# Performance Evaluation of low-cost sensors under different aerosols

**Deepali Agrawal [1], Jakka Venkat Chandan[1], Anil Kumar Saini [2], Aakash C. Rai [3], and Prateek Kala [1, *]**

[1]Department of Mechanical Engineering, Birla Institute of Technology and Science Pilani,

Pilani, Rajasthan, India – 333031.

[2] Senior Scientist, SEG Design Group, CSIR – Central Electronics Engineering Research

Institute (CSIR-CEERI) Pilani, Pilani, Rajasthan, India – 333031.

[3]Department of Sustainable Energy Engineering, Kotak School of Sustainability, Indian Institute of Technology Kanpur, Kanpur, Uttar Pradesh, India – 208016.

## Abstract

Air regulatory authorities have publicly recognized the paradigm shift that low-cost sensors (LCSs) have brought about in supplementary traditional air monitoring. However, its large-scale uses have been severely limited because of concerns about its performance stability and data quality. Also, these inexpensive sensors need to be regularly calibrated against a standard instrument in order to guarantee the quality of the data they produce. Although a number of calibration studies have already been published, there is currently no transparent and standardized approach for PM sensors. In this work, the author calibrates two optical particle counters, Plantower PMSA003 and Sensirion SPS30, using a few machine learning algorithms (artificial neural network, random forest, support vector machine, and XGBoost) and some basic statistical regression algorithms (linear and quadratic). For calibration, the low-cost sensors are exposed to four different aerosols: Arizona road dust, compressor oil, incense, and salt (NaCl) particles in a controlled environmental chamber. The coefficient of determination ($R^2$) and root mean square error (RMSE) are used to evaluate the sensors' performance. Sensirion SPS30 sensors outperformed Plantower sensors among the examined sensors. Amongst the tested algorithms, the machine learning algorithms gave better results than the statistical regression algorithms. When Sensirion sensors are calibrated using a machine learning calibration technique, the RMSEs are determined to be less than 7 $\mu g/m^3$.

**Keywords**: environmental chamber, low-cost sensor, machine learning, calibration, particulate matter, artificial neural network.

## 1. Introduction

Globally, air pollution is responsible for over 7 million fatalities annually, making it one of the top environmental risk factors(Fuller et al., 2022). Revised evidence from epidemiological studies led to the establishment of revised guidelines by the World Health Organization in 2021, which lowered most of the limits per pollutant(World Health Organization, 2021). Particulate matter (PM) is the most harmful air contaminant to people on a global scale. PM is a complicated mixture of several liquid and solid particles which are airborne(Adams et al., 2015; da et al., 2014). While PM is not directly linked to cancer, it can have negative effects on public health based on its size. There is mounting evidence that PM particles pose a greater harm to human health as their size decreases. This is because smaller particles are able to penetrate deeper into the respiratory system, increasing the risk to the lung(Wang et al., 2015; Xing et al., 2016).

So, in order to combat the potentially harmful impacts on humans and climate change, measuring $PM_{2.5}$ is important for the development of air pollution control strategies, policies, and frameworks. The Federal Reference Methods (FRM) and the Federal Equivalent Methods (FEM) are the two main ways to assess $PM_{2.5}$(Chow, 1995; Noble et al., 2001). Despite their high price tag, these $PM_{2.5}$ monitoring methods are dependable and accurate. Additionally, they aren't usually in the best spots to record population exposure, and they're not easy to relocate near people's breathing zones(Aix et al., 2023). Portable, tiny systems that measure PM at a fine spatial and temporal scale are known as low-cost sensors (LCS). They are inexpensive (sometimes less than $100 each device) and offer data that is nearly real-time(Li et al., 2020). The use of LCS is on the rise globally, and new research is coming out of developing nations that lack official measurement stations. Because LCS is not going to be able to match the precision of reference-grade instruments, it is crucial that they be calibrated before use(Zimmerman, 2022). They are a useful addition to regulatory-grade networks after calibration and quality assurance, and they perform comparably to conventional reference equipment(Malings et al., 2020).

