# Laboratory Evaluation of Low-cost PurpleAir PM Monitors and In-field Correction using Co-located Portable Filter Samplers

Jessica Tryner [a,b], Christian L'Orange [a], John Mehaffy [a], Daniel Miller-Lionberg [b], Josephine C. Hofstetter [b], Ander Wilson [c], and John Volckens [a,*]

[a] Department of Mechanical Engineering, Colorado State University, 1374 Campus Delivery, Fort Collins, Colorado, United States 80523

[b] Access Sensor Technologies, 2401 Research Blvd, Suite 107, Fort Collins, Colorado, United States 80526

[c] Department of Statistics, Colorado State University, 1801 Campus Delivery, Fort Collins, Colorado, United States 80523

* Corresponding author: john.volckens@colostate.edu

Authors' email addresses:

Jessica Tryner: jessica.tryner@colostate.edu

Christian L'Orange: christian.lorange@colostate.edu

John Mehaffy: john.mehaffy@colostate.edu

Daniel Miller-Lionberg: danlionberg@accsensors.com

Josephine C. Hofstetter: jhofstetter@accsensors.com

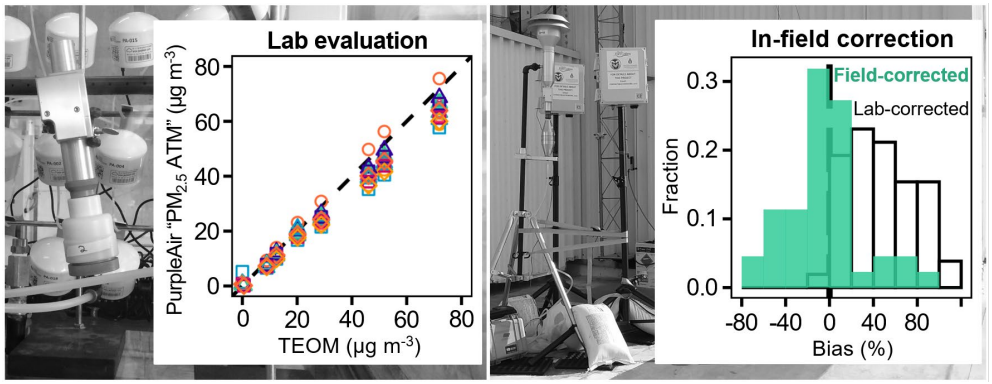
Ander Wilson: ander.wilson@colostate.edu

John Volckens: john.volckens@colostate.edu

ABSTRACT

Low-cost aerosol monitors can provide more spatially- and temporally-resolved data on ambient fine particulate matter ($PM_{2.5}$) concentrations than are available from regulatory monitoring networks; however, concentrations reported by low-cost monitors are sometimes inaccurate. We investigated laboratory- and field-based approaches for calibrating low-cost PurpleAir monitors. First, we investigated the linearity of the PurpleAir response to NIST Urban PM and derived a laboratory-based gravimetric correction factor. Then, we co-located PurpleAirs with portable filter samplers at 15 outdoor sites spanning 3×3-km in Fort Collins, CO, USA. We evaluated whether $PM_{2.5}$ correction factors calculated using ambient relative humidity data improved the accuracy of PurpleAir monitors (relative to reference filter samplers operated at 16.7 L $min^{-1}$). We also (1) evaluated gravimetric correction factors derived from periodic co-locations with portable filter samplers and (2) compared $PM_{2.5}$ concentrations measured using portable and reference filter samplers. Both before and after field deployment, a linear model relating NIST Urban PM concentrations reported by a tapered element oscillating microbalance and PurpleAir monitors ("$PM_{2.5}$ ATM") had $R^2$ = 99%; however, an F-test identified a significant lack of fit between the model and the data. The laboratory-based correction did not translate to the field. Over a 35-day period, time-averaged ambient $PM_{2.5}$ concentrations and RHs measured during 72- or 48-hour filter samples ranged from 1.5 to 8.3 μg $m^{-3}$ and 47% to 77%, respectively. Corrections calculated using ambient RH data increased the fraction of time-averaged PurpleAir $PM_{2.5}$ concentrations that were within 20% of the reference concentration from 24% (for uncorrected measurements) to 66%. Corrections derived from monthly, weekly, and concurrent in-field co-locations with portable filter samplers increased the fraction of time-averaged PurpleAir $PM_{2.5}$ concentrations that were within 20% of the reference to 46%, 54%, and 72%. $PM_{2.5}$ concentrations measured using portable filter samplers were within 20% of the reference for 69% of samples.

GRAPHICAL ABSTRACT




KEYWORDS

air quality; air pollution; particulate matter; Plantower; ultrasonic personal aerosol sampler; UPAS

## 1. INTRODUCTION

Thousands of low-cost (<$300) monitors for particulate matter (PM) air pollution, such as the PurpleAir (www.purpleair.com), have been deployed around the globe to provide more spatially-dense ambient air quality data than are available from networks used to monitor for regulatory compliance. Data from low-cost monitoring networks can: (1) assist in identification of pollution hotspots and sources (Gao et al., 2015; Rickenbacker et al., 2019; Zikova et al., 2017), (2) be used to identify targets for pollution reduction (Gillooly et al., 2019), and (3) enable individuals to adjust their behavior to reduce their contributions and exposure to air pollution (English et al., 2017). Data from such networks could be especially valuable to individuals that are more susceptible to adverse health impacts associated with exposure to particulate matter, such as those with asthma (Nelson, 2016).

Whereas many monitors used for regulatory compliance (e.g., conventional filter samplers, tapered element oscillating microbalances, beta attenuation monitors) measure PM mass directly or another quantity that depends on mass alone, low-cost monitors typically measure the amount of light scattered by particles (a quantity that does not depend on mass alone) to infer PM mass. As a result, measurements taken by low-cost monitors are sensitive to variations in particle properties such as size distribution and refractive index (Austin et al., 2015; Levy Zamora et al., 2019; Singer and Delp, 2018; Sousan et al., 2017; Tryner et al., 2019c; Wang et al., 2015). High ambient relative humidity (RH) can be especially problematic for low-cost PM monitors due to changes in particle size distribution and refractive index that result from aerosol water uptake when RH reaches 50% or more (Jayaratne et al., 2018; Levy Zamora et al., 2019; Magi et al., 2019; Malings et al., 2019). Even when regulatory monitors rely on light scattering techniques (e.g., particle spectrometers), they typically feature a mechanism to reduce the humidity of the sample prior to measurement.

Inaccuracies in the PM concentrations reported by low-cost monitors (as a result of the issues noted above) have been documented by researchers who co-located low-cost monitors with reference monitors in the field (Crilley et al., 2018; Feinberg et al., 2018; Gao et al., 2015; Holstius et al., 2014; Jiao et al., 2016; Kelly et al., 2017; Magi et al., 2019; Malings et al., 2019; Sayahi et al., 2019; Zheng et al., 2018). For example, Kelly et al. (2017) reported that PurpleAir monitors using Plantower PMS1003 sensors overestimated concentrations reported by regulatory-grade monitors when ambient $PM_{2.5}$ concentrations exceeded 10 μg $m^{-3}$ during wintertime inversions in Salt Lake City, UT, USA. If public understanding of the limitations of low-cost monitors is lacking, erroneously high readings could cause undue concern.

Given the benefits of spatially-dense PM monitoring networks and the limitations of the monitors that compose those networks, methods for calibrating low-cost monitors are needed to improve measurement accuracy (Bai et al., 2019; Holstius et al., 2014; Malings et al., 2019). Calibration of light-scattering sensors that report PM mass concentration is particularly challenging because particle properties, which affect the relationship between the amount of light scattered and particle mass, can vary both spatially and temporally (Sardar et al., 2005; Sioutas et al., 2000). As a result, a calibration derived in a certain place over a certain time period may not be applicable to other locations or time periods.

One correction approach is to adjust measurements reported by low-cost monitors (e.g., the PurpleAir) for aerosol water uptake using Equation 1 or another similar correction (Chakrabarti et al., 2004; Malings et al., 2019).

$$c_{PA}{}' = \frac{c_{PA}}{1 + \frac{0.25RH^2}{(1 - RH)}} \quad (1)$$

In Equation 1, $c_{PA}{}'$ is the corrected real-time $PM_{2.5}$ concentration reported by the PurpleAir ($\mu g\, m^{-3}$), $c_{PA}$ is the uncorrected real-time $PM_{2.5}$ concentration reported by the PurpleAir ($\mu g\, m^{-3}$), and $RH$ is a relative humidity value between 0 and 1.

Another approach is to temporarily co-locate a low-cost monitor (e.g., the PurpleAir) with a portable $PM_{2.5}$ gravimetric filter sampler, and then correct past and/or future concentrations reported by the low-cost monitor using Equation 2:

$$c_{PA}{}' = c_{PA}\frac{c_{filter}}{\bar{c}_{PA}} \quad (2)$$

where $c_{filter}$ is the $PM_{2.5}$ concentration derived from the filter sample, and $\bar{c}_{PA}$ is the sample-averaged $PM_{2.5}$ concentration reported by the PurpleAir (both in $\mu g\, m^{-3}$). Using this approach, the gravimetric correction factor is defined as $CF = c_{filter}/\bar{c}_{PA}$. Correction of real-time $PM_{2.5}$ concentrations (as opposed to just the sample-averaged $PM_{2.5}$ concentration) using Equation 2 assumes that the relationship between the concentration reported by the low-cost monitor and the true concentration: (1) is linear over the range recorded and (2) remains constant during the sample. If these assumptions do not hold (e.g., due to variations in particle sources and/or ambient relative humidity over time), an alternate function may be needed to relate the gravimetric correction factor ($CF$) to real-time $PM_{2.5}$ concentrations reported by the low-cost monitor.

