Annals of Work Exposures and Health, 2026, 70, wxag021
https://doi.org/10.1093/annweh/wxag021
Published: 20 April 2026
Original article

BOHS British Occupational Hygiene Society · The Chartered Society for Worker Health Protection · OXFORD

Comparison of calibration models for low-cost PM2.5 sensors in high-concentration occupational environments

Runcheng Fang[1], Scott Collingwood[2], Kerry Kelly[3], and Darrah Sleeth[1*]

[1]Division of Occupational & Environmental Health, University of Utah, 295 Chipeta Way, Salt Lake City, UT 84108, United States
[2]Department of Pediatrics, University of Utah, 295 Chipeta Way, Salt Lake City, UT 84108, United States
[3]Department of Chemical Engineering, University of Utah, 50 S Central Campus Drive, Room 3290 MEB, Salt Lake City, UT 84112, United States
*Corresponding author: Email: Darrah.Sleeth@hsc.utah.edu.

Abstract

Low-cost PM2.5 sensors are increasingly used for ambient and indoor air quality monitoring, but their performance in high-concentration occupational environments remains uncertain. This study compares linear regression (LR), polynomial regression (PR), and random forest (RF) calibration models for low-cost PM2.5 sensors operating in a controlled high-concentration chamber. This study assessed the calibration performance of low-cost aerosol sensors (LCS) by employing LR, PR, and RF models across various aerosol (PM2.5) concentration ranges. Conventional LR models demonstrated solid performance when dealing with lower concentrations (0 to 150 μg/m^3), but their accuracy diminished when confronted with higher concentrations. On the other hand, RF models continuously demonstrated better performance over all concentration ranges, making them potentially more appropriate for occupational situations with higher and fluctuating aerosol levels. PR models exhibited intermediate performance, surpassing that of LR but falling short of the robustness demonstrated by RF. The study highlights the importance of using modern calibration methods such as RF in locations with high concentrations to enable accurate aerosol monitoring using LCS. These results could lead to the development of efficient and cost-effective methods for monitoring air quality using LCS for high concentration environments.

Keywords air quality monitoring, PM2.5 concentration, low-cost sensors (LCS), sensor calibration, occupational health

What's Important About This Paper?

Low-cost PM2.5 sensors are increasingly used in workplaces, yet calibration performance at high and rapidly changing concentrations remains under-characterized. This study compared multiple calibration approaches across concentration ranges and identified a concentration where nonlinearity becomes pronounced. Results suggest that range-aware modeling—particularly ensemble methods—can sustain low error even above typical occupational peaks, offering practical guidance for exposure assessment with affordable instruments. By providing reproducible workflows and clear recommendations on when to favor data-driven models, the work supports more reliable decision-making in industrial hygiene.

Received: September 23, 2025. **Accepted:** February 23, 2026

Introduction

Airborne dust, or aerosol, is well known for its adverse health effects on the respiratory and cardiovascular system (Pope et al. 2002). Traditional time-integrated sampling methods remain the standard approach for regulatory compliance because they are well-established, robust, and cost-effective, but they do not provide information about temporal variability in exposure. The advent of low-cost sensors (LCS) offers a promising alternative, enabling affordable deployment of extensive monitoring networks across various environments (Liang 2021; Liang and Daniels 2022). Among the many available options, the Plantower series of sensors has become one of the most widely used low-cost PM sensors (eg PurpleAir, Tellus Network Solutions, etc.), valued for their affordability and reasonable accuracy in measuring PM2.5 concentrations (Liang 2021). These sensors estimate particle mass via light scattering, and are influenced by particle size distribution, composition, and environmental factors like relative humidity (RH) and temperature (Zou et al. 2021; Russell et al. 2022). Accurate calibration is therefore essential to mitigate these environmental influences and ensure sensor reliability.

Linear regression (LR) models have commonly been used to calibrate LCS effectively at lower PM2.5 concentrations (ie 0 to 150 μg/m^3) (Kelly et al. 2017; Levy Zamora et al. 2019; Sayahi et al. 2019; Hegde et al. 2020; Kuula et al. 2020; Liang and Daniels 2022). However, the performance of LR models at higher concentrations is less well understood, where nonlinear effects can become significant (Williams et al. 2015; Giordano et al. 2021; Fang et al. 2025). LR models, favored for simplicity and interpretability, can fail to adequately capture nonlinear sensor behaviors at higher aerosol concentrations (Zimmerman et al. 2018; Sayahi et al. 2019; Wang et al. 2019, 2020; Barkjohn et al. 2021; Liang and Daniels 2022). Therefore, a gap remains in our knowledge of LCS calibration at higher concentration levels (>150 μg/m^3) that are especially relevant for occupational exposure.

Another option is polynomial regression (PR) models, which extend LR by incorporating polynomial terms, capturing moderate nonlinearities and potentially providing better fits for transitional concentration ranges (Liang and Daniels 2022; Vajs et al. 2023). Additionally, random forest (RF), an ensemble learning approach, combines multiple decision trees and excels in capturing complex, nonlinear interactions among sensor readings, temperature, and RH, making it potentially ideal for higher concentration scenarios (Jaiswal and Samikannu 2017; Wang et al. 2019).

