sensors; thus, they require regular recalibrations to maintain their accuracy [2,4]. Traditional calibration methods remain, however, insufficient for all these problems holistically, which has thus raised a growing interest in advanced techniques like machine learning to enhance sensor performance and reliability [1,2].

Table 1 summarizes the low-cost sensor system project in this literature review.

Table 1. Summary of projects using machine learning for sensor calibration.

Project Name	Reference	Pollutant Detected	Sensor Used	Models Tested
Evaluation of Nine ML Algorithms for $PM_{2.5}$ Sensors	Kumar & Sahu, 2021 [1]	$PM_{2.5}$	Custom $PM_{2.5}$ Sensor	MLR, SVR, kNN, RT, RF, GB
Field Calibration Method for Low-Cost Sensors	Patra et al., 2021 [5]	$PM_{2.5}$, NO_2, O_3	Alphasense OPC-N3, NO_2-A43F	OLS, Elastic Net, RF, GAM
Calibration of a Low-Cost Methane Sensor	Mitchell et al., 2024 [2]	Methane	Figaro NGM2611-E13	Various ML models
Improving Data Reliability for Indoor PM Monitoring	Chojer et al., 2022 [6]	$PM_{2.5}$	Plantower PMS5003	Linear Regression, kNN, SVR
Humidity and Temperature Corrections for Low-Cost Sensors	Vajs et al., 2021 [7]	$PM_{2.5}$	Alphasense OPC-N2	Linear Regression, RF, ANN
Indoor Air Quality Monitoring and Source Apportionment	Higgins et al., 2024 [8]	$PM_{2.5}$, CO_2	Senseair S8, Plantower PMS7003	MLR, SVR, kNN, RF

DOI: 10.54254/2755-2721/71/20241635

Table 1. (continued).

Advancing Air Sensor Calibration in Stationary and Mobile Settings	Wang et al., 2022 [9]	$PM_{2.5}$, NO_2	Alphasense OPC-N3, NO_2-A43F	Linear Regression, RF, LightGBM
IoT LoRaWAN Connectivity and ML-Based Calibration	Ali et al., 2020 [10]	$PM_{2.5}$, NO_2, O_3	Custom IoT LoRaWAN Sensor	MLR, SVR, kNN, RF
Urban Air Quality Mapping with Mobile Sampling in Seoul	Lim et al., 2019 [11]	$PM_{2.5}$, NO_2, O_3	Alphasense OPC-N3, NO_2-A43F	Linear Regression, RF, LightGBM, SVR
ML Techniques to Improve Field Performance of Low-Cost Sensors	Bush et al., 2022 [12]	$PM_{2.5}$, NO_2, O_3, CO_2	Plantower PMS7003, Senseair S8, Alphasense OPC-N3	MLR, kNN, RF, SVR, ANN

2.2. Machine learning techniques

Various machine learning algorithms that have been used to enhance calibration for low-cost sensors include supervised, unsupervised, and reinforcement learning.

Supervised Learning: This involves training models on labelled datasets, in which the model learns how to map inputs into known outputs through the data. The common algorithm is linear regression, which predicts continuous output variables and is widely used since most sensor readings have linearity; support vector machines (SVR) performed better on calibration tasks because of the ability to balance between model complexity and accuracy; k-nearest neighbors (kNN), where the approach is simple for small datasets; random forest, where it constructs multiple decision trees and proves to be very robust for complex, non-linear relationships; and neural networks, where each node is interconnected with one another across layers and is said to learn complex patterns from large datasets [1].

Unsupervised Learning: Models are trained on data that does not contain responses labeled by humans; it aims to find structure among the data's natural groupings. The K-means Clustering partitions data into clusters nearest to the mean; it is effective when it comes to finding hidden patterns. Principal Component Analysis (PCA) reduces the dimensionality of the dataset while retaining the variance and outlining the principal variables affecting sensor readings [2].

Reinforcement Learning: This category considers the learning algorithm as an agent's decision according to the performance of the actions and the rewards observed. It finds applications in dynamic calibration scenarios, mainly because the environments keep changing. However, it is not frequently applied to sensor calibration, like supervised and unsupervised learning are.

2.3. Application of machine learning in calibration

Machine learning has seen widespread application to improve calibration with low-cost sensors for a range of pollutants, showing significant gains both in accuracy and reliability.

One study found that using Random Forest models for $PM_{2.5}$ sensors, such as the Alphasense OPC-N2, resulted in good accuracy improvement, with an R^2 value of 0.85 and a reduction in RMSE of 30% compared with traditional methods [1]. Support vector regression and neural networks have also been effective, with neural networks performing the best [1].