In this study, we first compared $PM_{2.5}$ mass concentrations reported by PurpleAir monitors to gravimetric measurements (collected using integrated filter samples and a tapered element oscillating microbalance [TEOM]) in a laboratory aerosol chamber. We then deployed 19 PurpleAir monitors for ~45 days in an outdoor network with 15 sites that spanned a ~3 × 3 km area in downtown Fort Collins, CO, USA. During the field deployment, each PurpleAir was co-located with a compact, portable, battery-powered filter sampler. At a subset of 4 sites, the PurpleAir monitors and portable filter samplers were co-located with conventional 16.7 $L\, min^{-1}$

filter samplers. Using data from the laboratory and field, we aimed to answer the following questions: (1) does the PurpleAir respond linearly to $PM_{2.5}$ concentrations ranging from 0 to 75 μg $m^{-3}$; (2) does a laboratory-derived gravimetric correction factor translate to the field; (3) does correcting for ambient relative humidity alone improve the accuracy of the ambient $PM_{2.5}$ concentrations reported by the PurpleAir monitors; (4) does a gravimetric correction factor derived from a 72-hour co-location with a portable filter sampler improve the accuracy of the ambient $PM_{2.5}$ concentrations reported by the PurpleAir monitors; (5) how does the frequency of field calibration using the portable filter sampler affect PurpleAir accuracy; and (6) do $PM_{2.5}$ concentrations measured using the portable and conventional filter samplers agree?

## 2. METHODS

The low-cost monitor evaluated in this study was the PurpleAir (PA-II-SD). Each PurpleAir monitor operated on firmware version 2.50i and contained two PMS5003 sensors (Plantower, Beijing, China). The PMS5003 estimates particle mass concentrations by measuring the amount of ~680 nm light scattered at ~90° (Sayahi et al., 2019). The PMS5003 reports mass concentrations of $PM_1$, $PM_{2.5}$, and $PM_{10}$. Each concentration is reported two ways: with a correction factor of one (e.g., $PM_{2.5}$ CF=1) and with a proprietary “atmospheric” correction factor (e.g., $PM_{2.5}$ ATM). According to the PMS5003 manual, CF=1 values should be used for indoor monitoring and ATM values should be used for atmospheric monitoring.

### 2.1. Laboratory Evaluations

The PurpleAir monitors were evaluated in a laboratory aerosol chamber before and after the field deployment to: (1) determine whether the monitors responded linearly to $PM_{2.5}$ concentrations ranging from 0 to 75 μg $m^{-3}$ and (2) obtain a laboratory-derived gravimetric correction factor for the $PM_{2.5}$ concentrations reported by the PurpleAirs. We also assessed whether or not the

response of the monitors to a fixed aerosol changed over the duration of the field deployment. Seventeen of the 19 PurpleAir monitors deployed in the field were evaluated before deployment and all 19 were evaluated after deployment. Aside from the difference in the number of monitors, the same methods were used for both of these laboratory evaluations.

Later, a third laboratory evaluation was conducted in which $PM_{2.5}$ concentrations reported by six of the 19 PurpleAir monitors were compared to $PM_{2.5}$ concentrations reported by eight independent Plantower PMS5003 sensors. The eight PMS5003 sensors were connected to NUCLEO-F767ZI development boards (STMicroelectronics, Geneva, Switzerland) and data from the sensors were logged every 3 seconds using Megunolink (Number Eight Innovation Limited, Hamilton, New Zealand). All other methods were the same as for the pre- and post-deployment laboratory evaluations.

#### *2.1.1. Data Collection*

The aerosol chamber had dimensions of 1.2 m × 0.85 m × 0.75 m (Tryner et al., 2019c). Flow into the chamber was provided by HEPA-filtered compressed air. Air inside the chamber was mixed using a small fan.

The calibration aerosol was NIST SRM 1648a Urban Particulate Matter (National Institute of Standards and Technology, Gaithersburg, MD, USA). The Urban PM was suspended in deionized water at a concentration of 0.5 g $L^{-1}$ and aerosolized using a six-jet Collison nebulizer (CH Technologies, Westwood, NJ, USA). Air flow through the nebulizer was provided by filtered compressed air and controlled using a solenoid valve. The compressed air pressure, interval at which the valve opened and closed, and the flow rate of dilution air into the chamber were varied over the course of the experiment to achieve seven unique $PM_{2.5}$ concentrations inside the chamber ranging up to 75 µg $m^{-3}$. A zero point was repeated at the start and end of the experiment.

The maximum concentration was 75 μg m$^{-3}$ because we did not expect to observe higher ambient concentrations during field deployment in Fort Collins. The concentration in the chamber was changed approximately once per hour, and each experiment lasted ~10 hours.

Measurements taken by PurpleAir monitors were compared to measurements taken by a TEOM (1405 TEOM, ThermoFisher Scientific, Waltham, MA, USA) and three integrated gravimetric samples. The PurpleAir monitors and the TEOM logged data at 80-s and 60-s intervals, respectively. Integrated gravimetric samples were collected on 47-mm PTFE filters (Tisch Scientific, North Bend, OH, USA) installed behind 16.7 L min$^{-1}$, 2.5 μm cyclones (URG Corp., Chapel Hill, NC, USA). Flow rates through the cyclones were checked before and after sampling (using a triCal, BGI Inc., Waltham, MA, USA or AFS-20, Alicat Scientific, Tucson, AZ, USA). Filters were pre- and post-weighed on a balance with 1 μg resolution (XS3DU, Mettler Toledo, Columbus, OH, USA). During the laboratory evaluation conducted after field deployment, only two of the filter samples were successful due to an error in setting the flow rate for the third.

#### *2.1.2. Data Analysis*

At each 80-s interval, the two concentrations labeled "$PM_{2.5}$ ATM" in each PurpleAir log file (i.e., the concentrations reported by PMS5003 sensors "A" and "B") were compared. The data recorded at a given timestamp were retained for analysis if these concentrations differed by ≤ 15 μg m$^{-3}$. For both the pre- and post-deployment evaluations, 98% of the 80-s readings were retained.

To calculate the laboratory correction factor ($CF$) for a PurpleAir monitor, the "$PM_{2.5}$ ATM" concentrations reported by sensors A and B were time-averaged (independently) over the duration of the experiment. Then, the time-averaged concentrations reported by sensors A and B were averaged. Finally, the average $PM_{2.5}$ concentration measured using the 16.7 L min$^{-1}$ filter samples was divided by this average "$PM_{2.5}$ ATM" concentration. This procedure was repeated to

calculate a correction factor for each PurpleAir monitor. The laboratory correction factors calculated during the pre- and post-deployment evaluations were compared using a paired t-test (Reimann et al., 2008). The assumption of normality was evaluated using the Shapiro-Wilk test.

The linearity of the PurpleAir response was evaluated using the data associated with each 1-hour long concentration point. For each concentration point, the transient PurpleAir and TEOM data recorded during the first 15 minutes were ignored. These data were ignored because, after the chamber settings (e.g., solenoid valve open/closed interval, dilution air flow rate) were adjusted, it took approximately 15 minutes for the chamber to reach a new steady-state $PM_{2.5}$ concentration and for the TEOM data to reflect that concentration. The steady-state $PM_{2.5}$ concentrations recorded by each instrument during the remaining ~45 minutes were then time-averaged (see **Figure S1**). At each concentration point, the time-averaged "$PM_{2.5}$ ATM" concentrations reported by sensors A and B in each PurpleAir monitor were averaged. The same was done for the time-averaged "$PM_{2.5}$ CF=1" concentrations reported by sensors A and B. The raw and concentration-point averaged data from the pre- and post-deployment laboratory evaluations are available through an online digital repository (Tryner et al., 2019b).

A linear model of the form shown in Equation 3 was fit to the time-averaged data, using weighted least-squares regression, with the concentration measured by the TEOM ($\bar{c}_{TEOM}$; μg $m^{-3}$) as the predictor and the "$PM_{2.5}$ ATM" concentration reported by the PurpleAir ($\bar{c}_{PA}$; μg $m^{-3}$) as the outcome.

$$\bar{c}_{PA} = \beta_0 + \beta_1 \bar{c}_{TEOM} + \epsilon \quad (3)$$

where $\beta_0$ was the y-intercept, $\beta_1$ was the slope, and $\epsilon$ was the random error.

The weight for each concentration point ($w_i$) was calculated as

$$w_i = 1/s_i^2 \quad (4)$$

where $s_i$ was the sample standard deviation of the time-averaged "$PM_{2.5}$ ATM" concentrations reported by the $J$ co-located PurpleAir monitors during concentration point $i$. An F-test was conducted to determine whether there was a significant lack of fit between the data and the linear model given by Equation 3 (Analytical Methods Committee, 1994).