It should be noted that most LCSs measure PM2.5 concentration (and sometimes PM10 and/or PM1), which is most relevant for ambient air quality regulation. Occupational settings are not typically regulated for PM2.5 concentrations (with the exception of some limited US state regulation of PM2.5 concentration as it relates to wildfire smoke). Instead, occupational exposure assessment focuses on particle size fractions that have been shown to be associated with health-based outcomes (ie respirable or inhalable aerosols). For example, the US Occupational Safety and Health Administration (OSHA) has established an 8-h time-weighted average permissible exposure limit of 15 mg/m^3 for total PM and 5 mg/m^3 for respirable PM (OSHA 2012). Although PM2.5 is a component of respirable PM, it is not separately regulated in occupational standards.

Occupational environments frequently experience concentrations far exceeding ambient regulatory limits—often surpassing 200 μg/m^3 and even 500 μg/m^3 (Shezi et al. 2020; Masri et al. 2022). If the use of LCS is going to continue to expand into such occupational environments, there is a critical need for accurate sensor calibration beyond standard ambient conditions, facilitating better occupational exposure assessments.

This study uniquely addresses this current gap by systematically evaluating LR, PR, and RF calibration models across distinct aerosol concentration segments relevant to occupational environments. Specifically, this study aims to: (i) assess model performance across low (0 to 150 μg/m^3), transitional (150 to 348 μg/m^3), and high (>348 μg/m^3) aerosol concentrations; (ii) identify optimal calibration strategies for each concentration range; and (iii) pinpoint the concentration breakpoint (BP) signaling a necessary shift from linear to more sophisticated calibration models. Unlike previous research that typically applied a single calibration algorithm uniformly across all exposure levels, this approach introduces a novel range-specific calibration strategy. By explicitly defining these concentration segments, this method should enhance model accuracy and practical applicability in occupational settings, providing a foundation for more precise and reliable workplace aerosol monitoring.

Methods

Calibration data were collected following the procedures of Sayahi et al. (2019), using ammonium nitrate aerosols in a controlled chamber environment. The setup included eight PMS3003 low-cost PM2.5 sensors and a TSI DustTrak 8530 (TSI Inc., Shoreview, MN) reference monitor. Aerosols were generated using a single-jet atomizer (TSI 9302) supplied with dry air, maintaining RH at 7.3% to 9.9% and temperatures between 21.7 and 25.1 °C. Reference PM2.5 concentrations ranged from <5 μg/m^3 (the sensor limit of

Table 1 Overall model performance for each sensor (bold indicates best model performance).

Sensor	Model	MAE (μg/m^3)	PE (%)	MSE (μg/m^3)^2	R^2
Sensor 187	Linear regression	8.15	1.35	130.50	0.971
	Polynomial regression	5.80	1.10	75.23	0.984
	Random forest	**3.01**	**0.65**	**25.67**	**0.992**
Sensor 188	Linear regression	10.25	2.00	150.30	0.953
	Polynomial regression	7.35	1.75	100.25	0.964
	Random forest	**4.50**	**1.10**	**30.50**	**0.981**
Sensor 189	Linear regression	12.50	2.10	160.50	0.942
	Polynomial regression	6.10	1.20	80.23	0.967
	Random forest	**3.80**	**0.75**	**20.45**	**0.992**
Sensor 190	Linear regression	9.30	1.75	140.25	0.957
	Polynomial regression	5.50	1.30	70.75	0.979
	Random forest	**3.25**	**0.85**	**27.45**	**0.990**
Sensor 191	Linear regression	11.75	2.20	175.25	0.934
	Polynomial regression	7.00	1.80	95.50	0.953
	Random forest	**4.10**	**0.95**	**35.60**	**0.977**
Sensor 192	Linear regression	8.80	1.90	135.75	0.956
	Polynomial regression	6.00	1.25	78.50	0.971
	Random forest	**3.45**	**0.80**	**28.75**	**0.990**
Sensor 193	Linear regression	10.50	2.05	145.50	0.952
	Polynomial regression	6.80	1.60	85.25	0.961
	Random forest	**4.25**	**1.00**	**32.50**	**0.984**
Sensor 194	Linear regression	11.00	2.15	150.75	0.942
	Polynomial regression	7.10	1.70	90.50	0.951
	Random forest	**3.75**	**0.90**	**26.75**	**0.992**

*R^2, coefficient of determination; MAE, mean absolute error; PE, percent error; MSE, mean squared error.

detection [LOD]) to approximately 1,000 μg/m^3, monitored at 1-min intervals after averaging 15-s measurements from both reference and sensor devices. Due to the use of a controlled chamber environment, this study sought to isolate model performance from environmental variability.

PMS3003 outputs were logged as 15-s averaged PM2.5 concentrations. For calibration, we averaged each set of four consecutive 15-s readings to obtain a single 1-min PMS3003 value, which was paired with the corresponding 1-min DustTrak PM2.5 concentration.

Three modeling approaches were applied to calibrate sensor outputs against the DustTrak reference: LR, Polynomial (PR), and RF. LR assumes a direct linear relationship between sensor readings (X) and reference concentrations (Y), and is most suitable for lower concentration ranges where sensor behavior is well understood to be linear:

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \tag{1}$$

where Y is the DustTrak PM2.5 concentration, X is the PMS3003 PM2.5 reading, β_0 is the intercept, β_1 is the slope relating the sensor to the reference, and ε is the error term.

PR extends LR by including a nonlinear term in the sensor PM2.5 reading, allowing the model to fit curved relationships in the data that may arise at higher concentrations. The PR model used in this study was a quadratic polynomial:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon \tag{2}$$

where β_2 captures curvature in the relationship between the PMS3003 and DustTrak measurements, and the other parameters have the same meaning as in Equation (1).