When calibrating LCS, a number of methods are employed, such as linear regression(Barkjohn et al., 2021; Wang et al., 2015), multiple linear regression(Barkjohn et al., 2021; Malings et al., 2020), and machine learning(Agrawal et al., 2024; Kumar & Sahu, 2021; Mahajan & Kumar, 2020). The use of regulatory and satellite data for spatial calibration also demonstrated efficiency in $PM_{2.5}$ calibration. When compared to multiple

linear regression (MLR), linear regression (LR) is often inaccurate for a number of reasons(Badura et al., 2018). When compared to machine learning, the performance of conventional linear approaches is sometimes severely lacking. In comparison to machine learning, they are simpler and more transparent, but they aren't very good at capturing cross-sensitivities between variables. An additional issue with ML models is the possibility of overfitting. In the case of overfitting, the model learns the training data excessively, leading to a high train score but a lower test score. The inability to generalize and accurate performance on fresh datasets are two problems that overfitted models suffer from. One approach to avoiding overfitting is to train and compare various ML algorithms with different underlying concepts in order to determine which one works best. Using multiple ML algorithms for LCS calibration is an uncharted territory.
The paper aims to accomplish two things: first, to compare and contrast the performance of various calibration algorithms for LCSs; and second, to assess the Plantower PMS A003 and Sensirion SPS30 against a research-grade optical particle spectrometer, the Grimm 11-A.

## 2. Methodology

### 2.1 Experimental chamber details

A 210-liter Plexiglas chamber (59.5 cm × 59.5 cm × 59.5 cm) has been built and constructed for the purpose of calibrating and evaluating low-cost $PM_{2.5}$ sensors, as seen in fig. 1. The chamber's temperature was regulated with a thermoelectric heating and cooling system at 25±1.0 °C, while relative humidity (RH) was controlled at 50±3% using a humidity controlling unit. The chamber's temperature and RH were consistently measured utilizing an air quality sensor (Greywolf DSIAQ-PLUSTAB10-DSII). The chamber received clean air via a vacuum pump equipped with a mass flow controller. To cleanse the recirculated air before it entered the experimental chamber, an activated carbon filter and a high-efficiency particulate air (HEPA) filter (Coda® XtraInline® Filters-GREEN CXGR-001) were installed in the air supply line. Seven fans were put within the chamber to ensure optimal mixing conditions: four on the rear wall, two on the front wall, and one on the ceiling. The chamber is linked to an atomizer-type particle generator to produce test aerosols with varying size distributions and compositions (Topas ATM 228). The particles are subsequently conveyed through the diffusion drier (Topas DDU 570) machine to produce dry aerosol particles within the chamber. A laser PM monitor quantified the size-resolved particle number concentration and mass (Grimm 11-A).

### 2.2 Instrument details

Plantower PMSA003-A and Sensirion SPS30 sensors, two distinct and low-cost particulate matter detectors, are utilized in the research. Particle aerosol spectrometer dust monitors (Grimm 11-A) are the gold standard for scientific study. Optical particle counters based on the light scattering principle are the sensors and the research-grade monitor. Table 1 provides details about the research-grade monitor, Grimm, and low-cost sensors that were used in this work. In order to gather information from the sensors, a printed circuit board (PCB) is linked to an Arduino Inc. programmed microcontroller Mega 2560 REV3.

