https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

Evaluation of Calibration Performance of a Low-cost Particulate Matter Sensor Using Colocated and Distant NO_2

[1], Kabseok Ko[1], Seokheon Cho[2], and Ramesh R. Rao[2]

[1]Department of Electronics Engineering, Kangwon National University, Chuncheon, 24341, Korea
[2]Qualcomm Institute, University of California, San Diego (UCSD), La Jolla, CA, 92093, USA

Correspondence: Seokheon Cho (justinshcho@gmail.com)

Abstract. Low-cost optical particle sensors have the potential to supplement existing particulate matter (PM) monitoring systems to provide high spatial and temporal resolution. However, low-cost PM sensors have often shown questionable performance under various ambient conditions. Temperature, relative humidity (RH), and particle composition have been identified as factors that directly affect the performance of low-cost PM sensors. This study investigated if NO_2, which creates $PM_{2.5}$ by chemical reactions in the atmosphere, can be used to improve the calibration performance of low-cost $PM_{2.5}$ sensors. To this end, we evaluated the PurpleAir PA-II, called PA-II, a popular air monitoring system that utilizes two low-cost PM sensors that is frequently deployed near air quality monitoring sites of the Environmental Protection Agency (EPA). We selected a single location where 14 PA-II units have operated for more than two years since July 2017. Based on the operating periods of the PA-II units, we then chose the period of Jan. 2018 to Dec. 2019 for study. Among the 14 units, a single unit containing more than 23 months of measurement data with a high correlation between the unit's two PMS sensors was selected for analysis. Daily and hourly $PM_{2.5}$ measurement data from the PA-II unit and a BAM 1020 instrument, respectively, were compared using the federal reference method (FRM), and a per-month analysis was conducted against the BAM-1020 using hourly $PM_{2.5}$ data. In the per-month analysis, three key features, temperature, relative humidity (RH), and NO_2, were considered. The NO_2, called colocated NO_2, was collected from the reliable instrument colocated with the PA-II unit. The per-month analysis showed the PA-II unit had a good correlation (coefficient of determination, $R^2 > 0.819$) with the BAM-1020 during the months of Nov., Dec., and Jan. in both 2018 and 2019, but their correlation intensity was moderate during other months, such as July and Sep. 2018, and Aug., Sep., and Oct. 2019. NO_2 was shown to be a key factor in increasing the value of R^2 in the months when moderate correlation based on only $PM_{2.5}$ was achieved. This study calibrated a PA-II unit using multiple linear regression (MLR) and random forest (RF) methods based on the same three features used in the analysis studies as well as their multiplicative terms. The addition of NO_2 had a much larger effect than that of RH when both $PM_{2.5}$ and temperature were considered for calibration in both models. When NO_2, temperature, and relative humidity were considered, the MLR method achieved similar calibration performance to the RF method. Since it is practically infeasible to colocate a reliable NO_2 instrument colocation with high accuracy at low-cost PM sensors, we investigated the effectiveness of using NO_2 data (which we call distant NO_2), collected from monitoring sites deployed at locations far from the considered low-cost PM sensor for calibration performance enhancement. It was shown that the use of distant NO_2 enhances the calibration performance compared to calibration without NO_2 when it is highly correlated with colocated NO_2. Overall, PA-II units have good agreement with $PM_{2.5}$ monitoring

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

systems of high quality. Moreover, the calibration performance can be improved by using machine learning algorithms and by considering temperature, RH, and especially NO_2.

1 Introduction

Recently, attention has been paid to particulate matter (PM), which not only has adverse effects on visibility but also can impact human health by contributing to conditions such as cardiovascular disease, asthma, and lung cancer (Liu et al., 2018, 2013). PM that is less than $2.5\,\mu m$ in diameter, referred to as $PM_{2.5}$, can penetrate the lungs and may thus increase the risk to human health. Globally, the estimated number of adult deaths attributable to $PM_{2.5}$ exposure is over 0.67, 1.6, and 2.1 million for lung cancer, cardiopulmonary disease, and all causes, respectively (Evans et al., 2013). To minimize the harmful effects, many countries regulate daily and annual $PM_{2.5}$ concentrations by monitoring $PM_{2.5}$ levels at air quality monitoring stations. The monitoring stations use Federal Reference Method (FRM) and/or Federal Equivalent Method (FEM) instruments, which have high precision and accuracy. These instruments can provide high quality measurements of $PM_{2.5}$ concentrations at the installed locations and nearby surroundings. However, these instruments are sparsely distributed due to the high cost of the equipment (ten thousand to tens of thousands of US dollars), so they cannot provide spatial variability. In other words, traditional monitoring stations frequently provide air quality data with poor spatio-temporal resolution, due to the limited number of high quality instruments.

As a cost-effective approach for a dense monitoring network, many stakeholders and researchers have turned to low-cost PM sensors that use a light scattering technique for measurement. In addition to low cost, these sensors have the advantages of low energy consumption and high sampling frequency, and they are easy to deploy and operate compared to traditional monitoring networks. Thus, low-cost PM sensors have been deployed in several communities to measure and report local air quality information (Jiao et al., 2016; PurpleAir, 2018).

Low-cost PM sensors are not suitable for regulatory purposes, however, because the data reported can be questionable in terms of accuracy, precision, and reliability. In worst case scenarios, low-cost sensors report no meaningful data all. Because manufacturers provide limited information on sensors' performance, some studies have been conducted to evaluate the performance of a variety of low-cost sensor models by comparing them with high-cost instruments in laboratory and outdoor ambient environments (Alvarado et al., 2015; Johnson et al., 2018; Wang et al., 2015; Holstius et al., 2014; Austin et al., 2015; Gao et al., 2015; Kelly et al., 2017; Mukherjee et al., 2017; Sousan et al., 2016; Feinberg et al., 2018; Crilley et al., 2018; Badura et al., 2018; Budde et al., 2018; Liu et al., 2019; Cavaliere et al., 2015; Kelly et al., 2017; Zheng et al., 2018). Most sensors showed good performance under laboratory tests where the sensors measured, known concentrations of particles, such as polystyrene latex, in a chamber. On the other hand, under ambient conditions, the performance of low cost sensors varied depending on the sensor model and its deployed location. Some PM sensor units have inconsistent precision between units of the same model (Feenstra et al., 2019; Feinberg et al., 2018), while other PM monitors, including the PurpleAir PA-II, have shown good precision (Barkjohn et al., 2020; Pawar and Sinha, 2020; Mailings et al., 2020). Field evaluations of PurpleAir PA-II units colocated with FEM instruments for approximately two months shown good correlation with the FEM instruments

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

(SCAQMD, 2017c). Furthermore, it was shown that PMS5003 sensors, which are used in PurpleAir PA-II monitors, have good a correlation with the FEM monitors (Kelly et al., 2017; Sayahi et al., 2019). However, the sensors still require calibration for better performance before use in ambient conditions.

Several studies have developed calibration models for low-cost PM sensors based on the following approaches: simple linear regression (Zheng et al., 2018), multiple linear regression (Zimmerman et al., 2018), random forest (Zimmerman et al., 2018), and neural networks (Si et al., 2020). Moreover, to improve calibration performance, several studies have identified other factors in addition to $PM_{2.5}$ concentration that can affect the performance of low-cost sensors. These typical factors include temperature, relative humidity, and particle properties (composition and size distribution) (Holstius et al., 2014; Gao et al., 2015; Kelly et al., 2017). In particular, some low-cost PM sensors have been shown to excessively overestimate $PM_{2.5}$ concentrations under high relative humidity conditions (Jayaratne et al., 2018). The reason for this overestimation is that some aerosols can uptake water via hygroscopy. To solve this problem, several correction models have been proposed, such as a correction model based on the κ-Köhler theory (Crilley et al., 2018, 2020), multiple linear regression (Barkjohn et al., 2021; Nilson et al., 2022), and generalized additive model (Hua et al., 2021). Analysis of direct factors, such as temperature, relative humidity, and particle composition, can enhance the performance of low-cost sensors. In addition to these direct factors, we examine in this study the impact of secondary formation by precursor gases, particularly NO_2, as a source of $PM_{2.5}$ emissions. This study aims to identify the significance of the precursor NO_2 and evaluate its potential for improving the performance of low-cost $PM_{2.5}$ sensors. To this end, we used a Multiple Linear Regression (MLR) model to evaluate the effect of meteorological factors and NO_2 on the performance of low-cost sensors compared to a BAM-1020 instrument. For the evaluation, we performed a per-month analysis on hourly $PM_{2.5}$ measurement data and considered the effect of several feature vectors, that are a combination of temperature, relative humidity, and NO_2. Based on the per-month analysis, we then considered two machine learning methods, MLR and Random Forest (RF), for calibration models. The trained MLR and RF models were evaluated on the test set, and their performance was compared. From an implementable perspective on NO_2 data, we investigated the feasibility of using data from far NO_2 regulatory instruments due to the questionable data quality of low-cost NO_2 sensors. The results of our study showed that incorporating far NO_2, in addition to temperature and relative humidity, into RF models yields lower errors than RF models that only include temperature and relative humidity.

2 Method

2.1 PurpleAir PA-II Units

The PurpleAir PA-II Outdoor air quality monitor was developed for measuring particulate matter of various sizes. PA-II units can measure various particulate matter as well as temperature, relative humidity, and barometric pressure. PurpleAir also developed a crowdsourcing platform to share publicly gathered PM measurements obtained from all PA units. From the PurpleAir website (https://www.purpleair.com/map), we can observe and download data reported by all installed PA units.

A PA-II unit includes two identical PMS 5003 sensors. The PMS 5003 sensors based on a light scattering principle measure concentrations of $PM_{1.0}$, $PM_{2.5}$, and PM_{10} in real-time. By counting the number of particles per each diameter range that flows

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

through a fan at a rate of 0.1L/min. Based on number of particles counted per diameter, each sensor estimates $PM_{1.0}$, $PM_{2.5}$, and PM_{10} concentrations and then averages the concentrations every 80 s [1]. The PA-II unit sends the averaged concentrations obtained from two PMS sensors (A and B) to the PurpleAir server without storing the data in the unit itself. The PA-II unit does not calibrate the data, which implies it just collects the measured data.

The PurpleAir website provides the following information about all PA-II units via a JSON formated file: a name, a unique ID, a latitude, a longitude, and an installation date. Each PA-II unit has two unique IDs for each of its PMS sensors A and B.