RF models were implemented as nonparametric regressors of the form $Y = f(X)$, where $f(\cdot)$ is an ensemble of decision trees (RandomForestRegressor in scikit-learn). The number of trees and the maximum tree depth were tuned using cross-validation to balance bias and variance and to reduce overfitting. For each sensor, LR, PR, and RF models were fit both (i) across the full concentration range and (ii) separately within each concentration segment defined by the BP analysis (segments A to C).

Model performance was evaluated using mean absolute error (MAE), percentage error (PE), mean squared error (MSE), and coefficient of determination (R^2). Analyses were performed using R software (tidyverse, ggplot2, lmtest) and Python (scikit-learn, pandas, NumPy, matplotlib). RF model evaluation additionally included feature importance and mean prediction to identify influential factors and model adjustment across concentration ranges.

A piecewise LR analysis was also used to identify a concentration BP marking the shift from linear to nonlinear sensor response. Using 100 candidate BPs across measured concentrations, the BP with the lowest combined sum of squared residuals was selected. Prior literature (Sayahi et al. 2019) suggests sensor nonlinearity might begin at lower concentrations, around 150 μg/m^3. Therefore, two thresholds were used for BP analysis of three concentration segments: Segment A (0 to 150 μg/m^3), Segment B (150 μg/m^3 to BP), and Segment C (>BP). Models were evaluated separately within each segment.

Results

Overall model analysis

Table 1 shows that the RF model consistently outperformed both LR and PR across all sensors for the entire concentration range. RF consistently achieved the lowest error metrics (MAE, PE, MSE) and the highest R^2 values, indicating strong model fit. Although LR performed reasonably well, with R^2 values above 0.93, it was less accurate at capturing nonlinear patterns, as evidence by the higher MAE (mean: 10.28 μg/m^3 vs. 3.76 μg/m^3). PR improved upon LR by accounting for some curvature in the data but remained less robust than RF. RF models had the lowest MAE and PE, minimized variance, and best explained the variance in PM2.5 concentrations. These findings demonstrate that RF was the most effective model for calibrating low-cost PM2.5 sensors across a wide concentration range. Figure 1 shows the relationship between the DustTrak PM2.5 concentrations and the corresponding PMS3003 PM2.5 readings for each of the eight sensors (Sensors 187 to 194). At lower concentrations (approximately below 150 μg/m^3), the PMS3003 readings closely follow the DustTrak values, indicating an approximately linear response. As concentrations increase, especially above about 350 μg/m^3, the PMS3003 readings begin to plateau relative to the DustTrak, and the sensors increasingly underestimate the reference. This pattern visually illustrates the transition from linear to nonlinear behavior and motivates the BP and segmented calibration analyses described below.

BP analysis

Piecewise LR identified an optimal BP at 347.78 μg/m^3, rounded to 348 μg/m^3 for practical purposes. Combined with the previously identified BP of 150 μg/m^3, three segments were then analyzed: Segment A (0 to 150 μg/m^3), Segment B (150 to 348 μg/m^3), and Segment C (>348 μg/m^3). Model performance was then evaluated across the three concentration segments for each type of model, and detailed numeric results are provided in Tables S2 (LR), S3 (PR), and S4 (RF).

In Segment A (0 to 150 μg/m^3), where sensor responses were mostly linear, all models performed relatively well (see Fig. 2). LR showed good accuracy, with low errors and high R^2 values, but RF outperformed both LR and PR across all sensors. RF achieved the lowest MAE ranging from 0.75 to 2.50 μg/m^3 and PE from 0.38% to 1.00%, with consistently high R^2 values between 0.98 and 0.99. PR had slightly higher errors (MAE: 1.40 to 4.50; PE: 0.70% to 1.80%) but remained reasonably accurate. Of note in this segment is that all the MAE values were below the sensor LOD (5 μg/m^3), which suggests excellent practical performance.

In Segment B (150 to 348 μg/m^3), as concentrations increased and sensor behavior began to deviate from linearity, LR performance slightly declined, with maximum MAE rising to 5.62 μg/m^3 (which is above the sensor LOD) and PE up to 2.21%, albeit maintaining high R^2. RF maintained better accuracy with lower MAE (0.87 to 2.43 μg/m^3), PE (0.38% to 0.99%), and near-perfect R^2 values (0.996 to 0.999). PR performed moderately, with MAE between 1.74 and 6.44 μg/m^3 and PE up to 2.65%, and similarly high R^2. These results are presented in Fig. 3.

In Segment C (>348 μg/m^3), where nonlinearity became most pronounced, LR showed the greatest limitations, with MAE ranging from 8.13 to 15.42 μg/m^3 and declining R^2 values (0.970 to 0.992). PR performed better than LR but was still prone to less accuracy (MAE: 4.51 to 6.74 μg/m^3; PE: 0.85% to 1.31%). RF continued to deliver the most accurate results, with MAE between 1.82 and 3.69 (still below the LOD of the sensor), PE between 0.34 to 0.74%, and R^2 values ranging from 0.994 to 0.999. These results are presented in Fig. 4.