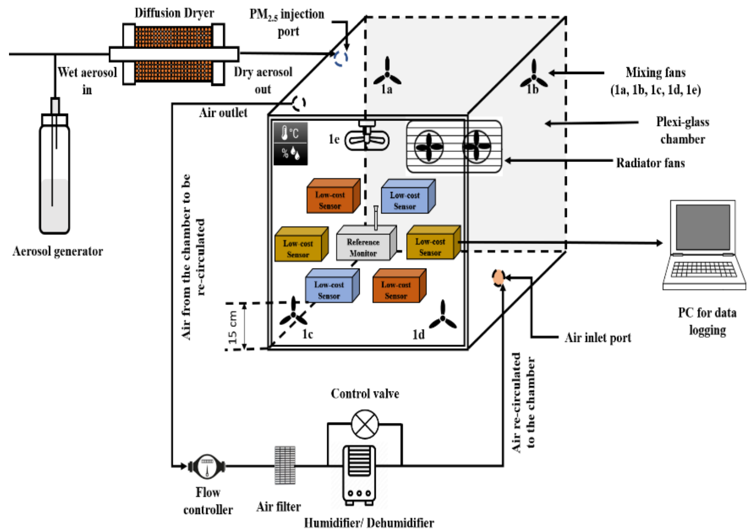



**Figure 1**. Schematic experimental setup diagram.

**Table 1**: Specifications of the research-grade monitor, Grimm, and the low-cost sensors.

| Sensor Details | Grimm 11-A | Plantower PMSA003 | Sensirion SPS30 |
|---|---|---|---|
| |  |  |  |
| Mini. size detection | 0.25 µm | 0.3 µm | 0.3 µm |
| Size ($mm^3$) | 240×130×70 | 38×35×12 | 41×41×12 |
| Power supply | 12 V DC (or 240 V AC) | 5V DC | 5V DC |
| Particle size range (µm) | 0.25 – 32 µm (31-size bins) | 0.3 – 10 µm (6-size bins) | 0.5 – 10 µm (5-size bins) |
| PM mass concentration | $PM_{10}$, $PM_{2.5}$, and $PM_{1.0}$ | $PM_{10}$, $PM_{2.5}$, and $PM_{1.0}$ | $PM_{10}$, $PM_{4.0}$, $PM_{2.5}$, and $PM_{1.0}$ |
| Concentration Range of measurement | 0 –1,00,000 µg/$m^3$ | 0 – 500 µg/$m^3$ | 0 – 1000µg/$m^3$ |
| Cost | $12,000 to $13,000 | $30 to $40 | $30 to $50 |

### 2.3 Experimental procedure

The chamber was sanitized with distilled (DI) water and dried, and the Grimm was positioned at its center. A 40 cm x 40 cm × 15 cm 3D-printed table was positioned centrally to accommodate low-cost $PM_{2.5}$ sensors. The chamber was vented until the $PM_{2.5}$ mass concentration fell below 1 µg/m³ or $N_{total}$ (total particle concentration) decreased to < 10,000 particles/L, while maintaining a temperature of 25 ± 1.0°C and the specified humidity. Aerosols were subsequently injected, elevating the $PM_{2.5}$ concentration to approximately 10,000 µg/m³. Upon reaching this level, particle injection ceased, and the system was enabled to equilibrate, therefore assuring a homogeneous particle distribution. Four out of seven mixing fans were deactivated after ten minutes. The particle concentration progressively diminished due to wall loss, attaining equilibrium in 60 minutes. Concurrent data from the reference monitor and economic sensors were documented from the initiation of particle injection until the concentration decreased below 1 µg/m³, with the entire experiment spanning around 2.5 to 3 hours.

## 3. Data Analysis
### 3.1 Low-cost sensor calibration

The low-cost sensors (Plantower and Sensirion) and the research-grade monitor Grimm are subjected to four distinct aerosols: ARD, compressor oil, incense, and NaCl. The outputs of these sensors are calibrated to the responses of a Grimm, employing six distinct calibration models for all evaluated aerosols. The six calibration models employed for testing are linear, quadratic, artificial neural network (ANN), random forest (RF), support vector regression (SVM), and extreme gradient boosting (XGBoost). Initially, the most fundamental type of regression model, known as the linear regression (LR) model, is employed. This model has a single independent variable (x), with the connection between the independent variable (x) and the dependent variable (y) presumed to be linear. Subsequently, quadratic calibration is generally employed when the anticipated connection between the measured and known data is curved or nonlinear. Additionally, Artificial neural networks (ANNs) are modeled after the human brain and replicate the signaling processes between biological neurons used in the study. The current study uses a feedforward network consisting of a single hidden layer with two nodes. Also, the network consists of a single input and a single output.
The random forest (RF) model tackles regression and classification challenges by generating several decision trees during training and averaging their predictions. Renowned for its resilience against overfitting and capacity to manage extensive, high-dimensional datasets, Random Forest is adept at identifying non-linear correlations and feature interdependencies.

Furthermore, the support vector regression (SVM) method is useful when dealing with nonlinear problems. The objective of employing SVR is to ascertain a function f(x) that delineates the link between the predictors and the target variable. Lastly, XG Boost (eXtreme Gradient Boosting) is a machine learning algorithm that uses gradient boosting to make predictions. It's a powerful and versatile algorithm known for its accuracy and efficiency, particularly when dealing with large datasets. XG Boost combines the predictions of multiple weak learners (usually decision trees) to create a stronger model.

### 3.2 Performance parameters

Model performances were evaluated using various measures, specifically the coefficient of determination ($R^2$), mean bias error (MBE), and root mean square error (RMSE). Initially, the coefficient of determination ($R^2$), a statistical metric assessing the predictive capacity of one variable's fluctuation on another, is employed. It assesses the strength of the linear relationship between two variables. However, $R^2$ alone is insufficient to assess the performance of low-cost sensors; thus, the RMSE metric is also used to assess the LCS's performance. RMSE is concerned with discrepancies between predicted and observed values, and it calculates an average error magnitude. It is computed by taking the mean squared deviation between the projected and actual values and taking the square root of it.