2.2 Air quality measurement data from EPA

Outdoor air quality data collected from across the US is publicly available through the U.S. Environmental Protection Agency (EPA) website (https://epa.gov/outdoor-air-quality-data). The EPA has a description file for monitors, which includes state code, county code, site number, location (latitude and longitude), parameter code, parameter occurrence code (POC), and last method. A combination of state code, county code, and site number can uniquely identify a monitoring site. For example, a monitoring station located at Bakersfield, CA has a state code of 06, a county code of 029, and a site number of 0014. The parameter code is an air quality system (AQS) code corresponding to the parameter measured by a monitor. For example, parameters regarding $PM_{2.5}$ and NO_2 are 88101 and 42602, respectively. A POC is used to identify an instrument among multiple ones with the same parameter code at a site. For example, two FRM instruments with a parameter of 88101 at the Bakersfield site are used to measure daily $PM_{2.5}$ concentrations and are identified with POC 1 and 2. The last method descriptor describes the measurement scheme used by the monitor for its most recent sample.

Hourly measurements of PM2.5, as well as other pollutants such as CO, NO2, SO2, and O3, obtained from FEM and non-FEM instruments can be downloaded via the EPA's application programming interface (https://aqs.epa.gov/data/api). Daily measurements of PM2.5 obtained from an FRM instrument are also available.

2.3 Selection of PA-II units and reference monitoring sites

To investigate the performance of a PA-II unit itself and evaluate its calibration, we focused on PA-II units that are installed close to an EPA monitoring site (i.e. reference site) that provides reliable hourly $PM_{2.5}$ concentrations. We use the location information of the PA-II units and reference monitors to find PA-II and reference monitor pairs that are located less than 100 m from each other (Wallace et al. (2021)). Among the identified pairs, we selected a monitoring site that has multiple PA-II units as pairs and can measure other pollutants such as NO_2 on an hourly basis.

2.4 Data Preprocessing of PA-II units

The PA-II units selected for study are long-term installations, i.e., they have been in operation for more than two years. Therefore, data from the unit's PM sensors may be abnormal due to failure and aging drift, data quality control is required before calibrating the PA-II sensors.

[1] After May 30, 2019, the averaging time is changed from 80 s to 120 s.

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

In this study, we used a correlation method to identify PA-II units with abnormal operations. Under normal operation, a PA-II unit's two PM sensors, A and B, are highly correlated since they are colocated in the same unit and have good precision. The correlation of sensors A and B in selected PA-II units was calculated based on the hourly $PM_{2.5}$ measurement data from each sensor without any data quality control. PA-II units that had a correlation larger than 0.99 between sensors A and B were selected, and a quality control (QC) measure developed by U.S. EAP for obtaining daily $PM_{2.5}$ was performed. The QC measure has been shown to be important for developing correction models of PA-II units (Barkjohn et al. (2021)). The QC measure has the following 3 steps: i) data from both channels A and B was removed when either channel A or B had a missing value, ii) data with abnormal temperature or relative humidity values was removed, and iii) data from channels A and B were compared. In the first step, when we calculate 1-hour averages of $PM_{2.5}$ measurement generated with 2 min (80 sec) intervals, we remove 1-hour average data if the number of $PM_{2.5}$ measurements is less than 27 (40). After calculating 1-hour average data, we removed all data points for the 1-hour interval, where either sensor A or B had a missing value. The second step deals with temperature and RH data. PA-II units occasionally report extremely high or low values of temperature and relative humidity that are inaccurate. Therefore, we removed the data points whose corresponding time interval contained unrealistic measurement of temperature or relative humidity. In this study, the acceptable ranges of temperature and RH are (0 °F, 140 °F) and (0%, 100%), respectively. Once the unacceptable data points were removed, we calculated the 1-hour average for temperature and RH. The last step was to compare results sensors A and B in a PA unit to check data consistency. To do this, we used symmetric percentage error (SPE) as follows:

$$SPE = \frac{2(|PM_{2.5}^{A}| - |PM_{2.5}^{B}|)}{|PM_{2.5}^{A} + PM_{2.5}^{B}|}, \tag{1}$$

where $PM_{2.5}^{A}$ and $PM_{2.5}^{B}$ are hourly averaged $PM_{2.5}$ concentrations from sensors A and B in the same PA-II unit, respectively. We removed the relevant data points with SPE larger than 0.61, which is 2 standard deviation. This value of SPE threshold has been used for 24-hr average $PM_{2.5}$ concentrations (Barkjohn et al. (2021)), but we use it here for 1-hour averaged $PM_{2.5}$ concentrations.

2.5 Calibration Methods

Two calibration methods, multiple linear regression (MLR) and random forest (RF), were evaluated. For both calibration methods, we considered various features, including $PM_{2.5}$ measured from a PA-II unit, temperature, relative humidity, NO_2, and their multiplicative interaction terms.

2.5.1 Multiple Linear Regression (MLR)

An MLR method can be expressed as follows:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n, \tag{2}$$

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

where $\hat{y}$ represents a response, n is the number of predictor variables, β_i for $i = 0, 1, \ldots, n$ are regression coefficients, and x_i for $i = 1, 2, \ldots, n$ represent predictor variables (called features). Using a linear equation with multiple variables, we investigated the relationship between features and a response.

All features in an MLR method should be independent. However, many studies have considered $PM_{2.5}$, temperature, and RH, which are not independent (Magi et al. (2019); Mailings et al. (2020)). Some studies have introduced multiplicative interaction terms (i.e., $PM_{2.5}$×RH) to exploit interdependence between features (Barkjohn et al. (2021)). We also consider multiplicative interaction terms in this study.

We use $PM_{2.5}$ concentrations obtained from a reference monitor as the response. As predictor variables, we consider multiple features, such as $PM_{2.5}$ measurement data from a PA-II unit, temperature, relative humidity, NO_2, and their multiplicative interaction terms (i.e., $PM_{2.5}$×RH, T×RH, $PM_{2.5}$×RH×T).

2.5.2 Random Forecast (RF)

An RF is an ensemble of K regression trees. Each regression tree is trained with a bootstrap sample of an original training dataset. The output of an RF is the aggregation of regression trees, i.e., averaging estimates over all trees. Each regression tree is grown by selecting random m features among M input features at each possible split. The best cut is calculated at the randomly chosen features. Optimal cuts can be achieved using the Classification and Regression Trees split criterion (CART), which compares the variance of the uncut node and one of all possible cuts along m directions. Every tree is fully grown with these splits (Breiman, 2001).

3 Results and Discussions

3.1 Selection of PA-II units and reference monitoring sites

We found one reference monitoring site at Rubidoux, CA whose pair has 14 PA-II units. The monitoring site is identified by a state code of 06, a county code of 065, and a site number of 8001 (i.e., 06-065-8001). This monitoring site is located in an urban residential area within the south coast air basin at an elevation of 248 m. Air pollutants from the Los Angeles and coastal areas are transported to this air basin, which is known to have poor ventilation and may experience air stagnation during the early evening and early morning periods. Local air pollution includes NOx from diesel trucks, since the city of Jurupa Valley, which includes the community of Rubidoux, is a main transportation corridor for diesel trucks serving three air cargo terminals and the ports of Los Angeles and Long Beach.

The monitoring site has an FRM method instrument and a BAM-1020 instrument with the parameter of 88502, which produces hourly $PM_{2.5}$ measurement data. Table 1 describes information about the 14 PA-II units, such as their IDs, location (latitude and longitude), sensor name, start time of measurement, end time of measurement, and non-operating months[2]. While we present the ID for only PMS sensor A of each PA-II unit, the ID of PMS sensor B is the ID of PMS sensor A plus 1. The

[2]We define non-operating month as the month, when the number of days without the measurement data is larger than 10 days.

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

geographic information on 14 PA-II units and the monitoring site is shown in Figure S1. Distances between PA-II units and the monitoring site are shown in Table S1. The minimum and maximum distance between a PA-II unit and the monitoring site is less than 10 m and 100 m, respectively.

Based on the non-operating months of the PA-II units found, we selected an appropriate period of sample data from Jan. 2018 to Dec. 2019 (24 months). Among the 14 identified PA-II units, we chose several that had more than 23 months of valid measurement data during the period selected for study. The selected units are RIVR_Co-loc2, 3, 6, 7, and 8, which we call PA-II 2, 3, 6, 7, and 8, respectively.

Before using $PM_{2.5}$ data from the PA-II units, we checked the unit's data quality. We first calculated the correlation among the selected PA-II units, considering both PMS 5003 sensors for each PA-II unit for the correlation analysis. Since these PA-II units are closely located, $PM_{2.5}$ data should be highly correlated. Figure 1 shows the correlation results for all PMS 5003 sensors included in the PA-II units. The numbers on each axis represent the number of the selected PA-II units. Boxes to the left and right of each number indicate PMS sensors A and B for its corresponding PA-II unit, respectively. The PMS sensor A of PA-II unit 2, PMS sensors A and B of PA-II unit 5, and PMS sensor A of PA-II unit 6 all have a poor correlation with other PMS sensors. In addition, sensor A of PA-II unit 3 has slightly poor correlation with other sensors. Based on these results, we selected PA-II unit 7 for application of the QC measure.

3.2 Daily $PM_{2.5}$ Measurement Data

We collected daily $PM_{2.5}$ measurement data for 2018 and 2019 from an FRM method instrument at Rubidoux, CA. During that two-year period, the daily $PM_{2.5}$ levels ranged from a minimum value of 1.2 $\mu g/m^3$ to a maximum value of 66.3 $\mu g/m^3$, with an average of 11.69 $\mu g/m^3$ and a standard deviation of 6.88 $\mu g/m^3$. Using Met-One BAM-1020 W/SCC, daily $PM_{2.5}$ concentrations were calculated by averaging 1-hour $PM_{2.5}$ concentrations over 24 hours from 2018 to 2019. The maximum and minimum values of daily $PM_{2.5}$ concentrations were 68.3 and 0 $\mu g/m^3$, respectively, and the mean and standard deviation were 12.13 and 9.16 $\mu g/m^3$, respectively. These data suggest that the non-FEM method compares well to the statistics achieved with the FRM method.

The maximum, minimum, mean, and standard deviation values for daily $PM_{2.5}$ concentrations of the PA-II 7 unit were 129.069, 0.199, 18.247, and 13.854, respectively. Compared to the FRM and BAM-1020 instruments, the PA-II 7 unit overestimates the maximum daily PM2.5 concentrations. Additionally, the mean daily PM2.5 concentration from the PA-II 7 unit was higher than that of the FRM and BAM-1020 instruments.