Discussion

The findings from this study demonstrate the varying effectiveness of different calibration models for LCS across different aerosol concentration ranges, especially relevant for use in occupational settings. LR models, while effective at lower concentrations (0 to 150 μg/m^3), showed limitations as concentrations increased. In contrast, RF

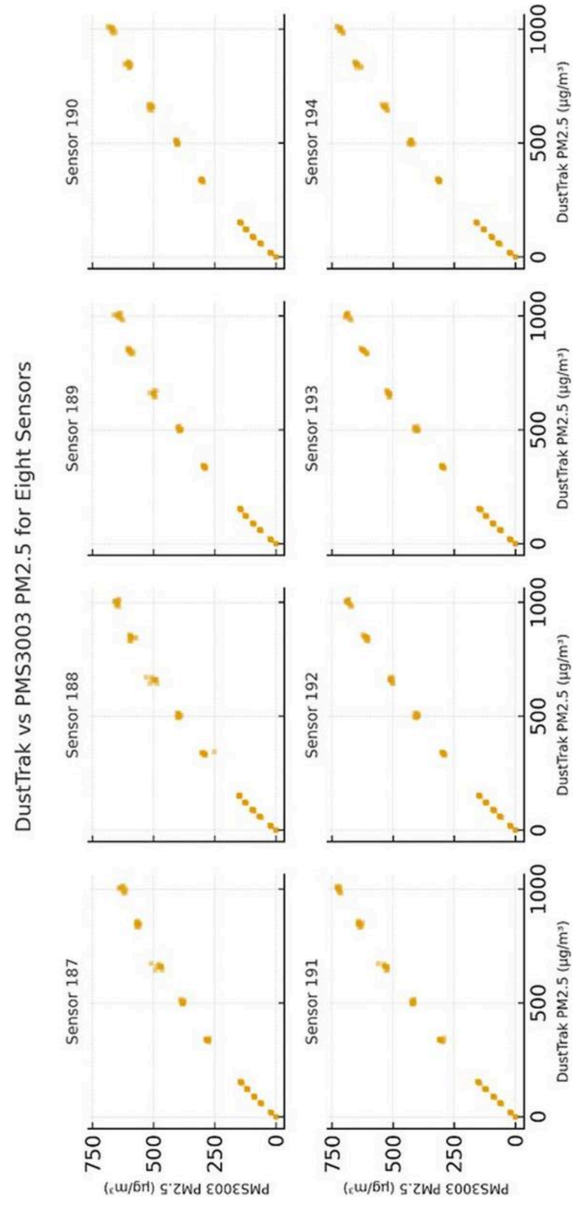

Figure 1 Relationship between DustTrak PM2.5 concentrations and PMS3003 PM2.5 readings for all sensors.

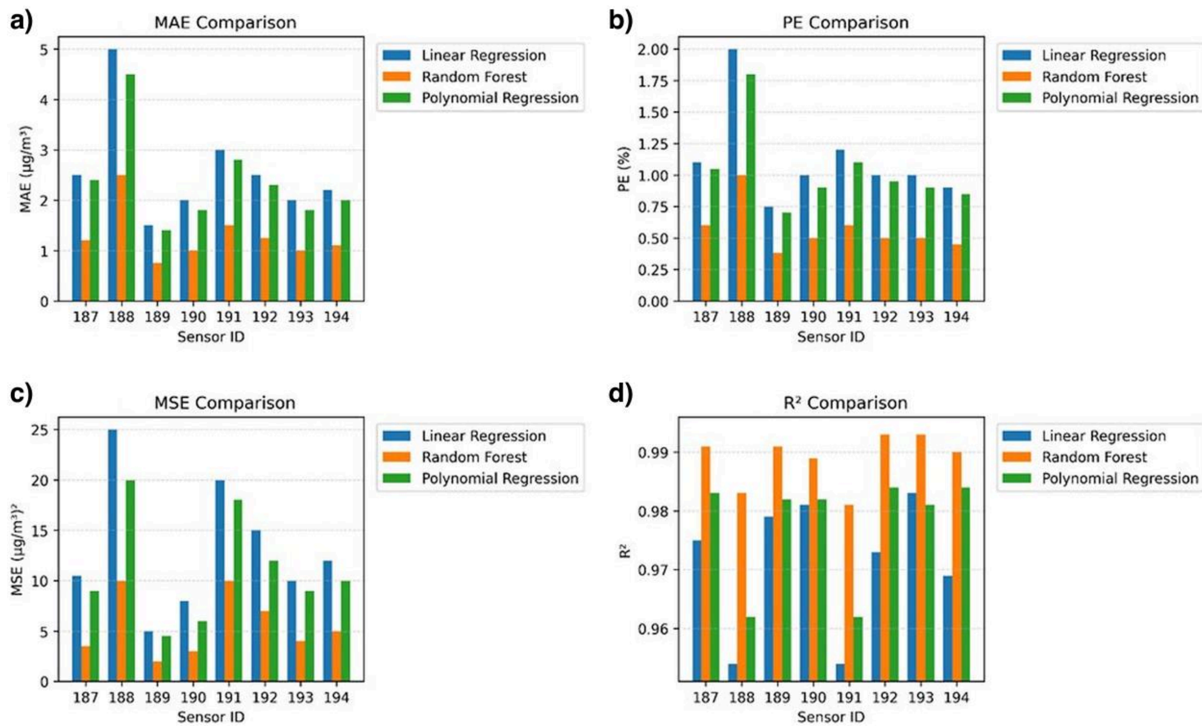

Figure 2 Performance comparison of calibration models in Segment A (a) MAE, (b) PE, (c) MSE, and (d) coefficient of determination (R^2) for different models and sensors.