Some very promising work has been done for NO_2 sensors using machine learning models—for example, Random Forests, Neural Networks, and SVR. Neural networks improved RMSE by about 30% and increased accuracy in measurements to 25% [1]. Models including environmental variables, such as Random Forests and Gradient Boosting, have improved calibration accuracy, among others [2].

For other pollutants studies, calibration of sensors for other pollutants such as methane has been made possible by machine learning methods; methane sensors, when calibrated using random forest and neural network methods in one study, delivered high accuracy at RMSE = 5.1 ppm and R^2 = 0.997 [2]. On the other hand, some studies have demonstrated that the calibration of multiple pollutant-sensing instruments including PM_{10}, NO_2, O_3, CO, and SO_2 could be greatly enhanced by the use of linear regression, random forest, and gradient boosting models among others [3,12,13].

Indoor air quality monitoring, based on low-cost sensor data, has greatly improved because of machine learning in improving reliability. Recent studies show that the employment of Alphasense OPC-N3 and NO_2-A43F sensors in indoor environments proves to improve data accuracy considerably; hereby, random forest and SVR models are addressed [5].

This has been furthered to include sensor calibration within mobile sampling and IoT integration by machine learning. Machine learning models, such as the random forest model used within mobile sampling, enhanced urban maps of air quality in Seoul, South Korea, contributing to a significant increase in spatial resolution and precision [11]. Further, the machine learning-calibrated technique, combined with IoT connectivity, also demonstrated greater accuracy, and monitoring real-time capacity compared to techniques without these capabilities [1,2,3,5].

2.4. General trends in machine learning applications

Several different machine-learning algorithms have been recently applied to improve the calibration of low-cost sensors. The most widely used models include Random Forest, Support Vector Regression, Neural Networks, and Gradient Boosting. These models address the intrinsic limitations of low-cost sensors such as sensor drift, environmental sensitivity, and inter-sensor variability. The model significantly increased the accuracy, with the sensor of $PM_{2.5}$ being between high R^2 and low RMSE in its calibration as compared to the traditional method. The use of Gradient Boosting and Neural Networks gave important improvements in the accuracy of NO_2 and O_3 sensors. The nearest to 0.998 R^2 for methane sensors showed high calibration accuracy. Random forest models more often showed the highest performance for most of the pollutants since they can handle non-linear relationships with good stability in the results. Particularly, for large datasets with complex patterns, Neural Networks can show a strong increase in calibration accuracy. SVR and Gradient Boosting models are also effective, often generating noticeable decreases in RMSE. Machine learning integrated with IoT, and mobile sampling technologies has enhanced monitoring capabilities and increased spatial resolution, both in a real-time situation, for an accurate approach with detailed environmental data collection.

2.5. Advantages and limitations

Machine learning models largely improve the calibration accuracy of low-cost sensors, providing much-needed reliable data. Random Forest and Neural Network models are good enough to capture the complex and non-linear relationships existing between the sensor readings and true pollutant concentrations. It is also scalable in regard to massive datasets and hence fits extensive sensor networks.

Nevertheless, the machine learning models are greatly dependent on the quality and quantity of the data. For more sophisticated models like Neural Networks, excessive computational resources, and vast expertise are required, which limits their more general application. Apart from the limitations, data inconsistency may arise with changing environmental conditions that influence sensor readings and hence may require frequent recalibration.

3. Methodology

This section describes the approach used to select and analyse literature on the application of machine learning in calibrating low-cost sensors. The focus was on studies discussing machine learning techniques and their application in calibrating low-cost environmental sensors.

DOI: 10.54254/2755-2721/71/20241635

3.1. Search strategy and databases used

A comprehensive literature search was conducted using databases including IEEE Xplore, PubMed, and Google Scholar. The search was limited to studies published between 2010 and 2023 to ensure the inclusion of the most recent advancements. Keywords used in the search included "machine learning calibration low-cost sensors," "low-cost air quality sensor machine learning," and "machine learning air quality monitoring." Additional filters applied included peer-reviewed articles and studies focusing on empirical data.

3.2. Inclusion and exclusion criteria

Papers were included if they discussed machine learning techniques specifically applied to low-cost sensor calibration, provided empirical data with quantitative results demonstrating the effectiveness of the calibration methods, and were published in peer-reviewed journals or conferences. Studies were excluded if they lacked sufficient empirical data, defined as quantitative results supporting the effectiveness of the machine learning models, or if they were not directly relevant to the calibration of low-cost environmental sensors.