In this study, we examined the root mean square error (RMSE), mean squared error (MSE), mean absolute error (MAE), and Pearson correlation coefficient r between the FRM instrument and the PA-II units for daily $PM_{2.5}$ data. These performance

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

metrics are expressed as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2}, \tag{3}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2, \tag{4}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - y_i|, \tag{5}$$

where x_i represents 1-hour averaged (24-hour period) sensor $PM_{2.5}$ concentrations for the ith hour (day) ($\mu g/m^3$), y_i represents 1-hour averaged (24-hour period) FRM or BAM-1020 $PM_{2.5}$ concentrations for the ith hour (day) ($\mu g/m^3$), and n is the number of data points. We matched the working period of the FRM instrument with that of the PA-II unit and then computed the above four performance metrics. The MSE, RMSE, MAE, and r for the selected PA-II unit are 102.565, 10.127, 7.156, and 0.928, respectively. These results show that the PA-II unit has a good correlation (r) with the FRM instrument for the two-year period of interest. However, a comparison of metrics from the FRM instrument and the BAM-1020 instrument did not correlate as favorably. When we evaluate the performance of the BAM-1020 against the FRM instrument, the BAM-1020 has MSE of 5.199 $\mu g/m^3$, RMSE of 2.280 $\mu g/m^3$, MAE of 1.685 $\mu g/m^3$, and r of 0.947.

The BAM-1020 instrument showed similar statistical results to that of the FRM method instrument. However, the measurements are not enough to evaluate how similar the performance of the BAM-1020 is to that of the FRM instrument. Hence, this study compared the performance of two instruments using a linear fitting scheme. Figure 2 shows the calibration performance using linear regression. The R^2, slope, and intercept are 0.896, 0.923, and 0.741, respectively. Also, the value of RMSE is 2.211. The BAM-1020 is close to an instrument with the parameter of 88101. In order for the BAM-1020 to attain the 88101 code in terms of performance, the following conditions must be satisfied: R^2 is larger than 0.9, a slope is larger than 0.9 and less than 1.1, and an absolute value of the intercept is less than 2.0. Slope and intercept are satisfied with the requirement, while R^2 is not. Nonetheless, the BAM-1020 instrument provides an acceptable level of performance to evaluate the calibration performance of PA-II units on an hourly basis.

3.3 Hourly $PM_{2.5}$ Measurement Data

Next, we compared the hourly $PM_{2.5}$ data of the PA-II unit with that of the BAM-1020 instrument over the course of the same two-year period. The maximum, minimum, mean, and standard deviation values for hourly PM2.5 concentration measurements of the BAM-1020 were 159.0, 0.0, 12.172, and 9.230 $\mu g/m^3$, respectively. The maximum, minimum, mean, and standard deviations values for hourly $PM_{2.5}$ of the selected PA-II unit were 263.062, 0.019, 18.367, and 17.610 $\mu g/m^3$, respectively. The PA-II unit's maximum hourly $PM_{2.5}$ measurement was almost twice that of the BAM-1020. In other words, the PA-II unit overestimates hourly $PM_{2.5}$ concentrations. Figure 3 shows the comparison of $PM_{2.5}$ measurement data obtained from the BAM-1020 and the selected PA-II 7 unit, as well as temperature and relative humidity measured from the selected PA-II 7 unit during winter season (from Dec. 2018 to Feb. 2019). The PA-II 7 unit showed a similar trend of $PM_{2.5}$ concentration

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

measurements to that of the BAM-1020 instrument, but it generally overestimated $PM_{2.5}$ concentrations more often than the BAM-1020.

In addition, we compared the hourly $PM_{2.5}$ concentrations of the PA-II unit with that of the BAM-1020 instrument in terms of RMSE, MSE, MAE, and r. The results are as follows: RMSE of 6.194 $\mu g/m^3$, MSE of 38.369 $\mu g/m^3$, MAE of 7.919 $\mu g/m^3$, and r of 0.876 $\mu g/m^3$. The PA-II unit had a good correlation with the BAM-1020 instrument based on r. However, other metrics, such as RMSE, MSE, and MAE, did not correlate well.

3.4 Per-month Analysis on Hourly $PM_{2.5}$ Measurement Data

We conducted a per-month analysis to evaluate the monthly performance based on hourly $PM_{2.5}$ data from the PA-II 7 unit compared to the BAM-1020 instrument. We did a linear fitting for the per-month analysis over two years. In other words, ordinary linear regression was applied to hourly $PM_{2.5}$ measurement data for each month. Table 2 shows the value of R^2, RMSE, and MAE of hourly $PM_{2.5}$ measurement data from the PA-II 7 unit compared to that of the BAM-1020 instrument and the corresponding slope and intercept of each optimal linear fitting. During the months of Nov., Dec., and Jan., the PA-II unit is shown to have a high correlation, R^2 of 0.813 to 0.936, with the BAM-1020 instrument. This result is supported by the field evaluation of PA-II units taken by the Air Quality Sensor Performance Evaluation Center (AQ-SPEC) during the period of Dec. 2016 - Jan. 2017, which showed the value of R^2 as being 0.868 to 0.921 when the PA-II units were compared with the FEM. (Sayahi et al., 2019) showed that PMS sensors have a high correlation with tapered element oscillating microbalances (TEOM) instruments during the winter season by providing R^2 of 0.866 to 0.892. That is, the hourly $PM_{2.5}$ measurement data from PA-II units seem to be highly correlated with that of FEM instruments during the months of November, December, and January, which implies the $PM_{2.5}$ measurement performance of PA-II is reliable, especially during winter seasons. These months have different slopes and intercepts; for example, Jan. 2018 has a slope of 0.502 and an intercept of 3.898, while Jan. 2019 has 0.397 and 1.961, respectively.

On the other hand, the PA-II 7 unit has a correlation lower than 0.6 for months of Jul. and Sep. 2018 as well as Aug., Sep., and Oct. 2019. These months, except Sep. 2019, have larger RMSE values compared to other months over the two-year period, which need to be calibrated.

3.5 Effect of Temperature, Relative Humidity, and NO_2

In the per-month analysis of PA-II units compared to the BAM-1020 instrument, non-winter seasons showed low correlations. In previous studies, temperature and relative humidity have been considered as key factors for effective calibration. In particular, relative humidity has been shown to affect low-cost PM sensors under high relative humidity conditions. Therefore, we investigated the effect of temperature and relative humidity on the fitting performance of PA-II units. In addition to temperature and relative humidity, we considered another factor, NO_2, which is known to be a precursor to the creation of $PM_{2.5}$ via chemical reactions in the atmosphere. NO_2 may indirectly affect $PM_{2.5}$ concentrations, some papers have considered NO_2 in calibration models (Hua et al., 2021). Thus, we investigated the possibility of improving the calibration performance of low-cost PM sensors by including NO_2 as a new feature.

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

For multiple features, such as temperature, relative humidity, and NO_2, we used an MLR approach for regression analysis of PA-II units compared to the BAM-1020 instrument. A per-month analysis was conducted based on hourly $PM_{2.5}$ measurements from the PA-II 7 unit under several feature vectors, such as ($PM_{2.5}$), ($PM_{2.5}$, T), ($PM_{2.5}$, RH), ($PM_{2.5}$, NO_2), ($PM_{2.5}$, T, RH), and ($PM_{2.5}$, T, NO_2), where T and RH represent temperature and relative humidity, respectively. For notational simplicity, we defined the above feature vectors ($PM_{2.5}$), ($PM_{2.5}$, T), ($PM_{2.5}$, RH), ($PM_{2.5}$, NO_2), ($PM_{2.5}$, T, RH), and ($PM_{2.5}$, T, NO_2) as 1, 2, 3, 4, 5, and 6, respectively. Figure 4 shows the R^2 and RMSE results of multiple linear regression for selected months with the above varying feature vectors. We considered feature vector 1 as a baseline for comparison among other feature vectors. On Jan. 2018, feature vector 5, referring to temperature and relative humidity, had little effect on regression performance of R^2 and RMSE. The amount of R^2 increase by feature vector 5 from the baseline was around 0.001, and the amount of RMSE decrease was 0.038 $\mu g/m^3$. In the case of feature vector 6, including NO_2 instead of RH, R^2 increased from the baseline by 0.015, while RMSE was improved by 0.518 $\mu g/m^3$. Similarly, for Apr. 2018, R^2 (or RMSE) for feature vector 5 increased (or decreased) by 0.01 (or 0.072 $\mu g/m^3$) compared to its baseline. R^2 and RMSE for feature vector 6 increase by 0.05 and decrease by 0.52 $\mu g/m^3$ from the baseline, respectively. For regressions on Aug. and Sep. 2019, an increase in R^2 was larger than 0.17 when feature vector 6 was considered, but it was less than 0.07 when feature vector 5 was considered. These remarkable results suggest that NO_2 is generally a key factor shown to be able to improve performance of PA-II units over an entire year, even though the enhanced performance for some months does not meet the same measurement data quality as the reference monitoring sites.

3.6 Calibration Performance

A per-month analysis with a combination of features, including T, RH, and NO_2, showed an effect on calibration for the PA-II unit. It is not simple to apply the per-month linear fitting result to calibrate PA-II units because month has a different slope and intercept defined for the linear fitting. Moreover, their values change according to year even for the same month. For example, notably, the linear fitting result in Apr. 2018 exhibited a higher RMSE than the fitting result in Apr. 2019. On the contrary, the calibration performance in Aug. 2018 was worse than that in Aug. 2019.

We used a machine learning approach to develop a calibration model, employing two machine learning algorithms, such as MLR and RF. The two-year dataset was divided into training and test sets at a 1:1 ratio, meaning the measurement data in the years 2018 and 2019 were used for training and testing, respectively. We used the training set to learn calibration models based on MLR and RF, and then used the test set to evaluate the calibration performance in terms of RMSE, MAE, and R^2. A calibration performance for the PA-II 7 unit using MLR and RF methods was compared with several features, including temperature, relative humidity, and NO_2, as well as their multiplicative terms.