## 4. Results

Fig. 2a) and b) show the $PM_{2.5}$ concentrations measured by the Plantower and Sensirion sensors four units each, along with the research-grade monitor Grimm 11-A, with respect to time. The figure illustrates that all the sensors can qualitatively detect the trend of particle decay within the chamber; however, the recorded $PM_{2.5}$ concentrations markedly diverge from the values obtained by the research-grade monitor.

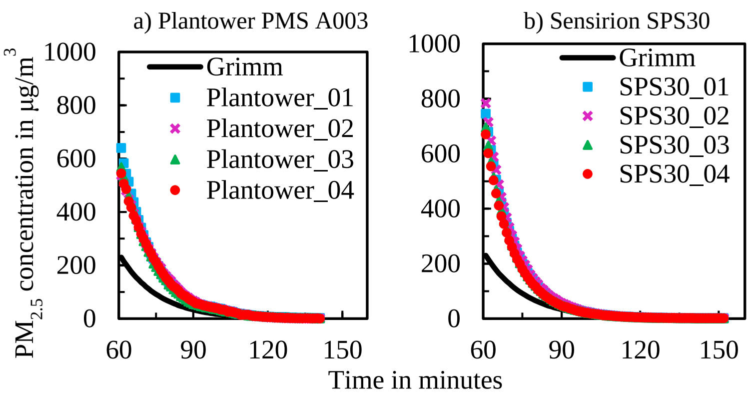



**Figure 2**: $PM_{2.5}$ concentration decay curves for salt particles (NaCl) for the a) Plantower sensors, and b) Sensirion sensors.

### 4.2 Combined calibration model

**Table 2**: Coefficient of determination ($R^2$) for different low-cost sensors: a) Plantower, and b) Sensirion when exposed to various aerosols.

| Plantower PMS-A003(A) | | | | |
|---|---|---|---|---|
| | ARD | Compressor oil | Incense | NaCl |
| Linear | 0.9813 | 0.9930 | 0.9879 | 0.9933 |
| Quadratic | 0.9928 | 0.9958 | 0.9958 | 0.9963 |
| ANN | 0.9997 | 0.9998 | 0.9998 | 0.9998 |
| RF | 0.9965 | 0.9974 | 0.9984 | 0.9978 |
| SVM | 0.9945 | 0.9962 | 0.9965 | 0.9937 |
| XG Boost | 0.9967 | 0.9969 | 0.9980 | 0.9957 |
| **Sensirion SPS30** | | | | |
| Linear | 0.9143 | 0.9679 | 0.9429 | 0.9705 |
| Quadratic | 0.9805 | 0.9941 | 0.9914 | 0.9943 |
| ANN | 0.9988 | 0.9972 | 0.9986 | 0.9972 |
| RF | 0.9994 | 0.9984 | 0.9990 | 0.9986 |
| SVM | 0.9983 | 0.9959 | 0.9966 | 0.9970 |
| XG Boost | 0.9986 | 0.9973 | 0.9990 | 0.9978 |

Fig. 3 a), b), c), and d) demonstrate the combined responses of Plantower sensors (all four units) against the research-grade monitor Grimm for four different aerosols (ARD, compressor oil, incense, and NaCl) for the random forest prediction model. Table 2 represents the $R^2$ values for LCSs (Plantower and Sensirion) when subjected to different aerosols. Interestingly, all six calibration models (linear, quadratic, ANN, RF, SVM, and XG Boost) yield outstanding results with $R^2 \geq 0.9900$ for Plantower sensors when exposed to different aerosols.

Analogous to Plantower sensors, fig. 4a–d) depict the aggregated responses of four Sensirion sensors plotted versus Grimm for four distinct aerosols (ARD, compressor oil, incense, and NaCl) for the RF prediction model. Unlike Plantower sensors, the $R^2$ values for the Sensirion sensors are not greater than 0.99 for all six calibration models for all the tested aerosols. For Sensirion sensors, all the ML models have $R^2 \geq 0.9970$ compared to statistical calibration models: quadratic ($R^2$ = 0.9805 – 0.9914) and linear ($R^2$ = 0.9143 – 0.9705) for all the tested aerosols. It is discovered that various sensor types, when calibrated using a common calibration equation, yield superior results with the machine learning calibration model, particularly the random forest (RF); however, the linear model produces the least accurate outcomes.