Additionally, the precision of the PurpleAir monitors was quantified at each nonzero concentration point using relative standard deviation (RSD) (Nees, 1993):

$$RSD_i = \frac{s_i}{(1/J)\sum \bar{c}_{ij}} \quad (5)$$

where $\bar{c}_{ij}$ was the time-averaged "$PM_{2.5}$ ATM" concentration ($\mu g\, m^{-3}$) reported by monitor $j$ during concentration point $i$.

## 2.2. Field Deployment

### *2.2.1. Break-in Period*

All 19 PurpleAir monitors were co-located on a single roof in Fort Collins from 22 August through 12 October 2018 (**Figure S2**). The goal of this initial co-location was to "break-in" the monitors so that the performance captured during network deployment would be representative of monitors that had been deployed outside for several weeks, as opposed to brand-new monitors.

### *2.2.2. Network Deployment*

Aerosol Sampler Plus Environmental Node (ASPEN) boxes were deployed at 15 sites in downtown Fort Collins between 22 October and 06 December 2018 (**Figure 1**). The ASPEN box was a compact, weather-proof unit that contained an Ultrasonic Personal Aerosol Sampler (UPAS, Access Sensor Technologies, Fort Collins, CO, USA), a stand-alone Plantower PMS5003 sensor, a PurpleAir monitor, and the batteries needed to power all components for 72 hours (**Figure 2**). The UPAS was used to sample $PM_{2.5}$ onto 37-mm PTFE filters (PT37P-PF03,

Measurement Technology Laboratories, Minneapolis, MN, USA). Each UPAS was equipped with a cyclone inlet that had a 2.5 µm cutpoint when operated at 2 L $min^{-1}$ (Volckens et al., 2017).

ASPEN boxes were installed approximately 3 m above the ground at sites E-O, on the roofs of the buildings at sites A-C, and approximately 2 m above the ground at site D. Duplicate ASPEN boxes were installed at sites A–D along with conventional 16.7 L $min^{-1}$ $PM_{2.5}$ filter samplers (**Figure S3**). A GRIMM Environmental Dust Monitor (EDM) 180 was located at site B (EPA AQS site 08-069-0009). The GRIMM EDM 180 is an optical aerosol spectrometer that has been designated as a Federal Equivalent Method for $PM_{2.5}$ by the U.S. Environmental Protection Agency (Spielvogel et al., 2009).

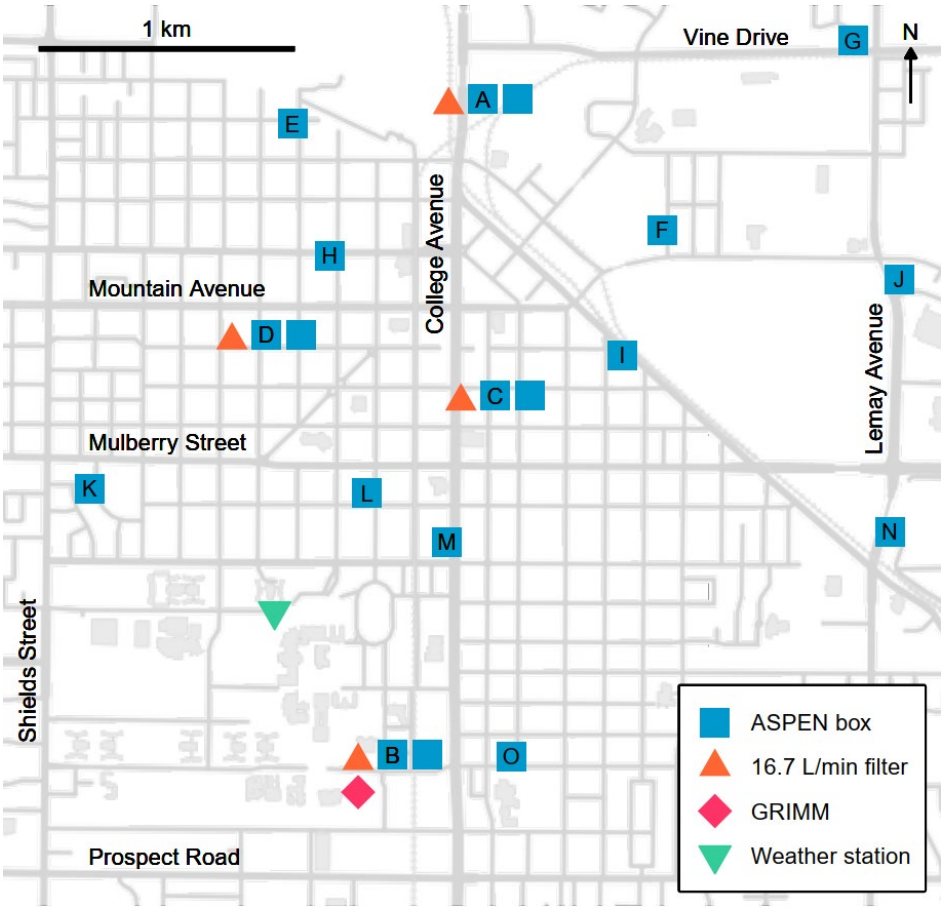



**Figure 1.** A map illustrating the co-located monitors at each of the 15 sites (A-O). The labeled ASPEN box represents the precise site location. Each ASPEN box contained a UPAS (which sampled $PM_{2.5}$ onto a 37-mm filter at 2 L $min^{-1}$) and a PurpleAir monitor. The location of the weather station on the Colorado State University (CSU) main campus is also shown.

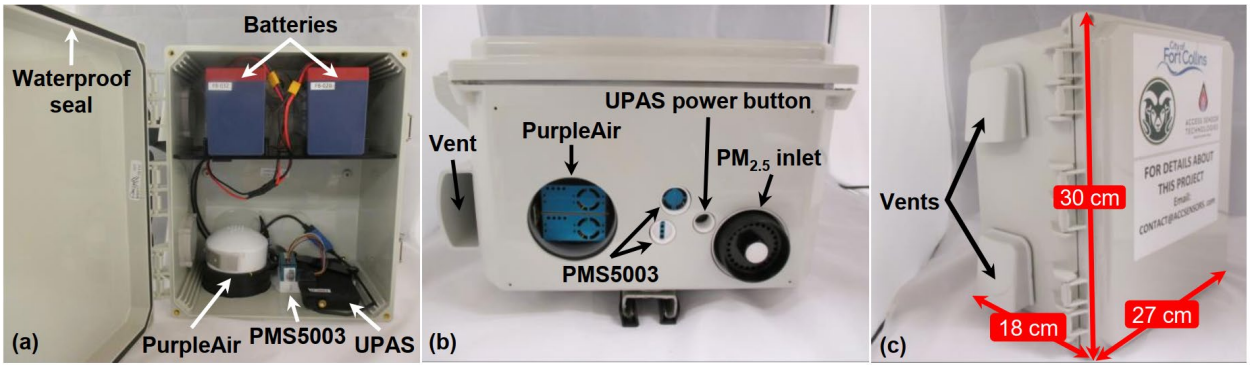



**Figure 2.** Components and dimensions associated with the ASPEN (Aerosol Sampler Plus Environmental Node) box. Each ASPEN box contained: a UPAS (which was used to collect filter samples of $PM_{2.5}$ at a flow rate of 2 L $min^{-1}$), a stand-alone Plantower PMS5003 sensor, and the paired PurpleAir monitor. Internal components are shown in panel (a), the inlets on the base of the box are shown in panel (b), and overall dimensions are shown in panel (c).

Each 16.7 L $min^{-1}$ filter sampler consisted of a standard EPA louvered $PM_{10}$ inlet (SSI2.5, Mesa Labs, Lakewood, CO, USA) followed by a $PM_{2.5}$ cyclone (URG Corp., Chapel Hill, NC, USA) and a cartridge containing a 47-mm PTFE filter (Tisch Scientific, North Bend, OH, USA). Air was pulled through the inlet, cyclone, and filter by a vacuum pump (86R142-P001B-N270X, Gast, Benton Harbor, MI, USA). The 16.7 L $min^{-1}$ flow rate was maintained by a mass flow controller (MCRW-20SLPM-D/5M, Alicat Scientific, Tucson, AZ, USA).

We collected 242 UPAS filter samples during network field deployment. Of these 242, 66 samples starting on or after 02 November 2018 were paired with 16.7 L $min^{-1}$ filter samples. Filter samples were started on Mondays and Fridays. The filter-based monitors typically sampled for 72-hour periods to ensure that sufficient mass accumulated on the 2 L $min^{-1}$ UPAS filter samples. Due to the holiday on 22 November, the 19 UPAS filter samples and three 16.7 L $min^{-1}$ filter samples that started on 19 November ran for 48 hours. We use the term “sample-averaged” to refer to quantities that were time-averaged over the 72- or 48-hour filter sample period.