3.3. Analysis of selected papers

The selected papers were analysed using a systematic approach. A data extraction form was used to collect key information, including study objectives, sensor types, pollutants monitored, machine learning models applied, and performance metrics such as R^2 and RMSE. The extracted data were then synthesized to identify common themes, methodological approaches, and performance outcomes.

4. Comparative analysis

4.1. Introduction to the comparative analysis section

This section presents a comparison of different machine learning models that have previously been applied in the calibration of low-cost sensors, using as performance indicators R^2 and RMSE.

These are common metrics that allow practitioners to qualitatively evaluate how effective a model is:

R^2 describes how much variance is explained by the model; a value near one is a better fit. RMSE describes the average magnitude of errors, with small values denoting good performance.

Some recent studies have shown appreciable improvements in sensor calibration using these metrics. For example, Kumar & Sahu reported improvements in $PM_{2.5}$ calibration, while Mitchell et al showed high accuracy in methane sensor calibration with $R^2 = 0.997$ and substantial RMSE reduction [1,2]. Such metrics convey consistent and insightful measures of model performance in improving the accuracy of sensor data.

4.2. Comparative performance of machine learning models

Table 2 shows the summary of sensor calibration performance metrics in terms of R^2 and RMSE from different studies.

Table 2. Performance metrics of machine learning models in sensor calibration.

Project Name	Reference	Pollutant(s)	Sensor	Models (bolded and underlined have the best performance)	R^2	RMSE
Evaluation of Nine ML Algorithms for $PM_{2.5}$ Sensors	Kumar & Sahu, 2021 [1]	$PM_{2.5}$	Custom $PM_{2.5}$ Sensor	MLR, SVR, **<u>kNN</u>**, RT, **<u>RF</u>**, GB	0.75 to 0.97	72.24 to 0.31

Table 2. (continued).

Field Calibration Method for Low-Cost Sensors	Patra et al., 2021 [5]	$PM_{2.5}$, NO_2, O_3	Alphasense OPC-N3, NO_2-A43F	OLS, Elastic Net, **RF**, **GAM**	0.65 to 0.91	50.1 to 1.32
Calibration of a Low-Cost Methane Sensor	Mitchell et al., 2024 [2]	Methane	Figaro NGM2611-E13	Various ML models	0.78 to 0.99	10.5 to 5.1
Improving Data Reliability for Indoor PM Monitoring	Chojer et al., 2022 [6]	$PM_{2.5}$	Plantower PMS5003	Linear Regression, kNN, **SVR**	0.60 to 0.89	25.8 to 8.2
Humidity and Temperature Corrections for Low-Cost Sensors	Vajs et al., 2021 [7]	$PM_{2.5}$	Alphasense OPC-N2	Linear Regression, **RF**, ANN	0.50 to 0.82	30.5 to 12.1
Indoor Air Quality Monitoring and Source Apportionment	Higgins et al., 2024 [8]	$PM_{2.5}$, CO_2	Senseair S8, Plantower PMS7003	MLR, **SVR**, kNN, **RF**	0.62 to 0.88	22.4 to 10.5
Advancing Air Sensor Calibration in Stationary and Mobile Settings	Wang et al., 2022 [9]	$PM_{2.5}$, NO_2	Alphasense OPC-N3, NO_2-A43F	Linear Regression, **RF**, **LightGBM**	0.60 to 0.90	15.0 to 7.5
IoT LoRaWAN Connectivity and ML-Based Calibration	Ali et al., 2020 [10]	$PM_{2.5}$, NO_2, O_3	Custom IoT LoRaWAN Sensor	MLR, **SVR**, kNN, **RF**	0.70 to 0.92	18.4 to 8.7
Urban Air Quality Mapping with Mobile Sampling in Seoul	Lim et al., 2019 [11]	$PM_{2.5}$, NO_2, O_3	Alphasense OPC-N3, NO_2-A43F	Linear Regression, **RF**, **LightGBM**, SVR	0.67 to 0.89	20.5 to 9.8
ML Techniques to Improve Field Performance of Low-Cost Sensors	Bush et al., 2022 [12]	$PM_{2.5}$, NO_2, O_3, CO_2	Plantower PMS7003, Senseair S8, Alphasense OPC-N3	MLR, kNN, **RF**, **SVR**, ANN	0.65 to 0.94	25.0 to 11.2

4.3. Detailed Analysis by Pollutant

ML models have shown significant effectiveness in improving $PM_{2.5}$ sensor calibration. In some studies, strong performance was particularly seen in kNN, Random Forest, and SVR models, where improvements had reached 0.97 for R^2 and 0.31 for RMSE [1,6,7]. These models can handle nonlinearity effectively; hence, robust results can be achieved.