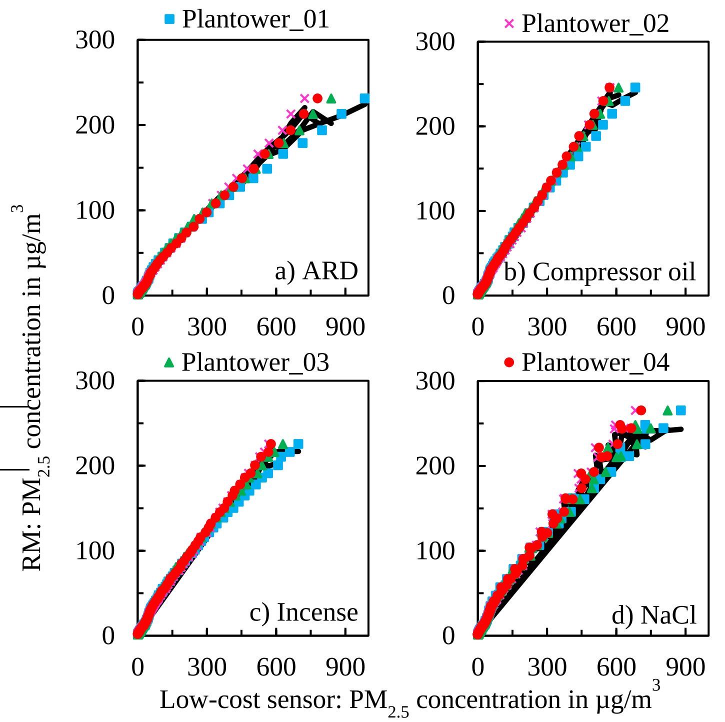



**Figure 3**: Pairwise correlation between combined sensor outputs and Grimm $PM_{2.5}$ measurements for Plantower sensors for four different aerosols: a) ARD, b) Compressor oil, c) Incense, and d) NaCl.

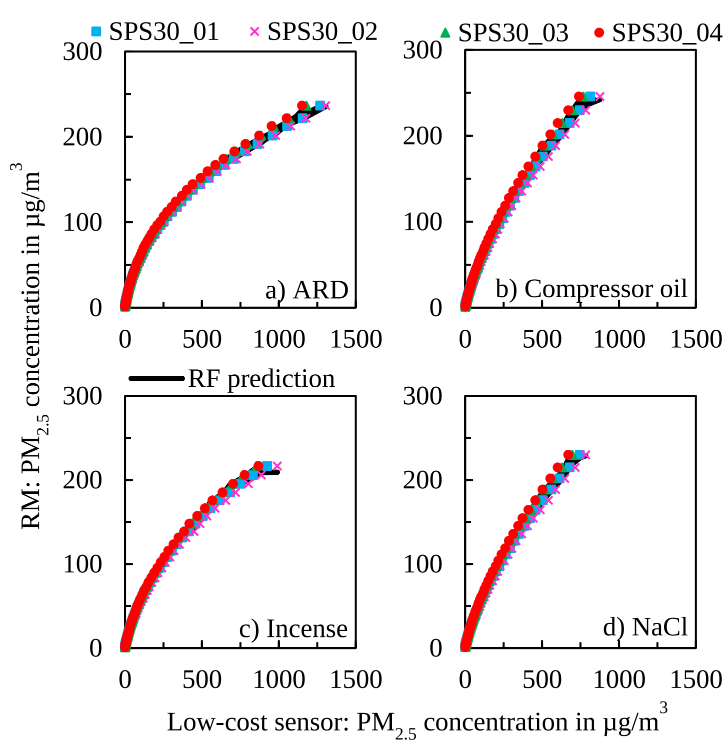


**Figure 4**: Pairwise correlation between combined sensor outputs and Grimm $PM_{2.5}$ measurements for Sensirion sensors for four different aerosols: a) ARD, b) Compressor oil, c) Incense, and d) NaCl.