### *2.2.3. Quality Assurance*

At each 80-s interval, the two concentrations labeled "$PM_{2.5}$ ATM" in each PurpleAir log file (i.e., the concentrations reported by PMS5003 sensors "A" and "B") were compared. Because of the low $PM_{2.5}$ concentrations measured in the field, the data recorded at a given timestamp were only retained if these two concentrations differed by ≤ 5 μg m$^{-3}$. The data reported by a PurpleAir monitor during a given filter sample period were only used in further analyses if (a) the data were available (i.e., recorded and retained) for at least 80% of the period (214/242 samples) and (b) the sample-averaged "$PM_{2.5}$ ATM" concentrations reported by sensors A and B were both greater than 1 μg m$^{-3}$ (211/214 samples; see SI Section S1.3).

The 2 L min$^{-1}$ filter samples collected using the UPAS were retained for analysis if: (1) the sampler ran for at least 80% of the target sample period (208/242 samples), (2) the mass accumulated on the filter was above the study limit of detection (197/208 samples), and (3) the $PM_{2.5}$ concentration derived from the sample was not more than 3× the maximum derived from the 30 successful 16.7 L min$^{-1}$ filter samples (194/197 samples). Overall, 194 samples were retained after applying these criteria. The LOD was calculated from 20 field blanks ($LOD = 3s_{blank} =$ 10 μg) and translated into an average mass concentration of 1.2 μg m$^{-3}$ over a 72-hour period. A sample-averaged $PM_{2.5}$ concentration that was more than 3× the maximum derived from the 16.7 L min$^{-1}$ filter samples (3 × 8.3 μg m$^{-3}$ = 25 μg m$^{-3}$) was assumed to result from filter mishandling.

The 16.7 L min$^{-1}$ filter samples were retained for analysis if: (1) the sampler ran for at least 80% of the target sample period (33/33 samples), (2) the ending flow rate was within 20% of 16.7 L min$^{-1}$ (30/33 samples), and (3) the mass accumulated on the filter was above the study LOD (31 μg; 30/30 samples). The LOD was calculated from 20 field blanks and translated into a mass concentration of 0.44 μg m$^{-3}$ over a 72-hour period.

### *2.2.4. Data Analysis*

Six different corrections for the PurpleAir "$PM_{2.5}$ ATM" concentrations were evaluated: (1) none ($c_{PA}{}' = c_{PA}$), (2) laboratory filter-based, (3) relative humidity-based, (4) monthly filter-based, (5) weekly filter-based, and (6) concurrent filter-based.

The relative humidity ($RH$) correction shown in Equation 1 was used (Chakrabarti et al., 2004). Relative humidity corrections calculated using Equation 1 were similar to hygroscopic growth factors calculated as described by Malings et al. (2019) using estimates of local particle composition and size. Equation 1 was used because it did not require local particle property data. This correction was applied to each 80-s "$PM_{2.5}$ ATM" concentration reported by each sensor in the PurpleAir monitor ($c_{PA}$), using the corresponding hourly ambient RH measured at the Colorado State University (CSU) main campus weather station (**Figure 1**). RH-corrected 80-s "$PM_{2.5}$ ATM" concentrations ($c_{PA}{}'$) were then time-averaged over the duration of the filter sample. Finally, the sample-averaged RH-corrected "$PM_{2.5}$ ATM" concentrations from sensors A and B were averaged for comparison to the $PM_{2.5}$ concentration derived from the co-located 16.7 L $min^{-1}$ filter sample.

Monthly, weekly, and concurrent filter-based correction factors were calculated as shown in Equation 2 with $c_{filter}$ equal to the $PM_{2.5}$ concentration derived from the paired UPAS filter sample and $\bar{c}_{PA}$ equal to the sample-averaged "$PM_{2.5}$ ATM" concentration reported by the PurpleAir (with no relative humidity correction). The sample-averaged "$PM_{2.5}$ ATM" concentrations reported by sensors A and B in the PurpleAir monitor were averaged to obtain $\bar{c}_{PA}$. The correction factor calculated for a given PurpleAir/UPAS pair was applied only to that PurpleAir monitor.

Monthly correction factors were calculated using UPAS filter samples that ended on 01 November and 29 November, respectively, and extended prospectively in time. Correction factors calculated from paired samples ending on 01 November were used to correct PurpleAir data collected

between 02 and 29 November. Correction factors calculated from paired samples ending on 29 November were used to correct PurpleAir data collected between 30 November and 06 December (see SI section S1.3 and **Figure S4**).

Weekly correction factors were calculated using UPAS filter samples that ended every Thursday between 01 and 29 November (except 22 November). The filter samples that ended each Thursday were used to calculate correction factors for PurpleAir data collected over the following 7 days. The filter samples that ended on Wednesday, 21 November were used to calculate correction factors for PurpleAir data collected between 23 and 29 November.

Concurrent correction factors were calculated using all UPAS filter samples. For example, filter samples collected between 02 and 05 November were used to calculate correction factors for PurpleAir data collected between 02 and 05 November.

The accuracy of the sample-averaged PurpleAir "$PM_{2.5}$ ATM" concentration after correction using each of the six approaches described above ($\bar{c}_{PA}{}'$; μg m$^{-3}$) was evaluated by calculating the bias and percent bias relative to a reference $PM_{2.5}$ concentration ($c_{ref}$; μg m$^{-3}$):

$$Bias = \bar{c}_{PA}{}' - c_{ref} \quad (6)$$

$$Bias\ (\%) = \frac{\bar{c}_{PA}{}' - c_{ref}}{c_{ref}} \quad (7)$$

where $c_{ref}$ was measured using a co-located 16.7 L min$^{-1}$ filter sample.

To estimate the fraction of the variance in the log-transformed concurrent gravimetric correction factor that was explained by differences (a) between dates and (b) between monitors, we fit a two-way random effects model to the data using the 'lme4' package in R (Bates et al., 2019):

$$(CF_{ij}) = \mu + r_i + c_j + \epsilon_{ij} \quad (8)$$

where $CF_{ij}$ was the concurrent gravimetric correction factor for PurpleAir monitor $j$ on sample date $i$, $\mu$ was the population mean, $r_i$ was the random effect associated with sample date $i$, $c_j$ was the random effect associated with PurpleAir monitor $j$, and $\epsilon_{ij}$ represented the random residual effects. The effects $r_i$, $c_j$, and $\epsilon_{ij}$ were assumed to be independent and normally distributed with means of zero and variances of $\sigma_r^2$, $\sigma_c^2$, and $\sigma_\epsilon^2$, respectively (McGraw and Wong, 1996). Intraclass correlation coefficients (ICCs) for sample date and PurpleAir monitor were defined as shown in Equations S3 and S4, respectively, and calculated using the 'sjstats' package (Lüdecke, 2019). We calculated 95% confidence intervals for the ICCs using parametric bootstrap resampling (n=1000).

Agreement between the UPAS filter samples and the conventional 16.7 L $min^{-1}$ filter samples was assessed using: (1) the bias of the $PM_{2.5}$ concentration derived from the UPAS filter sample relative to the $PM_{2.5}$ concentration derived from a co-located 16.7 L $min^{-1}$ filter sample, (2) a Bland-Altman plot, and (3) Deming regression. Bias was calculated using Equations 6 and 7 with the $PM_{2.5}$ concentration derived from the UPAS filter sample ($c_{filter}$) in place of $\bar{c}_{PA}{}'$. Deming regression, which assumes that both the x- and y- variables were measured with error, was performed using the 'deming' package in R (Therneau, 2018). Constant and equal coefficients of variation were assumed for the x- and y- error terms.

Data recorded during the field deployment are available through an online digital repository (Tryner et al., 2019b). This repository contains: (1) the raw data logged by the PurpleAir monitors, (2) the raw data recorded by the UPAS and standalone PMS5003 sensors installed in the ASPEN boxes, and (3) a file summarizing the sample-averaged $PM_{2.5}$ concentrations and correction factors derived from the raw data.

## 3. RESULTS AND DISCUSSION

### 3.1. LABORATORY EVALUATIONS

The $PM_{2.5}$ concentrations reported by the PurpleAir monitors during the laboratory evaluations are compared to the concentrations reported by the TEOM in **Figure 3**. At concentrations below 30 μg $m^{-3}$, the "$PM_{2.5}$ CF=1" and "$PM_{2.5}$ ATM" values were equal (**Figure S5**). At concentrations above 30 μg $m^{-3}$, the concentration labeled "$PM_{2.5}$ CF=1" in the PurpleAir log file did not increase linearly with the concentration recorded by the TEOM. Kelly et al. (2017) and Sayahi et al. (2019) reported a similar nonlinear response for "$PM_{2.5}$ CF=1" concentrations reported by PurpleAir monitors measuring ambient aerosols in Salt Lake City, UT, USA.

The concentration labeled "$PM_{2.5}$ ATM" in the PurpleAir log file did appear to increase linearly with the TEOM-reported concentration (**Figure 3**). Both before and after the field deployment, a linear model (Equation 3) explained 99% of the variance in the PurpleAir "$PM_{2.5}$ ATM" values. However, the F-test indicated that the lack of fit between the linear model and the data was greater than would be expected due to experimental uncertainty alone (Pre-deployment: $F_{(I-2)/(IJ-I)} = 3.09$, $F_C = 2.17$; Post-deployment: $F_{(I-2)/(IJ-I)} = 42.4$, $F_C = 2.16$). The weighted residuals for the linear model are shown in **Figure S6**. Whether the discrepancy between the linear model and the experimentally-observed relationship is important from a practical standpoint depends on the error that results from assuming a linear relationship between the PurpleAir "$PM_{2.5}$ ATM" concentration and a reference concentration in a given application (Analytical Methods Committee, 1994).