3.6.1 Calibration by MLR

Recently, calibration methods have employed multiplicative interaction terms, such as $PM_{2.5}$×RH and T×RH. In our MLR models, we considered both additive and multiplicative interaction terms. The additive terms in our models include raw PurpleAir $PM_{2.5}$, T, RH, and NO_2. We considered multiplicative interaction terms that involve less than four additive terms when

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

NO_2 was not included (i.e., we consider $PM_{2.5}$×T×RH), and less than three additive terms when NO_2 is included. There are 95 combinations of features. Out of 95 combinations tested, only 43 combinations had a p-value of less than 0.05. Of those, we select 20 combinations, among 43 combinations, by increasing the number of additive terms and the number of multiplicative interaction terms and identifying the combinations with the lowest RMSE among the same numbers of additive terms and multiplicative interaction terms. The selected combinations were as follows: 1) $PM_{2.5}$, 2) $PM_{2.5}$ and T, 3) $PM_{2.5}$ and RH, 4) $PM_{2.5}$, RH, and $PM_{2.5}$×RH, 5) $PM_{2.5}$, T, and RH, 6) $PM_{2.5}$, T, RH, and $PM_{2.5}$×RH, 7) $PM_{2.5}$, T, RH, $PM_{2.5}$×T, and $PM_{2.5}$×RH, 8) $PM_{2.5}$, T, RH, $PM_{2.5}$×T, $PM_{2.5}$×RH, and T×RH, 9) $PM_{2.5}$, T, RH, $PM_{2.5}$×T, $PM_{2.5}$×RH, T×RH, and $PM_{2.5}$×T×RH, 10) $PM_{2.5}$ and NO_2, 11) $PM_{2.5}$, NO_2, and $PM_{2.5}$×NO_2, 12) $PM_{2.5}$, T, and NO_2, 13) $PM_{2.5}$, T, NO_2, and $PM_{2.5}$×NO_2, 14) $PM_{2.5}$, T, NO_2, $PM_{2.5}$×T, and $PM_{2.5}$×NO_2, 15) $PM_{2.5}$, T, NO_2, $PM_{2.5}$×T, $PM_{2.5}$×NO_2, and T×NO_2, 16) $PM_{2.5}$, T, RH, NO_2, and $PM_{2.5}$×RH, 17) $PM_{2.5}$, T, RH, NO_2, $PM_{2.5}$×RH, and $PM_{2.5}$×NO_2, 18) $PM_{2.5}$, T, RH, NO_2, $PM_{2.5}$×RH, $PM_{2.5}$×NO_2, and T×NO_2, 19) $PM_{2.5}$, T, RH, NO_2, $PM_{2.5}$×T, $PM_{2.5}$×RH, T×RH, and RH×NO_2, and 20) $PM_{2.5}$, T, RH, NO_2, $PM_{2.5}$×T, $PM_{2.5}$×RH, T×RH, T×NO_2, and RH×NO_2.

The calibration results of the PA-II 7 unit for test datasets using the MLR method with 20 selected combinations are presented in Table 3. Multicollinearity is a known issue with MLR models, as it can cause instability. One common method to diagnose this issue is to use the variance inflation factor (VIF) test for multicollinearity (Mansfield and Helms , 1982). Out of the 20 combinations tested, most VIF values were less than 5, indicating the absence of collinearity issues.

When a single additive term, such as T or RH, was applied, the RMSE values for two combinations, #2 and #3, improved by more than 0.22 $\mu g/m^3$, compared to considering only $PM_{2.5}$. The inclusion of an additive RH term in an MLR yielded a lower error than an additive T term did, since both RMSE and MAE for combination #3 were less than those for combination #2. The MLR model with $PM_{2.5}$, the single additive term with RH, and its multiplicative interaction term with $PM_{2.5}$ yielded higher RMSE and MAE than the MLR model using $PM_{2.5}$ and two meteorological variables, such as T and RH, as demonstrated by the results of combinations #4 and #5. When we considered two meteorological variables and incorporated four multiplicative interaction terms, such as $PM_{2.5}$×T, $PM_{2.5}$×RH, T×RH, and $PM_{2.5}$×T×RH, the MLR model resulted in the lowest error with an RMSE of 4.171 and an MAE of 3.048, compared to all combinations generated from $PM_{2.5}$, T, RH, and their multiplicative terms.

The MLR model of combination #10 with $PM_{2.5}$ and NO_2 had an RMSE of 4.445, which was lower than that of the MLR model with only $PM_{2.5}$, whose RMSE was 4.534, but larger than that of combination #2 with a single environmental variable and an RMSE of 4.306. This implies that the addition of a single multiplicative term in that model has no performance enhancement. However, when the additive term T is incorporated into an MLR model with $PM_{2.5}$ and NO_2, an RMSE of 3.992 can be achieved, which is lower than the values of all combination cases, not including NO_2, i.e., combinations #1 to #9. Coefficients of $PM_{2.5}$, T, and NO_2 in the MLR model, including T and NO_2, were around 0.446, 0.110, and 0.112, respectively. The temperature had more impact on error than relative humidity when considering NO_2. Considering both temperature and relative humidity together with NO_2 may cause a non-zero correlation of relative humidity with other factors due to a p-value of 0.083. When some multiplicative terms were additionally integrated into T, RH, and NO_2, the MLR calibration models passed a p-value test. The model based on combination #20 with four additive terms, i.e., $PM_{2.5}$, T, RH, and NO_2, and

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

multiplicative interaction terms, including $PM_{2.5}$×T, $PM_{2.5}$×RH, T×RH, T×NO_2, and RH×NO_2, achieved the lowest RMSE of 3.938. Considering multiplicative terms with T and RH had little effect on calibration performance as shown in the results of combinations #15, #19, and #20. From these results, we conclude that considering NO_2 together with meteorological variables and their multiplicative terms or a single variable, such as temperature, can improve the calibration performance of PA-II units.

3.6.2 Calibration by RF

This study validated performance of RF-based calibration for PA-II units with 95 combinations of predictor variables mentioned in the previous subsection. An RF was implemented using the scikit-learn package in Python. An RF has several hyperparameters, such as n_estimators, max_depth, min_samples_leaf, and max_features, that need to be set for the best performance over each combination of features. For this study, the hyperparameters were tuned with a random search method by 5-fold cross-validation based on the training set. For a random search, the number of trees (n_estimators) was set to 10, 20, 50, 100, 200, and 400. The range of max_depth was set to 2, 4, 6, 8, 10, 16, and None. The range of min_samples_leaf was set to 1, 2, 3, 4, and 5. The range of min_samples_split was set to 2, 3, 5, 7, and 10. The range of max_features was set to None.

We selected 22 combinations according to the above mentioned method. The selected combinations were as follows: 1) $PM_{2.5}$, 2) $PM_{2.5}$ and T, 3) $PM_{2.5}$ and RH, 4) $PM_{2.5}$, RH, and $PM_{2.5}$×RH, 5) $PM_{2.5}$, T, and RH, 6) $PM_{2.5}$, T, RH, and $PM_{2.5}$×T, 7) $PM_{2.5}$, T, RH, $PM_{2.5}$×T, and T×RH, 8) $PM_{2.5}$, T, RH, $PM_{2.5}$×T, $PM_{2.5}$×RH, and T×RH, 9) $PM_{2.5}$, T, RH, $PM_{2.5}$×T, $PM_{2.5}$×RH, T×RH, and $PM_{2.5}$×T×RH, 10) $PM_{2.5}$ and NO_2, 11) $PM_{2.5}$, NO_2, and $PM_{2.5}$×NO_2, 12) $PM_{2.5}$, RH, and NO_2, 13) $PM_{2.5}$, T, NO_2, and T×NO_2, 14) $PM_{2.5}$, RH, NO_2, $PM_{2.5}$×RH, and RH×NO_2, 15) $PM_{2.5}$, T, NO_2, $PM_{2.5}$×T, $PM_{2.5}$×NO_2, and T×NO_2, 16) $PM_{2.5}$, T, RH, and NO_2, 17) $PM_{2.5}$, T, RH, NO_2, and T×RH, 18) $PM_{2.5}$, T, RH, NO_2, $PM_{2.5}$×RH, and $PM_{2.5}$×NO_2, 19) $PM_{2.5}$, T, RH, NO_2, $PM_{2.5}$×RH, $PM_{2.5}$×NO_2, and T×RH, 20) $PM_{2.5}$, T, RH, NO_2, $PM_{2.5}$×NO_2, T×RH, T×NO_2, and RH×NO_2, 21) $PM_{2.5}$, T, RH, NO_2, $PM_{2.5}$×T, $PM_{2.5}$×RH, $PM_{2.5}$×NO_2, T×NO_2, and RH×NO_2, and 22) $PM_{2.5}$, T, RH, NO_2, $PM_{2.5}$×T, $PM_{2.5}$×RH, $PM_{2.5}$×NO_2, T×RH, T×NO_2, and RH×NO_2. Table 4 summarizes calibration results, including R^2, RMSE, and MAE of test sets for PA-II units using the RF method with the selected combinations of features.

Like the MLR method, the RF method showed better performance on the training set than on the test set. Some combinations had larger RMSE differences than 0.6 between training and test sets, while others have differences smaller than 0.4. We note that some combinations with multiplicative terms showed significant RMSE differences between two datasets, which might have occurred because of overfitting the training dataset. Nonetheless, the RF models with the other combinations had lower RMSE than the model using only $PM_{2.5}$. Considering a single environmental variable together with $PM_{2.5}$ improved the calibration performance in terms of values of RMSE and MAE compared to the RF model with only $PM_{2.5}$. Specifically, RH had more significant impact on the performance enhancement of the RF calibration model than T as seen in the results of combinations #2 and #3. Including the additional multiplicative term of $PM_{2.5}$×RH had an insignificant effect on RMSE compared to the RF model with $PM_{2.5}$ and RH. Both meteorological variables together, i.e., combination #5, yielded lower RMSE compared to the RF model with $PM_{2.5}$ and RH, i.e., combination #3, but the improvement was insignificant. In contrast to MLR models, more than one multiplicative term, i.e., combinations #6 to #9, bring about worse RMSE compared to considering a single

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

meteorological variable. When we analyze calibration methods without NO_2, the RF model with $PM_{2.5}$, T, and RH improved RMSE by 0.135 $\mu g/m^3$, compared to the best MLR model.

Utilizing NO_2 on RF models had different effects on calibration performance, depending on the combinations of predictor variables. The RF model of combination #10 with the additional NO_2 term resulted in an RMSE of 4.461, which was little improvement compared to combination #1 with only $PM_{2.5}$ and an RMSE of 4.472. The RF model with $PM_{2.5}$ and NO_2 had a larger RMSE than the MLR model with the same features, but the difference was not significant, it did not show enough performance improvement to warrant adding the multiplicative term of $PM_{2.5} \times NO_2$ from combination #10. Adding single or two meteorological variables to RF models of combinations #12 and #16 lead to remarkable performance enhancement over combination #10, with RH, RMSE decreasing by 0.443 $\mu g/m^3$. Furthermore, RMSE dropped an additional 0.106 $\mu g/m^3$ when T was added as an additional feature. The combinations consisting of one or more multiplicative interaction terms resulted in either an insignificant improvement or a slight decline in the performance in terms of RMSE and MAE when compared with combination #16 consisting of $PM_{2.5}$, T, RH, and NO_2. In other words, there is no need to consider multiplicative interaction terms when we use the RF model, because there is no outstanding performance improvement.