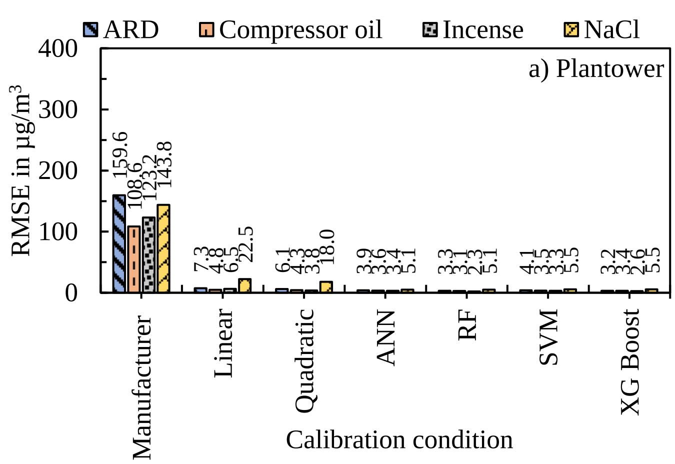
### 4.4 Sensor performance

Model performances of sensors when subjected to different aerosols are evaluated by calculating the root mean square error (RMSE). The subsequent section will provide a detailed discussion of the performance matrix as sensors are calibrated under varying situations.

#### 4.4.1. Root Mean Square Error (RMSE)

For the performance evaluation, the RMSE values are observed in fig. 5 when sensors were subjected to four different aerosols (ARD, compressor oil, incense, and NaCl). Figure 6 a) and b) shows the average RMSE for all kinds of tested LCSs (Plantower and Sensirion) obtained with the manufacturer's referred calibration and the different models used for the analysis.

When using different calibration models, the RMSEs are the least when sensors are calibrated with a random forest (RF) model for all kinds of sensors when exposed to different aerosols. In case of Plantower sensors (refer Fig. 6a), the RMSEs range between 108.6 μg/m$^3$ and 159.6 μg/m$^3$ with the manufacturer's calibration, while with the RF model, the errors were found to be between 2.3 μg/m$^3$ and 5.1 μg/m$^3$ for the tested aerosols. However, for Sensirion sensors (see Fig. 6 b), the RMSEs turn out to be 1.6 μg/ m$^3$ to 3.6 μg/m$^3$ with the RF calibration model, from 124.6 μg/m$^3$ to 236.3 μg/m$^3$ as per the manufacturer's preferred calibration. Amongst all the tested calibration models, the RF gave the best results, followed by XG Boost, ANN, SVM, and the statistical (linear and quadratic) models for both sensors when exposed to various aerosols.



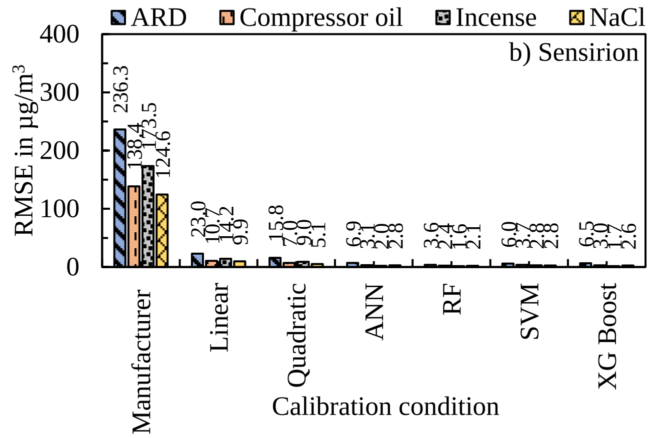




**Figure 5**: Root Mean Square Errors (RMSEs) for two different LCSs when subjected to various aerosols: a) Plantower and b) Sensirion sensors.

### 5. Conclusion

This study involves laboratory studies for calibration and performance assessment of two low-cost particulate matter sensors: the Plantower PMSA003-A and the Sensirion SPS30. These sensors are subjected to four distinct aerosols for calibration: ARD, compressor oil, incense, and NaCl. Six distinct calibration algorithms, linear, quadratic, artificial neural network (ANN), random forest (RF), support vector machine (SVM), and eXtreme gradient boosting (XG Boost), are utilized for the responses of LCSs.

When subjected to distinct aerosols, the Plantower and Sensirion sensors exhibited an $R^2$ value of 0.99 or higher when utilizing the machine learning calibration models. Both Plantower and Sensirion sensors performed better when calibrated with a random forest (RF) calibration model, with RMSEs less than 5.1 μg/m$^3$ for all the tested aerosols. This study indicates that low-cost sensors lack sufficient factory calibration. Calibration of all sensors is essential before any practical application. Under proper calibration, the Sensirion SPS30 sensors surpass all others. The machine learning calibration methods yielded the most favorable results, succeeded by the quadratic and linear calibration techniques.