Results from the third laboratory evaluation in which readings from six PurpleAir monitors were compared to readings from eight independent Plantower PMS5003 sensors indicated that the $PM_{2.5}$ CF=1 values output by the PMS5003 sensors were labeled "$PM_{2.5}$ ATM" in the PurpleAir log files (firmware version 2.50i). Similarly, the $PM_{2.5}$ ATM values output by the PMS5003 sensors

were labeled “$PM_{2.5}$ CF=1” in the PurpleAir log files (see **Figure S7**). Concentrations referred to as “$PM_{2.5}$ ATM" throughout this manuscript refer to values labeled “$PM_{2.5}$ ATM" in the PurpleAir log files (which are actually the $PM_{2.5}$ CF=1 values output by the PMS5003 sensors).

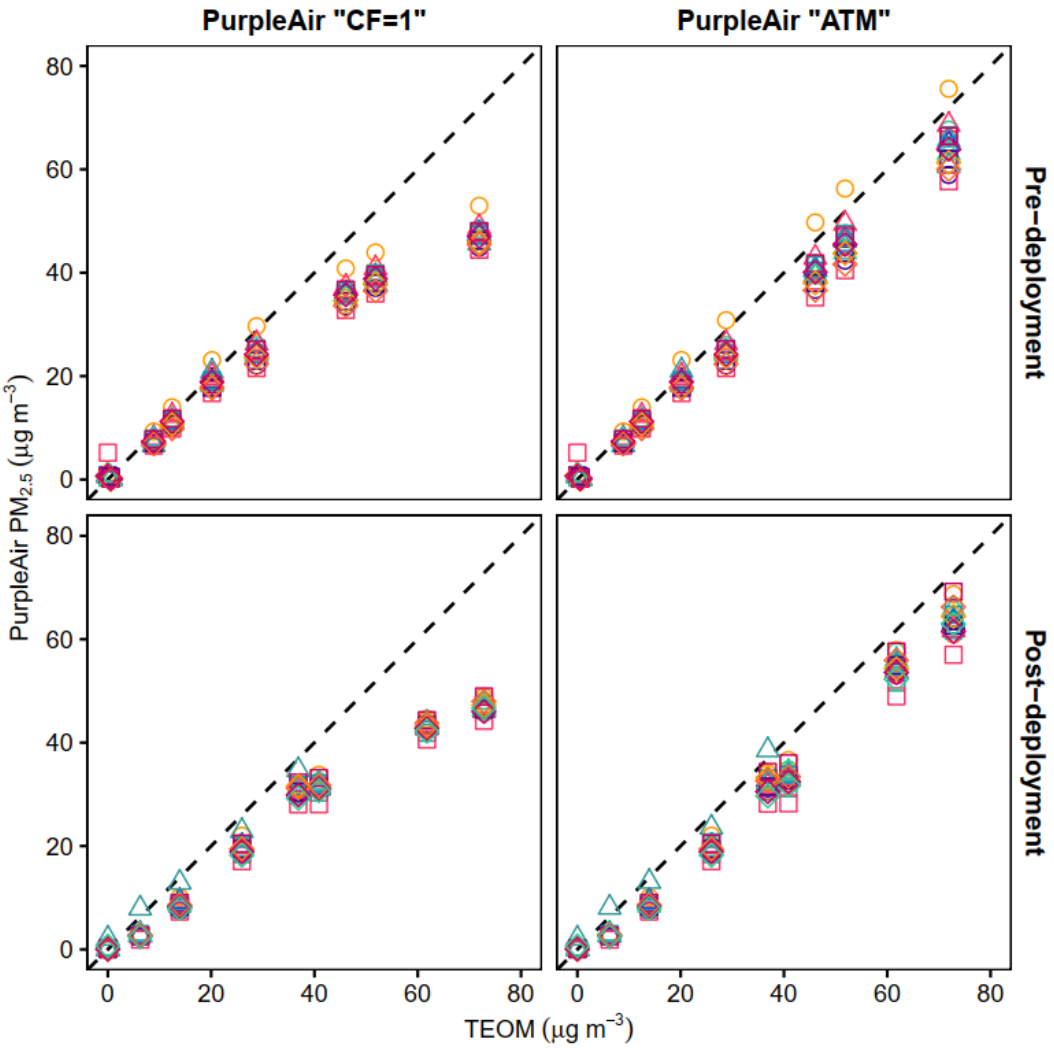



**Figure 3.** A comparison of the time-averaged $PM_{2.5}$ concentrations reported by the TEOM and the PurpleAir monitors before (top) and after (bottom) the field deployment. The concentrations labeled as “$PM_{2.5}$ CF=1” and “$PM_{2.5}$ ATM” in the PurpleAir log files (firmware version 2.50i) are shown in the left and right panels. Dashed lines represent y=x.

The average laboratory-derived correction factor ($CF = c_{filter}/\bar{c}_{PA}$) for the 17 PurpleAir monitors evaluated pre-deployment was 1.06 (calculated using the concentrations labeled “$PM_{2.5}$ ATM”). The average laboratory-derived correction factor for the 19 PurpleAir monitors evaluated post-deployment was 1.19. A paired t-test using the 17 PurpleAirs evaluated pre- and post-deployment

indicated that the correction factors were higher post-deployment than pre-deployment (mean difference = 0.11, p = 0.005). This change in correction factor might have resulted from PM accumulating on the optics within the PMS5003 sensors over the course of the break-in and network field deployment periods.

During our laboratory evaluations, the relationship between the NIST Urban PM concentrations reported by the TEOM and the "$PM_{2.5}$ ATM" values reported by PurpleAirs was approximately 1:1, with the PurpleAirs slightly underestimating the TEOM-reported concentration. In contrast, multiple studies have reported that the "$PM_{2.5}$ ATM" values reported by PurpleAir (PA-II) monitors often overestimate ambient $PM_{2.5}$ concentrations (Magi et al., 2019; Malings et al., 2019). Our results suggest that the Plantower sensors are calibrated using an aerosol with a size distribution and composition similar to that of the NIST Urban PM. The nebulized NIST Urban PM had a geometric mean diameter of 0.042 μm and a geometric standard deviation of 2.05 (**Figure S8**).

The relative standard deviation (RSD) of the "$PM_{2.5}$ ATM" values reported by the PurpleAir monitors decreased as concentration increased (**Figure 4**). The RSD was less than 20% at concentrations ≥ 9 $\mu g\, m^{-3}$ and less than 10% at concentrations ≥ 20 $\mu g\, m^{-3}$. The RSDs calculated before and after field deployment were similar.

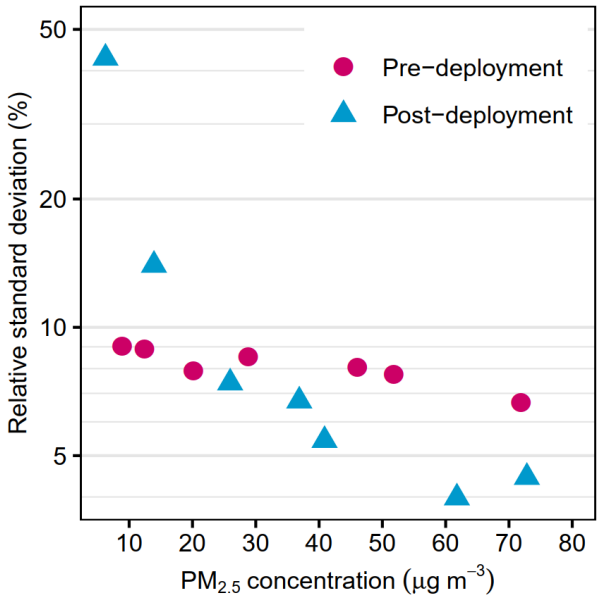



**Figure 4.** The relative standard deviation (RSD) of the time-averaged "$PM_{2.5}$ ATM" concentrations reported by the PurpleAir monitors vs. the $PM_{2.5}$ concentration measured by the TEOM. Different marker shapes represent data from the laboratory evaluations conducted pre- and post-deployment.

### 3.2. Field Deployment

The median $PM_{2.5}$ concentration derived from the 30 successful 16.7 L $min^{-1}$ filter samples collected at sites A-D between 02 November and 06 December was 5.1 µg $m^{-3}$ (range = 1.5 to 8.3 µg $m^{-3}$). During this time period, sample-averaged "$PM_{2.5}$ ATM" concentrations reported by PurpleAir monitors ranged from 1.3 to 19 µg $m^{-3}$. The median sample-averaged RH reported at the weather station was 63% (range = 47% to 77%). The hourly RH reported at the weather station ranged from 10% to 100% and was below 75% for 69% of the time (see **Figures S11-S13** for additional weather data).