As with the MLR method, it was shown that including NO_2 as a consideration in RF methods can improve calibration performance. Moreover, by integrating two additional variables, such as T and RH, even better calibration performance can be achieved.

The RF method was shown to have a better performance than the MLR method when NO_2 was not considered. From the viewpoint of RMSE, the best performance from MLR and RF methods was 4.171 and 4.036, respectively. However, when we consider NO_2, the best MLR model is not significantly different from the best RF model. For instance, the RMSE values from the best MLR and RF models were 3.938 and 3.906, respectively. Their corresponding R^2 values did not differ meaningfully. Nonetheless, the MAE of 2.789 achieved from the best MLR is lower than that achieved by the best RF (i.e., 2.863). From these results, we conclude that better calibration can be obtained by considering NO_2 additionally. Furthermore, when NO_2 is considered, the MLR model can enhance calibration performance without the need for an RF model.

3.6.3 Effect of Far NO_2 on Calibration Performance

In the previous subsections, it was demonstrated that including NO_2 as a consideration can effectively improve the calibration performance of PA-II units. However, it is not always feasible to have an NO_2 instrument with high accuracy colocated with a low-cost PM sensor. Instead, an alternative approach is to colocate a low-cost NO_2 sensor with a PA-II unit, but this approach is hindered by the unreliability of NO_2 sensors. To address this issue, we investigated the usefulness of using data from far NO_2 instruments installed with PA-II units for the calibration algorithm.

We selected two monitoring sites that measure NO_2 near the Rubidoux site. Two monitoring sites identified were 06-065-8005 and 06-071-0027. The distances between the two monitoring sites and the Rubidoux site are 7.05 and 18.87 km, respectively. The correlations of NO_2 measurements obtained from the Rubidoux site with that of 06-065-8005 and 06-071-0027 were 0.895 and 0.621, respectively. The 06-065-8005 site had NO_2 measurements that highly correlated with the Rubidoux site, while the 06-071-0027 measurements were only moderately correlated.

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

To evaluate the usefulness of measurements of far NO_2 on the calibration of a low-cost PM sensor, we used NO_2 data measured from monitoring sites near to the PA-II 7 unit as the test dataset rather than data from the Rubidoux site colocated with it. In other words, we used NO_2 data for training a calibration model and used the NO_2 data measured from sites 06-065-8005 and 06-071-0027 for testing it.

Table 5 shows calibration performance using MLR and RF methods with NO_2 collected from the air quality monitoring sites near the PA-II unit. In the case of MLR methods used with 06-065-8005 data, the difference in RMSE between NO_2 data obtained from a colocated NO_2 instrument, called colocated NO_2, and a distant NO_2 instrument, called distant NO_2, was less than 0.05 μgm^3 for every selected combination defined in previous two subsections for the MLR and RF methods. All MLR models using distant NO_2, except combinations #10 and #11, yielded lower errors than all MLR models without NO_2 as shown in Table 3. For example, the worst RMSE of the MLR methods using distant NO_2 data (except combinations #10 and #11) was 4.051 μgm^3, while the best RMSE without NO_2 was 4.171 μgm^3. Like RMSE, other metrics, such as R^2 and MAE, also showed a calibration performance enhancement for these combinations with distant NO_2.

When we used an MLR algorithm with NO_2 data, the result of the calibration performance for the monitoring site 06-071-0027 showed a new aspect from that of 06-065-8005. All MLR methods using far NO_2 data from site 06-071-0027 had a higher RMSE than was achieved with the MLR algorithm based on combination #1 without referring NO_2 data from the colocated Rubidoux instrument, which had an RMSE of 4.534. This result can be explained by comparing correlation of NO_2 measured from the Rubidoux site with measurements from site 06-065-8005 as well as site 06-071-0027. The NO_2 correlation between Rubidoux measurements and site 06-065-8005 was 0.895, while the correlation with site 06-071-0027 was 0.621. These results shows that 06-065-8005 data is much more correlated with the Rubidoux site in terms of NO_2.

In the case of RF models, the use of the distant NO_2 data from site 06-065-8005 increased RMSE compared to using colocated NO_2 data, but not significantly, since the maximum gap was just 0.051 μgm^3. Similar to the MLR method, all RF models referring to distant NO_2 from site 06-065-8005, except combinations #10 and #11, resulted in a better calibration performance than was seen in combination #1 without NO_2 which had an RMSE of 4.472 shown in Table 4. Other metrics, such as R^2 and MAE, also showed a calibration performance improvement. In the case of RF models using data from site 06-071-0027, calibration performance for each combination was degraded compared to the corresponding combination using colocated NO_2, which had similar results of the MLR model. As we explained previously, the higher the correlation of NO_2 measurements from the Rubidoux site with measurements from sites 06-065-8005 and 06-071-0027, the better the calibration performance of the RF model; that is, all combinations with far NO_2 from 06-065-8005 provide a lower RMSE than those from 06-071-0027. Moreover, when we consider 06-065-8005 having a high correlation of NO_2 with the expensive NO_2 instrument colocated with the PA-II 7 unit, the best RMSE for all combinations using the RF model is slightly lower than that based on the MLR method.

In the case of 06-065-8005, RF models using distant NO_2 resulted in lower, but insignificant, RMSE values, compared to MLR models using distant NO_2. From these results, we draw the conclusion that the use of NO_2 collected from distant instruments with a high correlation with a colocated NO_2 site of PA-II units can improve the PA-II unit's calibration performance.

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

Furthermore, both MLR and RF models can be good calibration models when distant NO_2 is considered. This is different from the conclusion that calibration performance of RF models is better than MLR models (Zimmerman et al., 2018).

4 Conclusions

As factors directly affecting the performance of low-cost PM sensors, temperature, relative humidity, and particle composition have been scrutinized. In addition to these direct factors, this study investigated whether NO_2, which is a precursor that creates $PM_{2.5}$ by chemical reactions in the atmosphere, has the potential to improve the performance of low-cost $PM_{2.5}$ sensors. To this end, we used the PurpleAir PA-II unit, which contains two Plantower PMS 5003 sensors, as a low-cost $PM_{2.5}$ sensor. The PA-II units need to be typically installed close to reference monitoring sites measuring $PM_{2.5}$ concentrations and other pollutants, such as NO_2, in order to analyze their calibration. We identified a EPA-certified monitoring instrument whose deployed location is within close proximity to the installed location of 14 PA-II units, which satisfied the condition for co-location with a reference monitoring site. The monitoring site is located in Rubidoux, CA, USA. A study period of two years, i.e., from Jan. 2018 to Dec. 2019, was selected to include all seasons. A single unit among 14 PA-II units was selected based on the availability of 23 months or more of measurement data from each PA-II unit, as well as its low intra-model variability through correlation analysis.

The selected PA-II unit was compared to FRM and BAM-1020 instruments based on daily and hourly $PM_{2.5}$ measurements. A comparison of the BAM-1020 instrument with the FRM instrument was also conducted on a daily $PM_{2.5}$ measurement basis to evaluate the performance of the BAM-1020. The BAM-1020 instrument had a slope of 0.923, an intercept of 0.741, and a R^2 of 0.896 against the FRM instrument, which implies it provides acceptable performance as a reference monitor for the calibration of low-cost $PM_{2.5}$ sensors. For a PA-II unit, the Pearson correlation coefficient against the BAM-1020 instrument was shown to be 0.928 on an hourly basis. The per-month analysis was conducted on hourly $PM_{2.5}$ measurements of the PA-II unit against the BAM-1020. Results showed the PA-II unit has a good correlation during the winter season, i.e., Nov., Dec., and Jan., with an R^2 value between 0.819 and 0.906, but a lower correlation during other months. The performance of the PA-II units was not notably affected by temperature or relative humidity (RH) during the winter months. Temperature and/or RH were found to improve R^2 during June and July 2018, but this effect in 2019 was not the same as in 2018.

A per-month analysis showed that NO_2 is a key factor that increased the value of R^2 during Sep. 2018, and Aug. and Sep. 2019. The effect of the addition of NO_2 for calibration of PA-II units was much larger when RH and temperature were considered together. In particular, NO_2 was shown to have more effect during months when the performance of PA-II units is moderate. It is expected that NO_2 can be used to improve the performance of low-cost $PM_{2.5}$ sensors, but the effect of NO_2 should be further investigated for various ambient conditions.

Two methods for calibrating PA-II units, the Multiple Linear Regression (MLR) and Random Forest (RF), were evaluated on a test set of one year of data. We considered additive and multiplicative terms in two calibration methods. The RF method yielded better performance than the MLR method because it provides a larger R^2 as well as smaller RMSE, and MAE when NO_2, called colocated NO_2, measured from the colocated monitoring site was not used for calibration. However, when colo-

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

cated NO_2 is considered, MLR models showed similar performance to RF models. When several features, such as $PM_{2.5}$, temperature, RH, NO_2, and their multiplicative terms, are considered together to calibrate $PM_{2.5}$ measurement data using the MLR method, the calibration performance was shown to increase remarkably compared to cases where only $PM_{2.5}$ are considered. For instance, the RMSE value decreased from 4.534 $\mu g/m^3$ to 3.938 $\mu g/m^3$. In RF models with colocated NO_2, inclusion of temperature and RH improved R^2, RMSE, and MAE by an increase of 0.013, a decrease of 0.125, and 0.094 $\mu g/m^3$, respectively, compared to the best RF models without NO_2. Contrary to the MLR model, multiplicative interaction terms do not affect calibration performance with a certain direction, compared to those without NO_2; some combinations of features provide slight enhancement, while the others cause worse performance.

We showed that NO_2 data could improve calibration performance in both MLR and RF models. The NO_2 data we referred to was measured from an expensive reference monitor and is very reliable. However, it is not always feasible to have an NO_2 instrument with high accuracy colocated with a low-cost PM sensor. An alternatives is to use low-cost NO_2 sensors. However, their performance remains questionable. To solve this issue, we investigated the effectiveness of using NO_2 measurements collected from distant reliable NO_2 monitoring sites, called distant NO_2, whose location is not that far from a low-cost $PM_{2.5}$ sensor. It was demonstrated that distant NO_2 is effective for calibration models based on the MLR and RF algorithms when distant NO_2 has a high correlation with collocated NO_2. Furthermore, we showed that MLR method can achieve a similar calibration performance to the RF method when reliable distant NO_2 is considered.