For NO_2 sensor calibration, Random Forest, Gradient Boosting, and Neural Networks have shown substantial improvements. These models registered an R^2 improvement to 0.92 and an RMSE reduction from 50.1 to 1.32 [9,10,11]. They capture the environmental variances effectively, thus increasing the accuracy of the measurements.

Other pollutants, including methane (CH_4), PM_{10}, O_3, and CO_2, also showed very good results through the machine learning models. For methane, the interaction model estimated RMSE between 4.5 to 5.1 ppm with R^2 between 0.997 and 0.998, which is a very high accuracy for methane measurements [2]. Like $PM_{2.5}$, PM_{10} calibration with Random Forest and Gradient Boosting models demonstrated high effectiveness, Calibration studies for O_3 and CO_2 using Random Forest and SVR methods showed R^2 improvements up to 0.94 and RMSE reductions up to 11.2 [8,12,13].

4.4. Improvements in Sensor Data Accuracy

The large decrease in RMSE and increase in R^2 is noticed with ML model calibration for all four pollutants. For instance, $PM_{2.5}$ calibration improved up to 0.97 in R^2 with an RMSE reduction to 0.31, NO_2 calibration up to 0.92 in R^2 with RMSE down to 1.32, and methane calibration up to 0.998 in R^2 with RMSE as low as 5.1 ppm.

4.5. Integration of Machine Learning with IoT and Mobile Sampling

Sampling Combining machine learning with IoT and mobile sampling technology improves sensor calibration accuracy to enhance real-time data collection and analysis. For instance, in South Korea, the spatial resolution and accuracy of an IoT system of air pollution monitoring systems were boosted by the employment of Random Forest models [10]. In Seoul, the use of mobile sensors integrated with the power of Random Forest and LightGBM models allowed a massive increase in the possibility of mapping urban air quality down to microscale levels [11]. These integrations provide fine-grained and real-time information about the environment at an increased level of detail, thus increasing the overall effectiveness of the sensor network.

5. Discussion

5.1. Key findings

The most important findings that emerge from the literature review and comparative analysis pertain to the fact that the calibration of low-cost sensors with machine-learning techniques yielded far better performance in terms of the significant improvement in average R^2 values, and reduction of RMSE for pollutants, particularly $PM_{2.5}$, NO_2, CH_4, O_3, and CO_2. These models added flexibility and ensured robustness when examining complex relationships between sensor readings and real pollutant concentrations in consideration of environmental variability and sensor drift. The joint combination of machine learning with IoT and mobile technologies improved the accuracy, real-time monitoring, and spatial resolution of information obtained from these low-cost sensors. For instance, they produced significant improvements in $PM_{2.5}$ calibration for both the Random Forest and SVR models, whereas very good accuracy concerning the methane calibration was obtained by the complex interaction models.

5.2. Gaps in the literature

Yet there exist several lacunas: the best model's performance depends largely on the quality and quantity of the training data used. A majority of these studies indicate that for higher accuracies and generalizations, large-scale, high-quality datasets are necessary. Though certain environmental parameters have been considered in the present models, a better approach is required to handle the environmental variability adequately. Very high computation resources and expertise are mandatory for dealing with advanced-level machine learning models, particularly Neural Networks. Simplification of such models without loss of accuracy should be helpful in their broader application. In addition, many machine-learning models are developed for site- or condition-specific applications and do not generalize well with other settings, urging the need for research to improve model transferability over environmental contexts.

5.3. Potential for future research

Future research would focus on the following lines: to improve the quality of diverse datasets by collecting more exhaustive data, dealing with sparsity, and ensuring consistency in quality across sensors and locations; use more advanced machine learning models, such as deep learning and ensemble methods, which hold better performance.

In such a case, real-time calibration methods should be developed for the continuous monitoring of sensors to ensure that consistent data accuracy exists, while IT and cloud-based technologies enable this purpose. Fourth is the enhancement of model transferability to different environmental scenarios and geographic locations. This relates to the specific identification of certain environmental and geographic variables that future models should consider enhancing their applicability across contexts.