The sample-averaged uncorrected "$PM_{2.5}$ ATM" concentrations reported by PurpleAir monitors at sites A-D were positively correlated with the concentrations derived from 16.7 L $min^{-1}$ filter samples (Spearman's ρ = 0.80; **Figure 5**). Uncorrected sample-averaged "$PM_{2.5}$ ATM" concentrations overestimated the filter-derived concentration for 55/58 pairs. Prior field studies

also reported that "$PM_{2.5}$ ATM" values reported by PA-II monitors tend to overestimate ambient $PM_{2.5}$ concentrations (Magi et al., 2019; Malings et al., 2019).

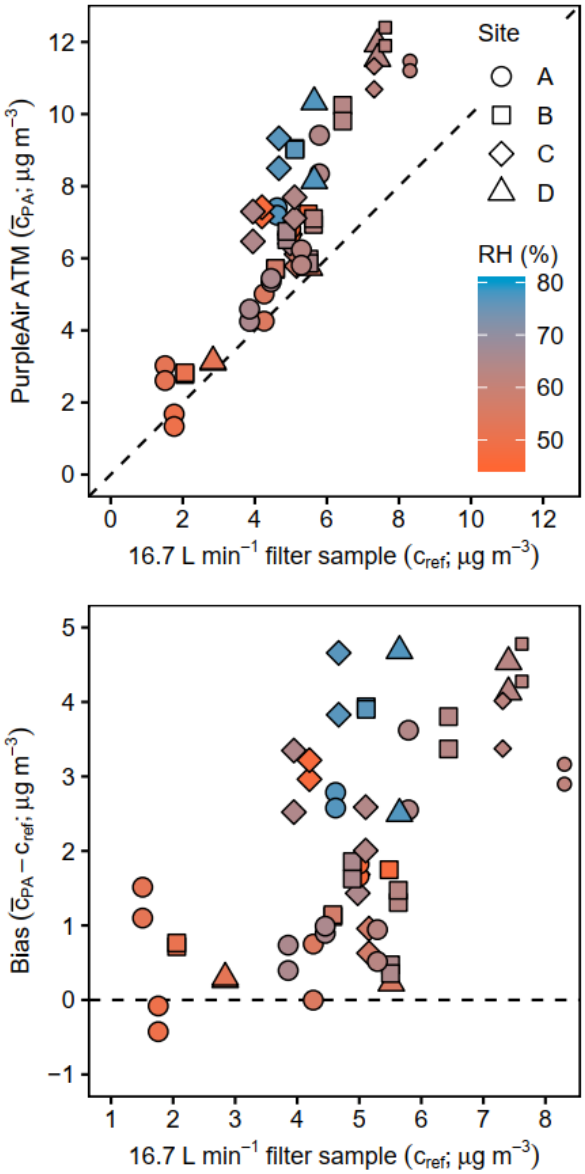



**Figure 5.** Top: A comparison of the sample-averaged $PM_{2.5}$ concentrations derived from the 16.7 L min$^{-1}$ filter samples and reported by the PurpleAir monitors ("$PM_{2.5}$ ATM" in the PurpleAir log file). The diagonal line represents y=x. Bottom: A Bland-Altman plot comparing the same two concentrations (with the filter-derived concentration on the x-axis). The marker color represents the average relative humidity (RH) reported at the main campus weather station during the sample period. Smaller markers represent the 48-hour samples collected between 19-21 November.

The bias of sample-averaged uncorrected "$PM_{2.5}$ ATM" concentrations reported during each sample period is shown in **Figure 6** (with bias in %) and **Figure S14** (with bias in $\mu g\, m^{-3}$). The median bias varied with sample date, but did not increase or decrease continuously during the field deployment. Changes in the response of the PurpleAir monitors over the course of the field deployment (as evidenced by the increase in the laboratory correction factor between the pre- and post-deployment evaluations) did not explain the variation in PurpleAir bias observed during field deployment. The pre- and post-deployment laboratory correction factors were 1.06 and 1.19, respectively—indicating that the PurpleAir monitors were biased low. Conversely, the concurrent gravimetric correction factor calculated in-field from paired PurpleAir monitors and UPAS filter samples was less than one for 140/175 samples—indicating that the PurpleAir monitors were biased high. These results suggest that the variations in PurpleAir bias observed in the field resulted from variations in the properties of the ambient aerosols measured by the monitors, as opposed to variations in the response of the monitors to an aerosol with fixed properties.

The second highest median PurpleAir bias was observed for samples collected 12-15 November, which was the sample period with the highest 72-hour average $PM_{2.5}$ concentration (**Figures 6 and S14**). The median bias was highest for samples collected 16-19 November, which was the period with the highest sample-averaged RH. The median bias was lowest for samples collected 23-26 November. These samples spanned a holiday weekend; consequently, there might have been less activity downtown than usual. This period had the lowest sample-averaged $PM_{2.5}$ concentration and the lowest sample-averaged RH. The median bias was higher for the subsequent 26-29 November sample period, when the 72-hour average $PM_{2.5}$ concentration was higher but the 72-hour average RH remained low. Overall, bias was likely influenced by the ambient RH and $PM_{2.5}$ concentration, as well as other factors.

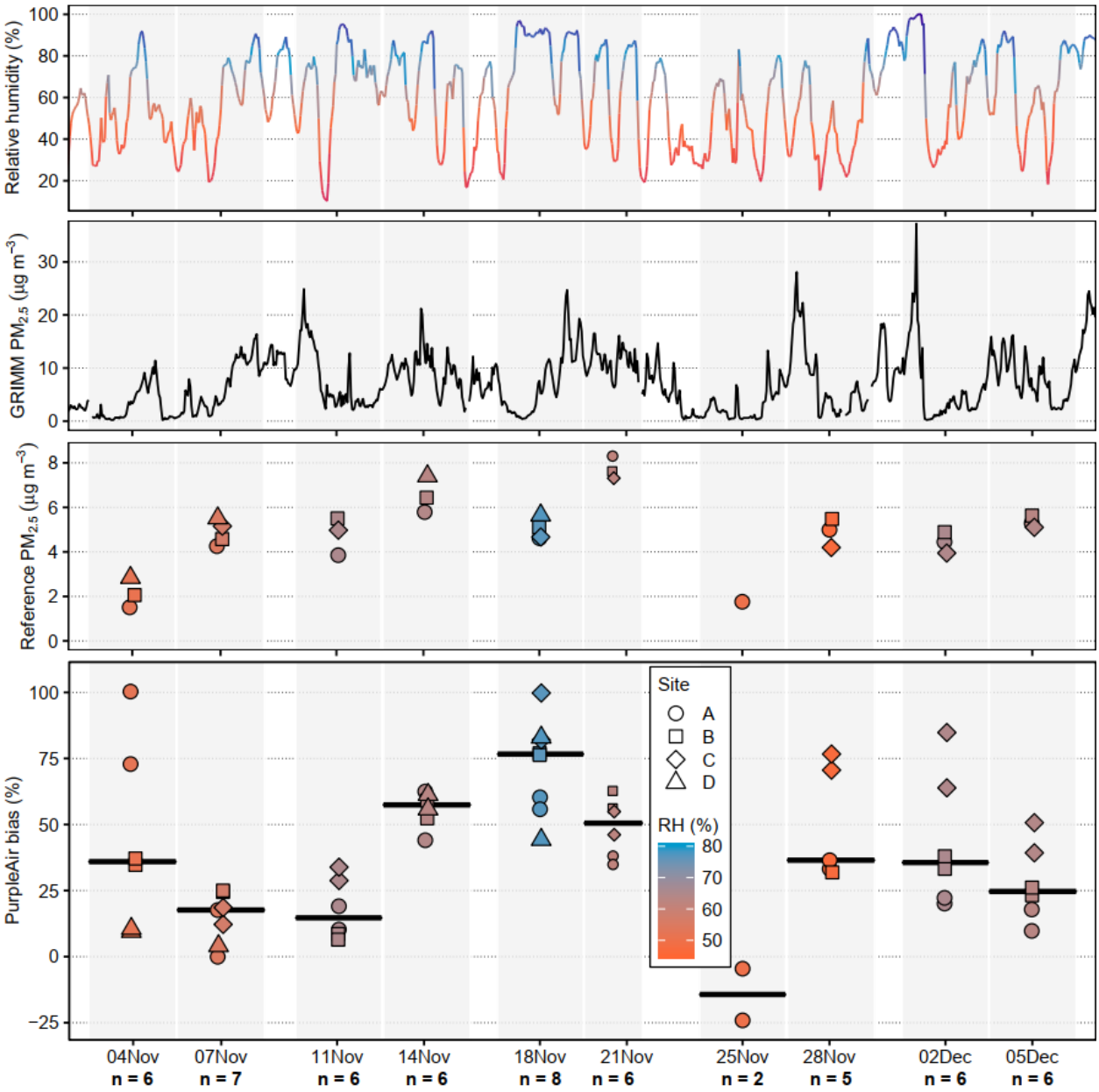



**Figure 6.** Top to bottom: Hourly relative humidity (RH) measured at the CSU main campus weather station, one-hour average $PM_{2.5}$ concentrations reported by the GRIMM EDM Model 180 at Site B, $PM_{2.5}$ concentrations derived from 16.7 L $min^{-1}$ "reference" filter samples, and percent bias of the uncorrected 72- or 48-hour average "$PM_{2.5}$ ATM" concentrations reported by PurpleAir monitors (relative to the concentrations derived from co-located 16.7 L $min^{-1}$ filter samples). Gray boxes represent the duration of each sample period. Horizontal black lines represent the median bias for each period. The number of valid comparisons obtained during each period is listed below the x-axis (maximum = 8). Smaller markers are used for the 48-hour samples collected on 19-21 November. The marker color represents the average RH reported at the weather station during the sample period. Note that the color scale for the hourly RH extends beyond the range shown in the legend.