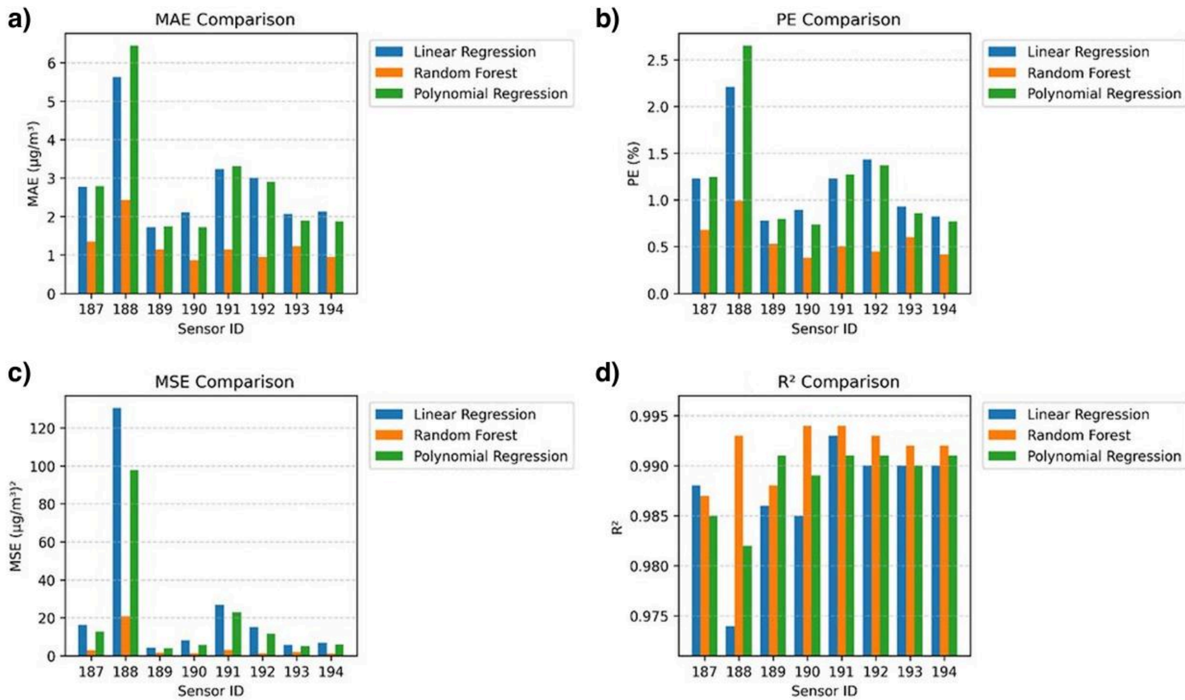

Figure 3 Performance comparison of calibration models in Segment B. (a) MAE, (b) PE, (c) MSE, and (d) coefficient of determination (R^2) for different models and sensors.

models consistently exhibited superior performance across all concentration ranges, including at higher levels more common in occupational settings.

Overall model comparison

The results revealed clear performance differences among the three calibration models across the entire PM2.5 concentration range. For example, Sensor 187 calibrated using LR showed relatively high errors (MAE: 8.15 μg/m^3, PE: 1.35%, MSE: 130.50, R^2: 0.97) but moderately improved with the use of PR (Sensor 187 MAE: 5.80 μg/m^3, PE: 1.10%, MSE: 75.23, R^2: 0.98). Presumably, PR helped address some curvature in the relationship. However, both models underperformed compared to RF, which provided improved accuracy (Sensor 187 MAE: 3.01 μg/m^3, PE: 0.65%, MSE: 25.67, R^2: 0.99) across all concentrations, underscoring its robustness and capability in capturing complex sensor responses. It should be noted that while the R^2 was high for all three models across the entire concentration range, this might not be the best metric of performance. When considering other metrics, such as MAE, the improved accuracy of the RF model becomes more obvious. For example, when compared to the sensor LOD, the RF model consistently maintained a low MAE, whereas the LR and PR did not. It is important to note that the LOD will be sensor-specific, which can then be taken into account for future application of such calibration models to other sensor types.

BP analysis

Another significant finding was the identification of an optimal BP around 350 μg/m^3, signaling a shift from linear to nonlinear sensor response. While valuable, relying solely on this BP has limitations due to variability in aerosol composition and environmental conditions, potentially necessitating adaptive calibration methods. Adopting a secondary BP at 150 μg/m^3, informed by prior literature, facilitated clearer segmentation and targeted calibration strategies. It should also be noted that this BP in sensor response might vary between different LCS models.

Evaluation of the different calibration methods for the different concentration segments provided a better understanding of how concentration might affect the overall usability of each method. In Segment A (0 to 150 μg/m^3), all models showed adequate performance, but RF consistently yielded lower errors and higher accuracy (MAE: 0.75 to 2.50 μg/m^3, PE: 0.38 to 1.00%, R^2: 0.98 to 0.99). Segment B (150 to 348 μg/m^3) illustrated RF's advantage in capturing transitional nonlinearity, outperforming LR and PR (RF MAE: 0.87 to 2.43 μg/m^3, PE: 0.38 to 0.99%, R^2: 0.996 to 0.999). In Segment C (>348 μg/m^3), RF maintained robust performance even under significant nonlinearities (MAE: 1.82 to 3.69 μg/m^3, PE: 0.34 to 0.74%, R^2: up to 0.999), highlighting LR and PR's inadequacies in highly variable conditions.