Data availability. All data can be provided by the authors upon request.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This paper is financially supported by the Ministry of Trade, Industry and Energy(MOTIE, Korea) through the fostering project of 'Smart City Urban Infrastructure Air Quality Real-time Monitoring and Prediction Platform Technology Development ' supervised by the Korea Institute for Advancement of Technology(KIAT).

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

References

Alvarado, M., Gonzalez, F., Fletcher, A., Doshi, A., Alvarado, M., Gonzalez, F., Fletcher, A., and Doshi, A.: Towards the Development of a Low Cost Airborne Sensing System to Monitor Dust Particles after Blasting at Open-Pit Mine Sites, Sensors, 15, 19667–19687, 2015.

Austin, E., Novosselov, I., Seto, E., and Yost, M.G.: Laboratory Evaluation of the Shinyei PPD42NS Low-Cost Particulate Matter Sensor, Plos ONE, 10, 2015.

Badura, M., Batog, P., Drzeniecka-Osciadacz, A., and Modzel, P.: Evaluation of low-cost sensors for ambient PM2.5 monitoring, J. Sens. 2018.

Nilson, B., Jackson, P. L., Schiller, C. L., and Parsons, M. T.: Development and evaluation of correction models for a low-cost fine particulate matter monitor, Atmos. Meas. Tech., 15, 3315-3328, 2022.

Barkjohn, K. K., Bergin, M. H., Norris, C., Schauer, J. J., Zhang, Y., Black, M., Hu, M., and Zhang, J.: Using Lowcost sensors to Quantify the Effects of Air Filtration on Indoor and Personal Exposure Relevant PM2:5 Concentrations in Beijing, China, Aerosol Air Qual. Res., 20, 297–313, https://doi.org/10.4209/aaqr.2018.11.0394, 2020.

Barkjohn, K. K., Gantt, B., and Clements A. L.: 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, 2021.

Breiman, L.: Random Forests, Machine Learning, 45, 5-32, 2001.

Budde, M., Müller, T., Laquai, B., Streibl, N., Schwarz, A., Schindler, G., Riedel, T., Beigl, M., and Dittler, A.: Suitability of the Low-Cost SDS011 Particle Sensor for Urban PM-Monitoring, In Proceedings of the 3rd International Conference on Atmospheric Dust, Bari, Italy, 29–31 May 2018.

Cavaliere, A., Carotenuto, F., Di Gennaro, F., Gioli, B., Gualtieri, G., Martelli, F., Matese, A., Toscano, P., Vagnoli, C., and Zaldei, A.: Development of Low-Cost Air Quality Stations for Next Generation Monitoring Networks: Calibration and Validation of PM2.5 and PM10 Sensors, Sensors, 18, 2843, 2018.

Crilley, L.R., Shaw, M., Pound, R., Kramer, L.J., Price, R., Young, S., Lewis, A.C., and Pope, F.D.: Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring, Atmos. Meas. Tech., 11, 709–720, 2018.

Crilley, L. R., Singh, A., Kramer, L. J., Shaw, M. D., Alam M. S., Apte, J. S., Bloss, W. J., Ruiz L. H., Fu, P., Fu, W., Gani, S. , Gatari M., Ilyinskaya, E., Lewis, A. C., Ng'ang'a, D. 6, Sun, Y. , Whitty R. C. W. , Yue S., Young, S., and Pope F. D.: Effect of aerosol composition on the performance of low-cost optical particle counter correction factors, Atmos. Meas. Tech., 13, 1181–1193, 2020.

Evans, J., van Donkelaar A., Martin, R. V., Burnett, R., Rainham, D. G., Birkett, N. J., and Krewski, D.: Estimates of globalmortality attributable to particulate air pollution using satellite imagery, Environ. Res. vol. 120, 33-42, 2013.

Feenstra, B., Papapostolou, V., Hasheminassab, S., Zhang, H., Boghossian, B. D., Cocker, D., and Polidori, A.: Performance evaluation of twelve low-cost PM2:5 sensors at an ambient air monitoring site, Atmos. Environ., 216, 116946, https://doi.org/10.1016/j.atmosenv.2019.116946, 2019.

Feinberg, S., Williams, R., Hagler, G.S.W., Rickard, J., Brown, R., Garver, D., Harshfield, G., Ster, P., Mattson, E., Judge, R., and Garvey, S.: Long-term evaluation of air sensor technology under ambient conditions in Denver, Colorado, Atmos. Meas. Tech., 11, 4605–4615, 2018.

Gao, M., Cao, J., and Seto, E.: A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of PM2.5 in Xi'an, China, Environ. Pollut., 199, 56–65, 2015.

Holstius, D.M., Pillarisetti, A., Smith, K.R., and Seto, E.: Field calibrations of a low-cost aerosol sensor at a regulatory monitoring site in California, Atmos. Meas. Tech., 7, 1121–1131, 2014.

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

Hua, J., Zhang, Y., Foy, B., Mei, X., Shang, J., Zhang, Y., Sulaymon, I. D., and Zhou, D.: Improved $PM_{2.5}$ concentration estimates from low-cost sensors using calibration models categorized by relative humidity, Aerosol Science and Technology, 55, 2021.

Jayaratne R., Liu1, X., Thai1, P., Dunbabin, M., and Morawska, L.: The influence of humidity on the performance of a low-cost air particle mass sensor and the effect of atmospheric fog, Atmos. Meas. Tech., 11, 4883–4890, 2018.

Jiao, W., Hagler, G., Williams, R., Sharpe, R., Brown, R., Garver, D., Judge, R., Caudill, M., Rickard, J., Davis, M., Weinstock, L., Zimmer-Dauphinee, S., and Buckley, K.: Community Air Sensor Network (CAIRSENSE) project: evaluation of low-cost sensor performance in a suburban environment in the southeastern United States, Atmos. Meas. Tech., vol. 9, no. 11, 5281–5292, 2016.

Johnson, K., Bergin, M., Russell, A., and Hagler, G.: Field Test of Several Low-Cost Particulate Matter Sensors in High and Low Concentration Urban Environments, Aerosol Air. Qual. Res. vol. 18, no. 3, 565–578, 2028.

Kelly, K.E., Whitaker, J., Petty, A.,Widmer, C., Dybwad, A., Sleeth, D., Martin, R., and Butterfield, A.: Ambient and laboratory evaluation of a low-cost particulate matter sensor, Environ. Pollut., 221, 491–500, 2017.

Liu, H.-Y, Bartonova, A., Schindler, M., Sharma, M., Behera, S.N., Katiyar, K., and Dikshit, O.: Respiratory Disease in Relation to Outdoor Air Pollution in Kanpur, India. Arch. Environ. Occup. Health, vol. 68, 204-217, 2013.

Liu, H.-Y., Dunea, D., Iordache, S., and Pohoata, A.: A Review of Airborne Particulate Matter Effects on Young Children's Respiratory Symptoms and Diseases, Atmosphere, vol. 9, 150, 2018.

Liu, H.-Y., Schneider, P., Haugen, R., and Vogt, M.: Performance Assessment of a Low-Cost PM2.5 Sensor for a near Four-Month Period in Oslo, Norway, Atmosphere, 10, 41, 2019.

Magi, B. I., Cupini, C., Francis, J., Green, M., and Hauser, C.: Evaluation of PM2:5 measured in an urban setting using a lowcost optical particle counter and a Federal Equivalent Method Beta Attenuation Monitor, Aerosol Sci. Tech., 54, 147–159, 2019.

Malings, C., Tanzer, R., Hauryliuk, A., Saha, P. K., Robinson, A. L., Presto, A. A., and Subramanian, R.: Fine particle mass monitoring with low-cost sensors: Corrections and longterm performance evaluation, Aerosol Sci. Tech., 54, 160–174, 2020.

Mansfield, E. R. and Helms, B. P.: Detecting Multicollinearity, The American Statistician, 36, 158-160, 1982.

Mukherjee, A., Stanton, L.G., Graham, A.R., and Roberts, P.T. Assessing the Utility of Low-Cost Particulate Matter Sensors over a 12-Week Period in the Cuyama Valley of California, Sensors, 17, 1805, 2017.

Olivares, G., and Edwards, S. The Outdoor Dust Information Node (ODIN) – development and performance assessment of a low cost ambient dust sensor, Atmos. Meas. Tech. Discuss., 8, 7511–7533, 2015.

Pawar, H. and Sinha, B.: Humidity, density and inlet aspiration efficiency correction improve accuracy of a low-cost sensor during field calibration at a suburban site in the north-western Indo- Gangetic Plain (NW-IGP), Aerosol Sci. Tech., 54, 685–703, https://doi.org/10.1080/02786826.2020.1719971, 2020.

PurpleAir Map, air quality Map [WWW Document] URL: https://map.purpleair.org (last access: 1 May 2020).

Sayahi, T., Butterfield, A., and Kelly, K. E.: Long-term field evaluation of the Plantower PMS low-cost particulate matter sensors, Environ. Pollut., vol. 245, 932-940, Feb. 2019.

South Cost Air Quality Management District (SCAQMD): Field Evaluation AirBeam PM Sensor, available at: http://www.aqmd.gov/docs/default-source/aq-spec/laboratory-evaluations/airbeam—laboratory-evaluation.pdf?sfvrsn=6 (last access: 1 May 2020), 2017.

South Cost Air Quality Management District (SCAQMD): Field Evaluation Purple Air (PA-II) PM Sensor, available at: http://www.aqmd.gov/docs/default-source/aq-spec/field-evaluations/purple-air-pa-ii—field-evaluation.pdf?sfvrsn=2 (last access: 1 May 2020), 2017.

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

South Cost Air Quality Management District (SCAQMD): Field Evaluation Laser Egg PM Sensor, available at: http://www.aqmd.gov/docs/default-source/aq-spec/field-evaluations/laser-egg—field-evaluation.pdf (last access: 1 May 2020), 2017.

Si, M., Xiong, Y., Du, S., and Du K.: Evaluation and calibration of a low-cost particle sensor in ambient conditions using machine-learning methods, Atmos. Meas. Tech., 13, 1693–1707, 2020.

Sousan, S., Koehler, K., Thomas, G., Park, J.H., Hillman, M., Halterman, A., and Peters, T.M.: Inter-comparison of low-cost sensors for measuring the mass concentration of occupational aerosols, Aerosol Sci. Technol., 50, 462–473, 2016.