With no correction, or with a laboratory-derived correction factor of 1.06, PurpleAir monitors tended to overestimate the $PM_{2.5}$ concentration measured using 16.7 L $min^{-1}$ filter samples [median bias = 1.7 µg $m^{-3}$ (or 37%) for no correction and 2.1 µg $m^{-3}$ (or 45%) for laboratory correction; **Table 1**, **Figure 7,** and **Figure S15**]. For uncorrected PurpleAir monitors, 55% of the samples had an absolute bias ≤ 2 µg $m^{-3}$, and 24% of the samples had an absolute bias ≤ 20%.

**Table 1.** Bias of sample-averaged $PM_{2.5}$ concentrations reported by PurpleAir monitors, relative to concentrations measured using co-located 16.7 L $min^{-1}$ filter samples, after applying different corrections.

| Correction factor type | None | Lab | Relative humidity | Monthly† | Weekly† | Concurrent† |
|---|---|---|---|---|---|---|
| Median bias (µg $m^{-3}$) | 1.7 | 2.1 | -0.54 | -0.12 | -0.26 | -0.01 |
| Min., max. bias (µg $m^{-3}$) | -0.4, 4.8 | -0.3, 5.5 | -3.0, 1.9 | -3.8, 3.4 | -4.5, 4.7 | -4.0, 2.6 |
| $25^{th}$, $75^{th}$ percentile (µg $m^{-3}$) | 0.80, 3.3 | 1.1, 3.8 | -1.1, 0.04 | -1.1, 1.0 | -1.1, 0.65 | -0.29, 0.41 |
| Number (%) of samples with \|bias\| ≤ 2 µg $m^{-3}$ | 32/58 (55%) | 26/58 (45%) | 54/58 (93%) | 37/48 (77%) | 39/46 (85%) | 40/46 (87%) |
| Median bias (%) | 37 | 45 | -15 | -2 | -5 | 0 |
| Min., max. bias (%) | -24, 100 | -20, 112 | -66, 44 | -65, 72 | -65, 81 | -61, 83 |
| $25^{th}$, $75^{th}$ percentile (%) | 19, 60 | 27, 70 | -23, 0 | -21, 23 | -24, 14 | -7, 14 |
| Number (%) of samples with \|bias\| ≤ 20% | 14/58 (24%) | 11/58 (19%) | 38/58 (66%) | 22/48 (46%) | 25/46 (54%) | 33/46 (72%) |

*† indicates periodic correction using the co-located UPAS filter sample*

The RH-based correction reduced the median bias to -0.54 µg $m^{-3}$ (or -15%). This correction produced the largest fraction of samples with an absolute bias of ≤ 2 µg $m^{-3}$ (93%) and increased the fraction of samples with an absolute bias of ≤ 20% to 66%. Malings et al. (2019) found that applying a similar RH correction to hourly "$PM_{2.5}$ ATM" concentrations reported by PurpleAir monitors in Pittsburgh, PA, USA resulted in a negative mean bias (relative to a beta attenuation monitor) and increased the absolute value of the mean bias compared to uncorrected measurements (from 1.9 to 2.7 µg $m^{-3}$). We found that the RH-based correction decreased the absolute value of the mean bias (from 2.1 to 0.56 µg $m^{-3}$). The difference between our results and those of Malings et al. (2019) could be due to (a) differences in the composition and size of the $PM_{2.5}$ measured in Fort Collins versus Pittsburgh and (b) differences in the distribution of hourly ambient RH values in Fort Collins versus Pittsburgh. In other words, the underestimation that

results from applying an RH-based correction similar to Equation 1 may be less severe in a less humid environment like Fort Collins.

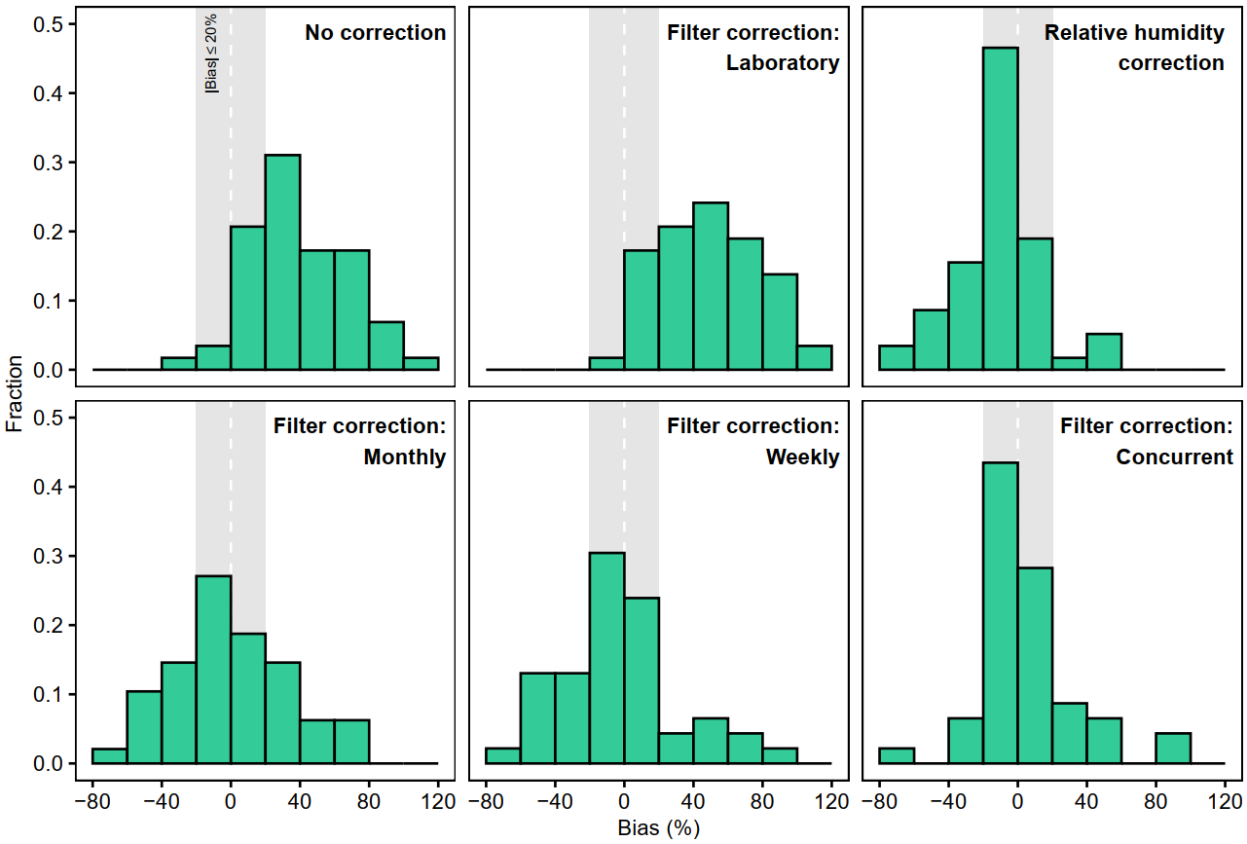



**Figure 7.** Bias of sample-averaged $PM_{2.5}$ concentrations reported by the PurpleAir monitors, relative to the concentrations measured using co-located 16.7 L $min^{-1}$ filter samples, after applying different corrections. The y-axis represents the fraction of samples that fell into each bin. The gray shaded area represents an absolute bias less than or equal to 20%.

Correction factors derived from monthly and weekly in-field co-location with UPAS filter samples reduced the median bias compared to uncorrected, lab-corrected, and RH-corrected PurpleAir "$PM_{2.5}$ ATM" concentrations (see **Table 1**). Applying monthly and weekly filter-based correction factors increased the fraction of samples with an absolute bias of ≤ 2 µg $m^{-3}$ or ≤ 20%, compared to uncorrected PurpleAir monitors, but not as much as the RH-based correction. For example, 77% and 46% of the samples, respectively had absolute biases of ≤ 2 µg $m^{-3}$ and ≤ 20% after applying monthly correction factors (72% and 38%, respectively, if data collected between 30

November and 06 December were excluded; see **Table S1** and **Figure S17**). This result might have been observed because monthly and weekly correction factors accounted for the general tendency of the PurpleAir "$PM_{2.5}$ ATM" values to overestimate the reference $PM_{2.5}$ concentration, but did not account for the impact of shorter-term variations in relative humidity on particle properties. Monthly co-locations with filter samplers could be helpful for detecting long-term (e.g., seasonal) variations in the mean bias of PurpleAir monitors. For example, Sayahi et al. (2019) reported that the mean bias (relative to a TEOM) of hourly "$PM_{2.5}$ CF=1" values reported by PurpleAir monitors in Salt Lake City, UT, USA was negative during the spring (when ambient $PM_{2.5}$ concentrations were low) but was positive during wildfire season (when ambient $PM_{2.5}$ concentrations were higher and PM composition was dominated by organic carbon and crustal compounds). However, our field deployment spanned a single season and, as a result, we were not able to evaluate seasonal variations in PurpleAir bias.