Across all segments, RF demonstrated the most reliable performance and robustness, adapting effectively to both linear and nonlinear sensor behaviors. EPA guidelines recommend sensor calibration models achieve a slope of 0.75 to 1.25, intercept within ±5 μg/m^3, $R^2 \geq 0.7$, and MAE ≤ 10 μg/m^3 (EPA 2021). It could also be argued that maintaining the MAE below the sensor-specific LOD is a robust way to ensure accuracy. In Segment A, all models generally met these criteria, but RF performed best. In Segment B, RF and PR still met the guidelines, while LR started to exceed the MAE threshold. In Segment C, only RF consistently stayed within all EPA-recommended ranges. These results support using RF for both ambient and occupational applications, especially at higher PM2.5 levels.

Implications for calibration in occupational environments

The findings of this study highlight the varying effectiveness of different calibration models across a wider concentration range than previously reported, emphasizing the need for more advanced methods in high-concentration occupational environments. While LR models performed well at lower concentrations (0 to 150 μg/m^3), their accuracy declined as levels increased. This was especially evident in Segment C (> 348 μg/m^3), where LR models exhibited greater errors and lower R^2 values. For example, Sensor 187 showed a slope of 2.00 and an intercept of −275.58, indicating a deviation from linearity. These limitations suggest that LR models are only suitable for environments where concentrations remain within the lower linear response range of the sensors, ie, below 150 μg/m^3.

PR models provided an intermediate level of accuracy, improving upon LR but still underperforming compared to RF. PR models captured some nonlinear behavior, but their performance declined at higher concentrations, particularly in Segment C. While PR can be a useful alternative when computational simplicity is needed, it does not provide the same level of reliability and adaptability as RF. In industrial hygiene applications, where environmental conditions are highly dynamic, RF remains the superior choice.

In occupational settings such as industrial sites, construction areas, and mining operations, aerosol concentrations frequently exceed 150 μg/m^3, often reaching

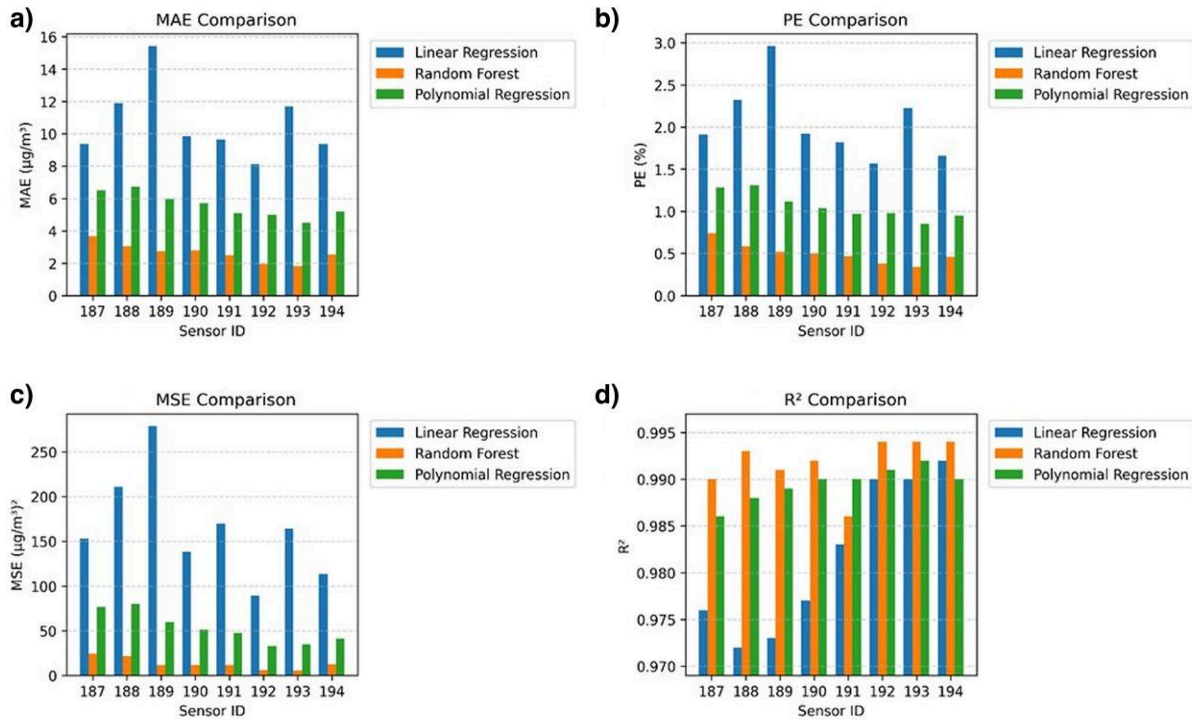

Figure 4 Performance comparison of calibration models in Segment C. (a) MAE, (b) PE, (c) MSE, and (d) coefficient of determination (R^2) for different models and sensors.

several hundred micrograms per cubic meter (Balanay and Lungu 2009; Madueño et al. 2019). Given these extreme conditions, relying solely on LR models could lead to substantial underestimation or overestimation of actual concentrations, potentially compromising exposure assessments and worker protection efforts. Robust and reliable data is critical for industrial hygiene professionals who require accurate real-time exposure assessments to inform protective measures (Jaiswal and Samikannu 2017; Wang et al. 2019). Therefore, if LCS are to continue being used in such environments, calibration models that can account for the nonlinear behavior of sensors at high concentrations are required.