Wallace, L., Bi, J., Ott, W. R., Sarnat, J., and Liu, Y.: Calibration of low-cost PurpleAir outdoor monitors using an improved method of calculating PM2.5, Atmos. Environ., vol. 256, Jul. 2021.

Wang, Y., Li, J., Jing, H., Zhang, Q., Jiang, J., and Biswas, P.: Laboratory Evaluation and Calibration of Three Low-Cost Particle Sensors for Particulate Matter Measurement, Aerosol Sci. Technol., 49, 1063–1077, 2015.

Zheng, T., Bergin, M.H., Johnson, K.K., Tripathi, S.N., Shirodkar, S., Landis, M.S., Sutaria, R., and Carlson, D.E.: Field evaluation of low-cost particulate matter sensors in high and low concentration environments, Atmos. Meas. Tech. Discuss., 11, 4823–4846, 2018.

Zimmerman, N., Presto, A. A., Kumar, S. P. N., Gu, J., Hauryliuk, A., Robinson, E. S., Robinson, A. L., and R. Subramanian: A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring, Atmos. Meas. Tech., 11, 291–313, 2018.

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

Table 1. Information about 14 PA-II units, such as their ID, location (latitude and longitude), sensor name, start time of measurement, end time of measurement, and non-operating months.

ID	Latitude	Longitude	Sensor Name	Start Time of Measurement	End Time of Measurement	Non-Operating Months
1866	33.999978	-117.41676	RIVR_Co-loc1	7/10/17	4/27/20	Sep., Oct., Nov., and Dec. 2018
1854	33.999503	-117.41602	RIVR_Co-loc2	7/10/17	4/27/20	
2346	33.999978	-117.41676	RIVR_Co-loc3	7/31/17	4/27/20	
2325	33.999978	-117.41676	RIVR_Co-loc4	7/31/17	4/27/20	Sep., Oct., Nov., and Dec. 2018
2167	33.999978	-117.41676	RIVR_Co-loc5	7/17/17	4/27/20	
2155	33.999978	-117.41676	RIVR_Co-loc6	7/17/17	4/27/20	May, 2018
2612	33.999515	-117.41595	RIVR_Co-loc7	8/7/17	4/27/20	
2758	33.999978	-117.41676	RIVR_Co-loc8	8/11/17	4/27/20	Sep. 2018
3537	33.999381	-117.41601	RIVR_Co-loc9	9/20/17	4/27/20	May, Sep., Oct., Nov., and Dec. 2018
4748	33.999516	-117.41594	RIVR_Co-loc10	11/22/17	4/27/20	May, Aug., Sep., Oct., Nov., and Dec. 2018 and Jan 2019
4731	33.999504	-117.41593	RIVR_Co-loc11	11/22/17	3/1/19	Jan., Feb., Mar., Sep., Oct., Nov., and Dec. 2018
5280	33.99946	-117.41594	RIVR_Co-loc12	12/12/17	4/27/20	May, Sep., Oct., Nov., and Dec. 2018
5284	33.999451	-117.41591	RIVR_Co-loc13	12/12/17	4/27/20	May, Sep., Oct., Nov., and Dec. 2018
6806	33.999583	-117.41621	RIVR_Co-loc14	1/30/18	4/27/20	Apr., Sep., Oct., and Nov. 2018
6912	33.999482	-117.41627	RIVR_Co-loc15	1/31/18	4/27/20	Apr., Sep., Oct., and Nov. 2018
9226	33.999389	-117.41633	RIVR_Co-loc16	3/24/18	4/27/20	Apr., Sep., Oct., Nov., and Dec. 2018
9358	33.999319	-117.41638	RIVR_Co-loc17	3/25/18	4/27/20	Apr., Sep., Oct., Nov., and Dec. 2018

Table 2. R^2, RMSE, and MAE of the PA-II unit against the BAM-1020 based on the hourly PM2.5 measurement data for each month.

	Jan-18	Feb-18	Mar-18	Apr-18	May-18	Jun-18	Jul-18	Aug-18	Sep-18	Oct-18	Nov-18	Dec-18
R^2	0.936	0.799	0.845	0.759	0.659	0.695	0.359	0.816	0.591	0.784	0.829	0.905
RMSE	4.201	3.735	2.932	3.938	3.477	4.097	5.615	3.204	4.550	3.650	3.832	3.765
MAE	3.171	2.721	2.196	3.098	2.716	3.267	3.597	2.424	3.358	2.844	2.913	2.743
Intercept	3.898	4.229	2.898	7.090	4.694	7.925	6.721	4.692	6.357	2.682	3.269	1.445
Slope	0.502	0.475	0.525	0.446	0.486	0.475	0.434	0.459	0.382	0.420	0.409	0.472
	Jan-19	Feb-19	Mar-19	Apr-19	May-19	Jun-19	Jul-19	Aug-19	Sep-19	Oct-19	Nov-19	Dec-19
R^2	0.884	0.750	0.735	0.618	0.801	0.730	0.893	0.405	0.441	0.523	0.880	0.813
RMSE	3.326	2.940	2.753	3.703	3.146	3.403	4.127	4.220	3.292	4.768	4.474	3.866
MAE	2.485	2.216	2.124	2.892	2.349	2.700	3.082	2.564	2.558	3.360	3.238	2.934
Intercept	1.961	2.190	1.881	4.065	2.525	3.225	3.070	5.649	5.312	5.088	2.976	1.165
Slope	0.397	0.354	0.427	0.385	0.418	0.383	0.575	0.428	0.511	0.483	0.497	0.572

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

Table 3. Calibration result (R^2, RMSE ($\mu g/m^3$), and MAE ($\mu g/m^3$)) of hourly $PM_{2.5}$ concentrations using MLR for the PA-II 7 unit based on the selected combinations.

NO_2 not included							NO_2 included						
Feature	Training Set			Test Set			Feature	Training Set			Test Set		
Vector	R^2	RMSE	MAE	R^2	RMSE	MAE	Vector	R^2	RMSE	MAE	R^2	RMSE	MAE
1	0.786	4.466	3.338	0.726	4.534	3.447	10	0.789	4.436	3.319	0.737	4.445	3.359
2	0.800	4.321	3.236	0.753	4.306	3.198	11	0.789	4.432	3.315	0.737	4.444	3.358
3	0.797	4.348	3.256	0.757	4.273	3.179	12	0.813	4.173	3.119	0.788	3.992	2.865
4	0.804	4.278	3.161	0.759	4.256	3.153	13	0.814	4.161	3.115	0.788	3.988	2.854
5	0.802	4.301	3.225	0.761	4.241	3.138	14	0.815	4.147	3.091	0.792	3.957	2.838
6	0.806	4.250	3.145	0.762	4.231	3.114	15	0.816	4.145	3.089	0.792	3.950	2.834
7	0.807	4.238	3.134	0.760	4.251	3.117	16	0.815	4.148	3.073	0.788	3.996	2.862
8	0.809	4.221	3.133	0.767	4.188	3.049	17	0.817	4.131	3.060	0.786	4.006	2.861
9	0.809	4.216	3.130	0.768	4.171	3.048	18	0.817	4.127	3.057	0.788	3.990	2.852
							19	0.818	4.118	3.068	0.793	3.942	2.793
							20	0.818	4.113	3.061	0.794	3.938	2.789

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

Table 4. Calibration result (R^2, RMSE ($\mu g/m^3$), and MAE ($\mu g/m^3$)) of hourly $PM_{2.5}$ concentrations using RF for for the PA-II 7 unit based on the selected combinations.

NO_2 not included							NO_2 included						
Feature	Training Set			Test Set			Feature	Training Set			Test Set		
Vector	R^2	RMSE	MAE	R^2	RMSE	MAE	Vector	R^2	RMSE	MAE	R^2	RMSE	MAE
1	0.813	4.180	3.122	0.736	4.472	3.342	10	0.807	4.244	3.157	0.737	4.461	3.321
2	0.831	3.965	2.983	0.763	4.236	3.169	11	0.807	4.244	3.151	0.736	4.470	3.319
3	0.844	3.811	2.836	0.781	4.076	2.991	12	0.849	3.751	2.789	0.787	4.018	2.956
4	0.844	3.809	2.834	0.784	4.046	2.979	13	0.878	3.367	2.567	0.798	3.910	2.893
5	0.847	3.777	2.813	0.785	4.036	2.973	14	0.876	3.405	2.570	0.784	4.044	2.951
6	0.851	3.730	2.801	0.777	4.108	2.989	15	0.876	3.406	2.580	0.787	4.021	2.932
7	0.875	3.417	2.572	0.777	4.108	2.995	16	0.884	3.284	2.491	0.798	3.911	2.879
8	0.848	3.761	2.813	0.776	4.118	2.994	17	0.885	3.278	2.490	0.797	3.918	2.876
9	0.851	3.727	2.795	0.773	4.151	3.005	18	0.911	2.887	2.183	0.799	3.906	2.863
							19	0.914	2.832	2.159	0.800	3.896	2.862
							20	0.915	2.813	2.137	0.794	3.955	2.885
							21	0.909	2.918	2.191	0.793	3.964	2.890
							22	0.886	3.256	2.466	0.791	3.983	2.907

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

Table 5. Calibration result (R^2, RMSE ($\mu g/m^3$), and MAE ($\mu g/m^3$)) of hourly $PM_{2.5}$ concentrations using MLR and RF models for the PA-II 7 unit based on the selected combinations additionally with distant NO_2.