6. Conclusion

6.1. Summary of key points

The literature is focused on the importance of machine learning techniques for the calibration of low-cost sensors measuring environmental parameters. In the majority of reviewed studies, such important improvements in both sensor accuracy and reliability are shown results in their being very cost-effective for large-scale environmental data collection. Machine learning models, such as Random Forest, SVR, and Neural Networks, have consistently proven to outperform traditional calibration techniques. These models attained remarkable improvement in R^2 with a significant decrease of RMSE in the estimation of pollutants including $PM_{2.5}$, NO_2, CH_4, O_3, and CO_2. The machine learning models Random Forest and Neural Networks appropriately accommodated sensor readings and their complicated and nonlinear associations with actual pollutant concentrations, controlling the drift of sensors and the environmental variability.

The combination of machine learning with IoT and mobile technologies enables the monitoring to be real-time and at higher spatial resolution, with improved data accuracy. Accordingly, various calibration approaches have been found to be optimal for the calibration of individual pollutants; for example, good $PM_{2.5}$ calibration was achieved with both Random Forest and SVR models, whereas high methane sensor accuracy was obtained with complex interaction models.

6.2. Implications for practice

The findings have a few implications in practice. Machine learning is accurate in calibration, and, for that matter, the natural environment will likely become increasingly represented with a high level of reliance in terms of accuracy related to policy decisions and public health programs. The use of machine learning allows continuous low-cost monitoring for real-life solutions at large scales. Such calibrations and algorithms integrated into the IoT scheme provide a path for real-time processing and calibration at any point where timely and relevant environmental assessment is required.

6.3. Potential for future research

Several important aspects deserve more attention in future research. Quality and diversity of the training data are very important, as they can impact the richness of collected data, sparsity in data, and quality data consistency in all the sensors and locations. Moreover, it would further push the calibration performance through advanced machine learning models—especially in deep learning and ensemble methods. Real-time calibration methods will, therefore, be developed using IoT and cloud-based technologies to allow on-the-fly monitoring and updating of the sensor data. The models of machine learning in transferability across environmental conditions and locations will, thus, guarantee robust performance under different contexts.

Conclusively, machine learning thoroughly improved the calibration of the low-cost sensors and hence provides more accurate and reliable environmental monitoring. This would signify a further step toward using innovative solutions for better collection and analysis of data to solve global environmental challenges.

DOI: 10.54254/2755-2721/71/20241635

References

[1] Kumar V and Sahu M 2021 Evaluation of nine machine learning regression algorithms for calibration of low-cost $PM_{2.5}$ sensor Journal of aerosol science

[2] Mitchell H, Cox S J and Lewis H G 2024 Calibration of a low-cost methane sensor using machine learning Sensors

[3] Morawska L et al. 2018 Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone? Environment international

[4] Ray S 2019 A quick review of machine learning algorithms

[5] Patra S S, Ramsisaria R, Du R, Wu T and Boor B E 2021 A machine learning field calibration method for improving the performance of low-cost particle sensors Building and environment

[6] Chojer H, Branco P T B S, Martins F G, Alvim-Ferraz M C M and Sousa S I V 2022 Can data reliability of low-cost sensor devices for indoor air particulate matter monitoring be improved? – An approach using machine learning Atmospheric environment

[7] Vajs I, Drajic D, Gligoric N, Radovanovic I and Popovic I 2021 Developing relative humidity and temperature corrections for low-cost sensors using machine learning Sensors

[8] Higgins C, Kumar P and Morawska L 2024 Indoor air quality monitoring and source apportionment using low-cost sensors. Environmental research communications

[9] Wang A, Machida Y, deSouza P, Mora S, Duhl T, Hudda N, Durant J L, Duarte F and Ratti C 2023 Leveraging machine learning algorithms to advance low-cost air sensor calibration in stationary and mobile settings Atmospheric environment

[10] Ali S, Glass T, Parr B, Potgieter J and Alam F 2020 Low cost sensor with IoT LoRaWAN connectivity and machine learning based calibration for air pollution monitoring. IEEE transactions on instrumentation and measurement

[11] Lim C C, Kim H, Vilcassim M J R, Thurston G D, Gordon T, Chen L C, Lee K, Heimbinder M and Kim S Y 2019 Mapping urban air quality using mobile sampling with low-cost sensors and machine learning in Seoul, South Korea Environment international

[12] Bush T, Papaioannou N, Leach F, Pope F D, Singh A, Thomas G N, Stacey B and Bartington S 2022 Machine learning techniques to improve the field performance of low-cost air quality sensors. Atmospheric measurement techniques

[13] Han P, Mei H, Liu D, Zeng N, Tang X, Wang Y and Pan Y 2021 Calibrations of low-cost air pollution monitoring sensors for CO, NO_2, O_3, and SO_2 Sensors