Future directions and practical implications

Recent work by Aix et al. (2023) proposed a structured calibration protocol for low-cost PM sensors that integrates data preprocessing, model tuning, and machine learning approaches, and showed that RF and multilinear regression outperform traditional linear methods for PM1 and PM2.5. Our findings are consistent with this framework in demonstrating superior performance of RF relative to linear and polynomial models. Importantly, we extend this prior work by showing that calibration performance is strongly concentration-dependent, and that a single global model may be insufficient across low, transitional, and high PM2.5 ranges. The concentration-segmented strategy presented here therefore complements the protocol of Aix et al. and provides additional guidance for high-concentration occupational environments (Aix et al. 2023). Our findings align with and extend recent best-practice recommendations for calibrating low-cost PM sensors. Frameworks such as those by Giordano et al. and Liang and Daniels emphasize careful model selection, cross-validation, and transparent reporting of calibration performance across different concentration regimes. In the present study, we applied these principles in a high-concentration chamber context and further introduced a concentration-segmented strategy that explicitly distinguishes low (0 to 150 μg/m^3), transitional (150 to 348 μg/m^3), and high (>348 μg/m^3) ranges. This approach highlights that calibration performance, particularly for simple linear models, can deteriorate markedly at higher concentrations, and that nonlinear machine learning approaches such as RF are better suited to occupational settings where PM2.5 concentration frequently exceeds several hundred μg/m^3.

These findings reinforce the need to adopt more advanced calibration techniques, particularly in occupational settings with high and variable aerosol levels. Future research should focus on optimizing machine learning-based calibration approaches, such as expanding RF models with additional environmental parameters (eg temperature, humidity) to further refine predictions. Additionally, field validation studies in real-world occupational environments are necessary to confirm the applicability of RF models beyond controlled laboratory conditions.

From a practical perspective, the use of RF models can enhance workplace safety by enabling more accurate and cost-effective air quality monitoring using LCS. This would allow for better regulatory compliance, improved exposure assessments, and more effective interventions to reduce worker health risks. Given the increasing reliance on LCS for air quality monitoring, integrating RF-based calibration approaches represents a critical step forward in industrial hygiene and occupational health research.

However, implementing RF models in practice may require specialized knowledge in machine learning and access to greater computing resources than traditional linear models. These requirements could pose challenges for smaller organizations or in settings with limited technical capacity, highlighting the need for the development of user-friendly tools and training to support broader adoption. Additionally, the complexity and computational intensity of RF models could pose challenges for real-time deployment in resource-constrained settings.

Study limitations

While this study advances LCS calibration for high concentrations, several limitations remain. First, results derive from controlled chamber tests and should be validated under more variable field conditions (Jovašević-Stojanović et al. 2015). Second, using ammonium nitrate as the challenge aerosol—chosen for local relevance—may limit generalizability to sites with different aerosol composition. Third, findings are specific to Plantower PMS sensors, which are less sensitive to particles >1 μm; performance can vary with particle size distributions and deployment context, and should be verified for other LCS types (Ouimette et al. 2022; Kaur and Kelly 2023). Fourth, we did not systematically vary temperature or RH, and light-scattering LCS are sensitive to both, so condition-specific corrections may be needed. Fifth, the chamber reference has ~±10% accuracy, and reported model performance may partly reflect overfitting to chamber conditions. Sixth, analysis was restricted to PM2.5 because the DustTrak 8530 was operated with a PM2.5 impactor and the chamber aerosol was dominated by small (<2.5 μm) particles. As a result, PM10 output or other coarse/inhalable fractions could not be meaningfully evaluated. Future chamber and field studies should explicitly assess the PMS3003 PM10 channel using appropriate PM10 or inhalable reference samplers and aerosols with substantial coarse particle fractions, particularly in occupational settings where coarse particles may dominate. Finally, PM2.5 is not an occupational exposure standard; future work should assess LCS performance for occupationally relevant size fractions (total, inhalable, respirable) and consider co-occurring gaseous pollutants.

Conclusion

In conclusion, while LR models were effective for calibrating LCS in low concentration environments, occupational settings with higher concentrations require more advanced calibration methods. The RF model provided the most accurate calibration method across all concentration ranges in this laboratory experiment. The insights from this study can be used to enhance the development of effective monitoring solutions tailored to the specific needs of different environments.

Acknowledgments

The authors would like to acknowledge the contributions of the research team and laboratory personnel in Dr. Kelly's laboratory, whose support in the setup and operation of the calibration experiments was instrumental to the successful completion of this study. We also thank all team members involved in data collection and analysis for their valuable assistance throughout the project. This work was supported in part by a training grant from the US National Institute for Occupational Safety and Health (T42OH008414).

Supplementary material

Supplementary material is available at *Annals of Work Exposures and Health* online.

Funding

This research received no external funding.

Conflicts of interest

The authors declare no conflict of interest.

Data availability

All data supporting the results of this study are provided in the Supplementary Materials.

References

Aix M-L, Schmitz S, Bicout DJ. 2023. Calibration methodology of low-cost sensors for high-quality monitoring of fine particulate matter. Sci Total Environ. 889. https://doi.org/10.1016/j.scitotenv.2023.164063.

Balanay JA, Lungu C. 2009. Exposure of jeepney drivers in Manila, Philippines, to selected volatile organic compounds (VOCs). Ind Health. 47:33–42. https://doi.org/10.2486/indhealth.47.33.

Barkjohn KK, Gantt B, Clements AL. 2021. Development and application of a United States-wide correction for $PM_{2.5}$ data collected with the PurpleAir sensor. Atmos Meas Tech. 14:4617–4637. https://doi.org/10.5194/amt-14-4617-2021.