		MLR						RF					
Site	Feature	Colocated NO_2			Nearby NO_2			Colocated NO_2			Nearby NO_2		
ID	Vector	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
	10	0.738	4.453	3.361	0.739	4.448	3.353	0.737	4.461	3.321	0.736	4.477	3.333
0	11	0.739	4.452	3.360	0.739	4.445	3.349	0.736	4.470	3.319	0.735	4.482	3.327
6	12	0.789	4.002	2.869	0.785	4.034	2.882	0.787	4.018	2.956	0.786	4.025	2.965
-	13	0.789	3.999	2.857	0.786	4.032	2.870	0.798	3.910	2.893	0.794	3.955	2.929
0	14	0.792	3.967	2.841	0.789	4.003	2.852	0.784	4.044	2.951	0.782	4.069	2.967
6	15	0.793	3.960	2.837	0.789	4.004	2.855	0.787	4.021	2.932	0.781	4.072	2.963
5	16	0.788	4.006	2.866	0.785	4.036	2.878	0.798	3.911	2.879	0.796	3.933	2.892
-	17	0.787	4.016	2.864	0.784	4.051	2.874	0.797	3.918	2.876	0.795	3.940	2.896
8	18	0.789	4.000	2.854	0.784	4.045	2.874	0.799	3.906	2.863	0.796	3.932	2.876
0	19	0.794	3.954	2.797	0.791	3.980	2.802	0.800	3.896	2.862	0.797	3.924	2.879
0	20	0.794	3.949	2.793	0.791	3.979	2.802	0.794	3.955	2.885	0.791	3.982	2.899
5	21							0.793	3.964	2.890	0.790	3.989	2.905
	22							0.791	3.983	2.907	0.788	4.012	2.927
	10	0.738	4.453	3.361	0.710	4.690	3.613	0.737	4.461	3.321	0.736	4.529	3.379
0	11	0.739	4.452	3.360	0.711	4.679	3.595	0.736	4.470	3.319	0.735	4.572	3.405
6	12	0.789	4.002	2.869	0.687	4.874	3.806	0.787	4.018	2.956	0.786	4.180	3.065
-	13	0.789	3.999	2.857	0.689	4.853	3.763	0.798	3.910	2.893	0.794	4.821	3.648
0	14	0.792	3.967	2.841	0.691	4.838	3.728	0.784	4.044	2.951	0.782	4.325	3.155
7	15	0.793	3.960	2.837	0.676	4.959	3.841	0.787	4.021	2.932	0.781	4.966	3.728
1	16	0.788	4.006	2.866	0.702	4.751	3.674	0.798	3.911	2.879	0.796	4.337	3.216
-	17	0.787	4.016	2.864	0.700	4.773	3.665	0.797	3.918	2.876	0.795	4.339	3.224
0	18	0.789	4.000	2.854	0.681	4.915	3.794	0.799	3.906	2.863	0.796	4.480	3.341
0	19	0.794	3.954	2.797	0.708	4.702	3.624	0.800	3.896	2.862	0.797	4.476	3.340
2	20	0.794	3.949	2.793	0.688	4.864	3.784	0.794	3.955	2.885	0.791	4.653	3.469
7								0.793	3.964	2.890	0.790	4.626	3.436
								0.791	3.983	2.907	0.788	4.457	3.287

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

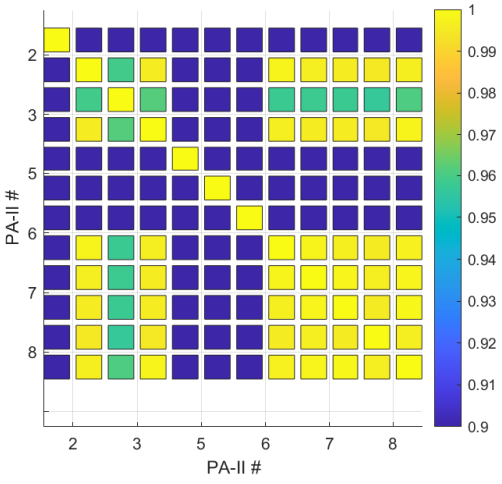

Figure 1. Correlation among all PMS 5003 sensors of the selected PA-II 2, 3, 5, 6, 7, and 8.

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

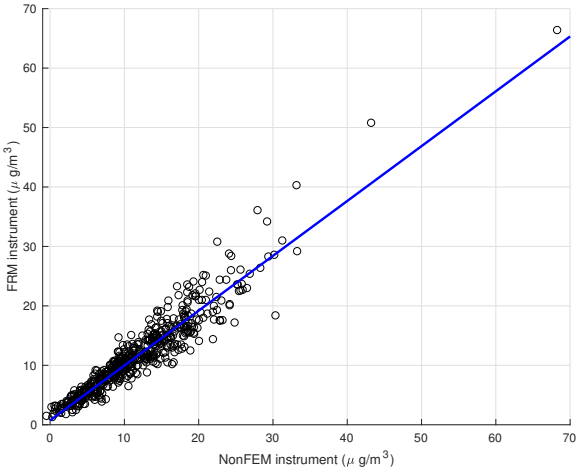

Figure 2. Scatter plot for daily $PM_{2.5}$ comparison of BAM-1020 (Non-FEM) instrument with the FRM instrument

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

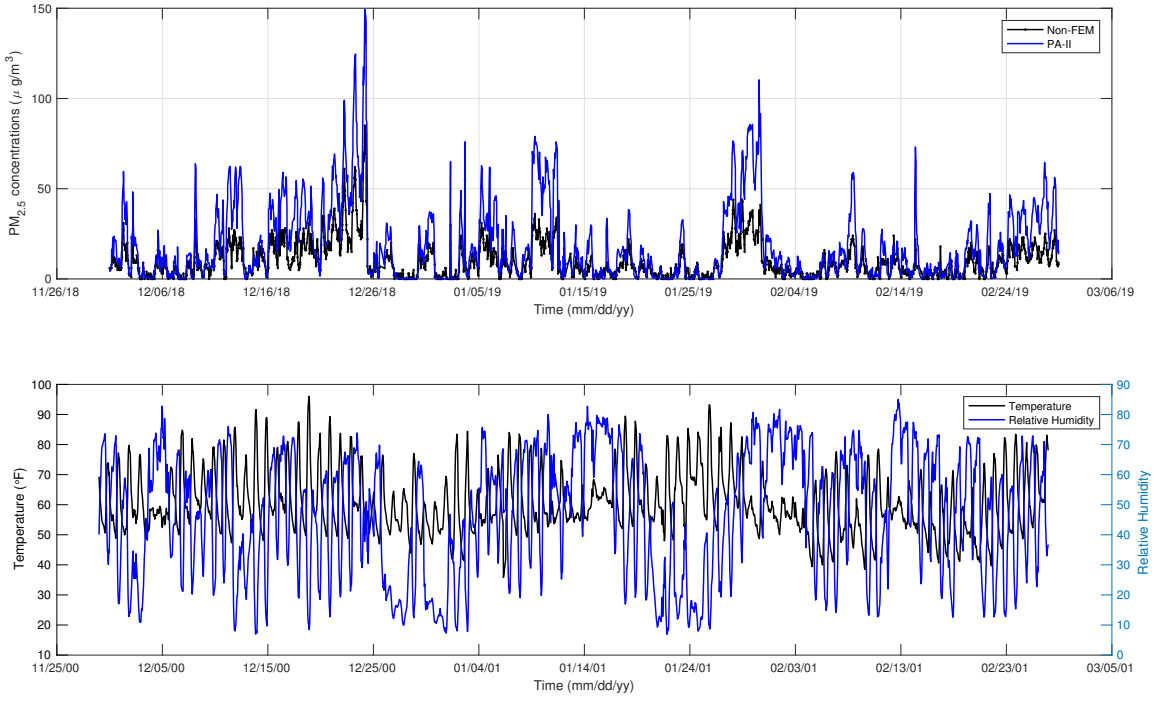

Figure 3. Hourly $PM_{2.5}$ concentrations measured from BAM-1020 (Non-FEM) and PA-II 7, and hourly temperature and relative humidity measured from PA-II 7 from Dec. 2018 to Feb. 2019.

https://doi.org/10.5194/egusphere-2023-1344
Preprint. Discussion started: 25 July 2023

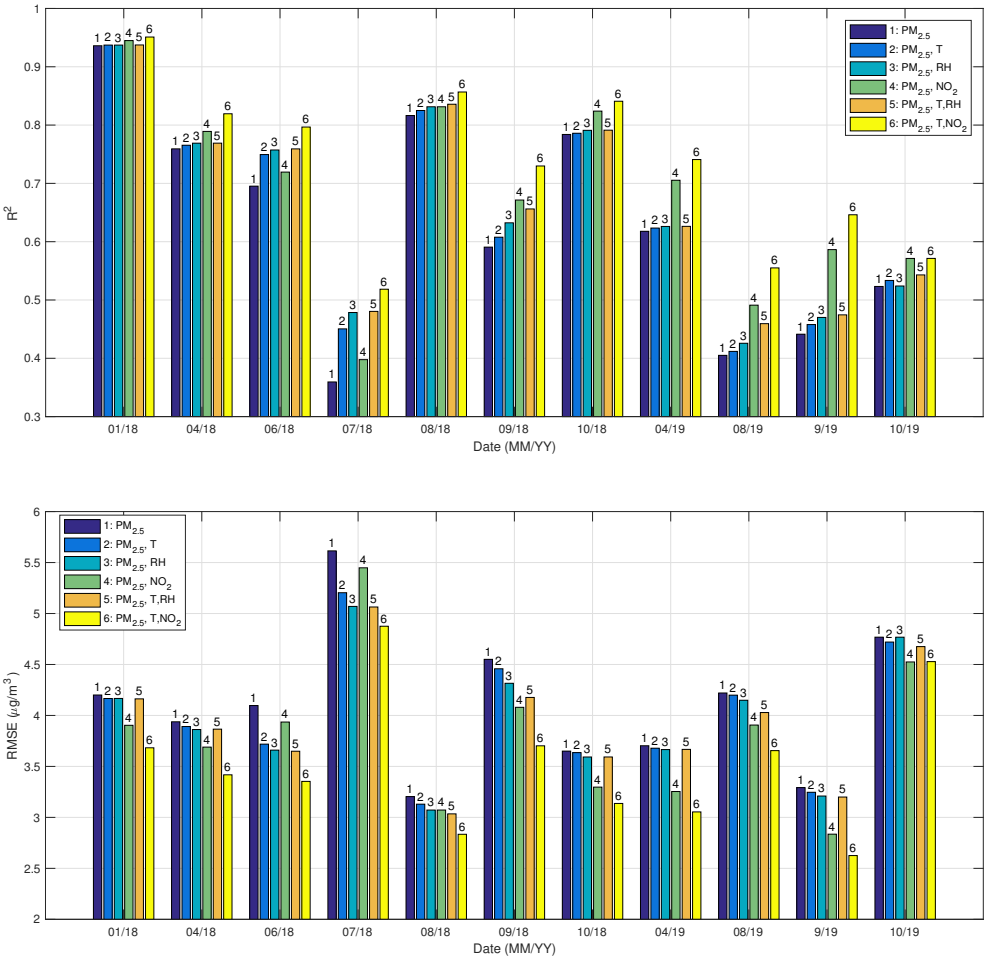

Figure 4. R^2 and RMSE using MLR method for the PA-II unit with the BAM-1020 for the selected months based on the following feature vectors; 1:($PM_{2.5}$), 2:($PM_{2.5}$, T), 3:($PM_{2.5}$, RH), 4:($PM_{2.5}$, NO_2), 5:($PM_{2.5}$, T, RH), and 6:($PM_{2.5}$, T, NO_2).