## References

Adams, K., Greenbaum, D. S., Shaikh, R., van Erp, A. M., & Russell, A. G. (2015). *Particulate matter components, sources, and health: Systematic approaches to testing effects. Journal of the Air and Waste Management Association*, *65*(5), 544–558.

Agrawal, D., Saini, A. K., Rai, A. C., & Kala, P. (2024). *In Chamber Calibration and Performance Evaluation of Air Quality Low-Cost Sensors. Atmospheric Pollution Research*, *15*(12), 102299.

Aix, M. L., Schmitz, S., & Bicout, D. J. (2023). *Calibration* *methodology of low-cost sensors for high-quality monitoring of fine particulate matter*. *Science of the Total Environment*, *889*(May).

Badura, M., Batog, P., Drzeniecka-Osiadacz, A., & Modzel, P. (2018). *Evaluation of low-cost sensors for ambient PM2.5 monitoring*. *Journal of Sensors*, *2018*.

Barkjohn, K. K., Gantt, B., & Clements, A. L. (2021). *Development and application of a United States-wide correction for PM2.5 data collected with the PurpleAir sensor*. *Atmospheric Measurement Techniques*, *14*(6), 4617–4637.

Chow, J. C. (1995). *Measurement methods to determine compliance with ambient air quality standards for suspended particles*. *Journal of the Air and Waste Management Association*, *45*(5), 320–382.

da, P., Levizzani Agostino Rappazzo, V., & Franco Belosi Dott Fabrizio Ravegnani, D. (2014). *Characterisation and calibration of optical counters for airborne particulate matter*.

Fuller, R., Landrigan, P. J., Balakrishnan, K., Bathan, G., Bose-O'Reilly, S., Brauer, M., Caravanos, J., Chiles, T., Cohen, A., Corra, L., Cropper, M., Ferraro, G., Hanna, J., Hanrahan, D., Hu, H., Hunter, D., Janata, G., Kupka, R., Lanphear, B., … Yan, C. (2022). *Pollution and health: a progress update*. *The Lancet Planetary Health*, *6*(6), e535–e547.

Kumar, V., & Sahu, M. (2021). *Evaluation of nine machine learning regression algorithms for calibration of low-cost PM2.5 sensor*. *Journal of Aerosol Science*, *157*(May), 105809.

Li, J., Zhang, H., Chao, C. Y., Chien, C. H., Wu, C. Y., Luo, C. H., Chen, L. J., & Biswas, P. (2020). *Integrating low-cost air quality sensor networks with fixed and satellite monitoring systems to study ground-level PM2.5*. *Atmospheric Environment*, *223*(January), 117293.

Mahajan, S., & Kumar, P. (2020). *Evaluation of low-cost sensors for quantitative personal exposure monitoring*. *Sustainable Cities and Society*, *57*, 102076.

Malings, C., Tanzer, R., Hauryliuk, A., Saha, P. K., Robinson, A. L., Presto, A. A., & Subramanian, R. (2020). *Fine particle mass monitoring with low-cost sensors: Corrections and long-term performance evaluation*. *Aerosol Science and Technology*, *54*(2), 160–174.

Noble, C. A., Vanderpool, R. W., Peters, T. M., McElroy, F. F., Gemmill, D. B., & Wiener, R. W. (2001). *Federal reference and equivalent methods for measuring fine particulate matter*. *Aerosol Science and Technology*, *34*(5), 457–464.

Wang, Y., Li, J., Jing, H., Zhang, Q., Jiang, J., & Biswas, P. (2015). *Laboratory Evaluation and Calibration of Three Low-Cost Particle Sensors for Particulate Matter Measurement*. *Aerosol Science and Technology*, *49*(11), 1063–1077.

World Health Organization. (2021). *WHO ambient air quality database 2021 update*.

Xing, Y. F., Xu, Y. H., Shi, M. H., & Lian, Y. X. (2016). *The impact of PM2.5 on the human respiratory system*. *Journal of Thoracic Disease*, *8*(1), E69–E74.

Zimmerman, N. (2022). *Tutorial: Guidelines for implementing low-cost sensor networks for aerosol monitoring*. *Journal of Aerosol Science*, *159*(August 2021), 105872.