EPA. 2021. *Air sensor performance targets recommendations for PM2.5 and ozone*. EPA/600/R-21/003. EPA Office of Research and Development.

Fang R et al. 2025. Optimizing air quality monitoring: comparative analysis of linear regression and machine learning in low-cost sensor calibration. Aerosol Air Qual Res. 25. https://doi.org/10.1007/s44408-025-00009-x.

Giordano MR et al. 2021. 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. 158. https://doi.org/10.1016/j.jaerosci.2021.105833.

Hegde S et al. 2020. Indoor household particulate matter measurements using a network of low-cost sensors. Aerosol Air Qual Res. 20:381–394. https://doi.org/10.4209/aaqr.2019.01.0046.

Jaiswal JK, Samikannu R. 2017. Application of random forest algorithm on feature subset selection and classification and regression. 2017 World Congress on Computing and Communication Technologies (WCCCT), IEEE: Tiruchirappalli, India.

Jovašević-Stojanović M et al. 2015. On the use of small and cheaper sensors and devices for indicative citizen-based monitoring of respirable particulate matter. Environ Pollut. 206:696–704. https://doi.org/10.1016/j.envpol.2015.08.035.

Kaur K, Kelly KE. 2023. Performance evaluation of the Alphasense OPC-N3 and Plantower PMS5003 sensor in measuring dust events in the Salt Lake Valley, Utah. Atmos Meas Tech. 16:2455–2470. https://doi.org/10.5194/amt-16-2455-2023.

Kelly KE et al. 2017. Ambient and laboratory evaluation of a low-cost particulate matter sensor. Environ Pollut. 221:491–500. https://doi.org/10.1016/j.envpol.2016.12.039.

Kuula J et al. 2020. Laboratory evaluation of particle-size selectivity of optical low-cost particulate matter sensors. Atmos Meas Tech. 13:2413–2423. https://doi.org/10.5194/amt-13-2413-2020.

Levy Zamora M et al. 2019. Field and laboratory evaluations of the low-cost plantower particulate matter sensor. Environ Sci Technol. 53:838–849. https://doi.org/10.1021/acs.est.8b05174.

Liang L. 2021. Calibrating low-cost sensors for ambient air monitoring: techniques, trends, and challenges. Environ Res. 197. https://doi.org/10.1016/j.envres.2021.111163.

Liang L, Daniels J. 2022. What influences low-cost sensor data calibration?—A systematic assessment of algorithms, duration, and predictor selection. Aerosol Air Qual Res. 22. https://doi.org/10.4209/aaqr.220076.

Madueño L et al. 2019. A new method to measure real-world respiratory tract deposition of inhaled ambient black carbon. Environ Pollut. 248:295–303. https://doi.org/10.1016/j.envpol.2019.02.021.

Masri S, Rea J, Wu J. 2022. Use of low-cost sensors to characterize occupational exposure to PM2.5 concentrations inside an industrial facility in Santa Ana, CA: results from a worker- and community-led pilot study. Atmosphere (Basel). 13. https://doi.org/10.3390/atmos13050722.

OSHA. 2012. *Particulates not otherwise regulated, total and respirable dust (PNOR)*. [accessed 2026 Mar 30]. https://www.osha.gov/chemicaldata/801

Ouimette JR et al. 2022. Evaluating the PurpleAir monitor as an aerosol light scattering instrument. Atmos Meas Tech. 15:655–676. https://doi.org/10.5194/amt-15-655-2022.

Pope CA, III et al. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA. 287:1132–1141. https://doi.org/10.1001/jama.287.9.1132.

Russell HS et al. 2022. Enhanced ambient sensing environment-A new method for calibrating low-cost gas sensors. Sensors (Basel). 22. https://doi.org/10.3390/s22197238.

Sayahi T et al. 2019. Development of a calibration chamber to evaluate the performance of low-cost particulate matter sensors. Environ Pollut. 255. https://doi.org/10.1016/j.envpol.2019.113131.

Shezi B et al. 2020. Occupational exposure to fine particulate matter (PM(4) and PM(2.5)) during hand-made cookware operation: personal, indoor and outdoor levels. Int J Environ Res Public Health. 17. https://doi.org/10.3390/ijerph17207522.

Vajs I, Drajic D, Cica Z. 2023. Data-driven machine learning calibration propagation in A hybrid sensor network for air quality monitoring. Sensors. 23. https://doi.org/10.3390/s23052815.

Wang Z, Delp WW, Singer BC. 2020. Performance of low-cost indoor air quality monitors for $PM_{2.5}$ and PM_{10} from residential sources. Build Environ. 171:106654. https://doi.org/10.1016/j.buildenv.2020.106654.

Wang Y, Du Y, Wang J, Li T. 2019. Calibration of a low-cost PM2. 5 monitor using a random forest model. Environ Int. 133. https://doi.org/10.1016/j.envint.2019.105161.

Williams R, Kaufman A, Hanley T, Rice J, Garvey S. 2015. *Evaluation of field-deployed low cost PM sensors*. Research Triangle Park, NC: U.S. Environmental Protection Agency, Office of Research and Development, National Exposure Research Laboratory.

Zimmerman N et al. 2018. A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring. Atmos Meas Tech. 11:291–313. https://doi.org/10.5194/amt-11-291-2018.

Zou Y, Clark JD, May AA. 2021. A systematic investigation on the effects of temperature and relative humidity on the performance of eight low-cost particle sensors and devices. J Aerosol Sci. 152. https://doi.org/10.1016/j.jaerosci.2020.105715.