Concurrent correction factors derived from in-field co-location with UPAS filter samples resulted in the lowest median concentration bias (-0.01 $\mu g\, m^{-3}$) and the largest fraction of samples with absolute bias ≤ 20% (72%). The disadvantage of concurrent correction factors was that they had to be evaluated and applied to the PurpleAir data during post-processing, whereas monthly and weekly filter-based correction factors could be determined in advance.

The concurrent gravimetric correction factors (n=175) calculated for each sample period are shown in **Figure S18**. The intraclass correlation coefficients calculated using the model shown in Equation 8 were 0.44 (95% CI = 0.18–0.63) for sample date and 0.03 (95% CI = 0.00–0.11) for PurpleAir monitor. These ICCs indicated that approximately 44% and 3% of the variability in the log-transformed concurrent gravimetric correction factor was explained by differences between sample dates and differences between PurpleAir monitors, respectively. Note that differences between PurpleAir monitors might represent differences between the 19 different monitors

themselves and/or differences between the 15 different monitoring sites. Given these results, variability in correction factors might be related to changes in aerosol properties and concentrations resulting from changes in weather as well as weekday vs. weekend differences in traffic patterns, commercial activities, and industrial activities (samples that started on Mondays ran on weekdays, whereas samples that started on Fridays ran over the weekend).

Another approach for correcting $PM_{2.5}$ concentrations reported by PurpleAir monitors would be to use a set of PurpleAir data and co-located filter samples to build a model that predicts the correction factor (i.e., the factor by which $c_{PA}$ is multiplied in Equation 2) using weather data and/or other predictors (Tryner et al., 2019a). Both Malings et al. (2019) and Magi et al. (2019) reported that models that included weather variables (e.g., relative humidity and temperature) as predictors improved agreement between hourly concentrations reported by PurpleAirs and beta attenuation monitors. We did not take such an approach in this study because the dataset from our field deployment was relatively small. Weather-based predictors, such as relative humidity, could be confounded with other variables that vary with sample date and use of weather data from a single central site was likely to result in overfitting of the model.

Correcting the $PM_{2.5}$ concentrations reported by the PurpleAir monitors using factors derived from the UPAS filter samples assumes that the $PM_{2.5}$ concentrations measured using the UPAS filter samples are accurate. To evaluate this assumption, $PM_{2.5}$ concentrations derived from the UPAS filter samples were compared to $PM_{2.5}$ concentrations derived from the 16.7 L $min^{-1}$ filter samples. These two quantities agreed well considering the low concentrations measured during the field campaign (**Figure 8**). The median bias of the UPAS filter sample, relative to the 16.7 L $min^{-1}$ filter sample, was 0%, bias was ≤ 2 μg $m^{-3}$ for 42/48 samples (88%), and percent bias was ≤ 20% for 33/48 samples (69%). Note that the percent bias of the UPAS filter samples, relative to the 16.7 L $min^{-1}$ filter samples, is illustrated in the “Concurrent” panel of **Figure 7**. The linear model fit using

Deming regression had an intercept of 0.50 (95% CI = -0.37 to 1.36) and a slope of 0.90 (95% CI = 0.67 to 1.12).

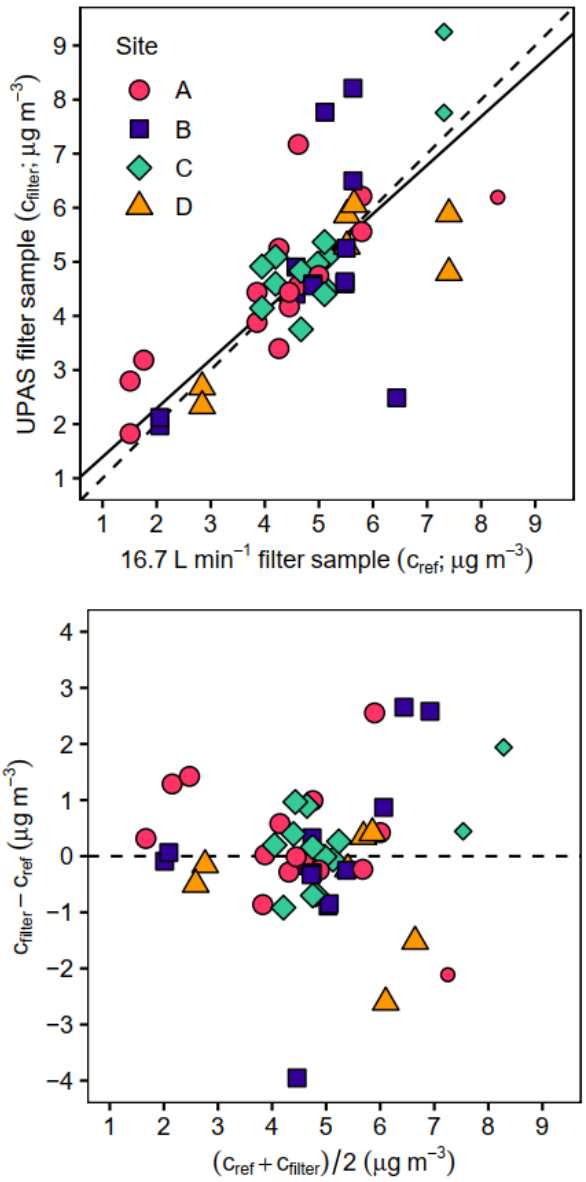



**Figure 8.** Top: A comparison of the $PM_{2.5}$ concentrations derived from the 16.7 L min$^{-1}$ filter samples and the UPAS filter samples. The dashed diagonal line is y = x and the solid diagonal line is the result of the Deming regression. Bottom: A Bland-Altman plot comparing the $PM_{2.5}$ concentrations measured using the 16.7 L min$^{-1}$ filter samples and the UPAS filter samples. In both graphs, different marker colors represent different sites (A, B, C, and D). Smaller markers represent the 48-hour samples collected between 19-21 November.

## 4. CONCLUSIONS

The "$PM_{2.5}$ ATM" concentrations reported by PurpleAir monitors appeared to increase linearly with the concentration of NIST Urban PM measured using a TEOM, and a linear model relating the two concentrations had $R^2$ = 99%; however, an F-test identified a significant lack of fit between the linear model and the data ($\alpha = 0.05$). The laboratory-derived gravimetric correction factor for the PurpleAir monitors did not translate to the field. Whereas the relationship between the NIST Urban PM concentrations measured by the TEOM and the PurpleAir "$PM_{2.5}$ ATM" values was approximately 1:1 during the pre-deployment laboratory evaluation, PurpleAir monitors overestimated $PM_{2.5}$ concentrations measured using conventional 16.7 L $min^{-1}$ filter samplers for 49/52 72-hour outdoor samples and 6/6 48-hour outdoor samples.

A $PM_{2.5}$ correction factor calculated from hourly ambient RH values increased the fraction of sample-averaged PurpleAir "$PM_{2.5}$ ATM" concentrations with bias ≤ 20% (relative to the 16.7 L $min^{-1}$ filter samplers) from 24% (with no correction) to 66%. Gravimetric correction factors derived from monthly, weekly, and concurrent co-location with portable filter samplers increased the fractions of sample-averaged PurpleAir "$PM_{2.5}$ ATM" concentrations with bias ≤ 20% from 24% (with no correction) to 46%, 54%, and 72%, respectively. The increase in the fraction of samples with bias ≤ 20% that was achieved with more frequent correction intervals indicated that the correction factor relating the 72- or 48-hour $PM_{2.5}$ concentration measured using the PurpleAir to the $PM_{2.5}$ concentration derived from a co-located gravimetric sample varied with time—presumably due to variations in particle composition and size distribution resulting from variations in particle sources and/or ambient relative humidity. Given the low 72- and 48-hour average ambient $PM_{2.5}$ concentrations derived from the 16.7 L $min^{-1}$ filter samples collected during our field deployment (median = 5.1 μg $m^{-3}$, range = 1.5 to 8.3 μg $m^{-3}$), future work investigating whether the results presented here are generalizable to more polluted regions would be beneficial.

Finally, co-locations with conventional 16.7 L $min^{-1}$ filter samplers demonstrated that the UPAS technology in the ASPEN boxes accurately measured 72- and 48-hour mean $PM_{2.5}$ concentrations below 10 µg $m^{-3}$. The median bias of the UPAS filter samples (relative to the 16.7 L $min^{-1}$ filter samples) was 0%, and bias was ≤ 20% for 69% of samples.

FUNDING

This work was funded by the National Oceanic and Atmospheric Administration under grant no. 1305M218CNRMW0048. The funding source had no role in the study design; in collection, analysis, and interpretation of data; in the writing of this article; or in the decision to submit this article for publication.

DECLARATION OF COMPETING INTEREST

John Volckens is a scientific founder of Access Sensor Technologies, LLC and has an equity interest in the company. The terms of this arrangement have been reviewed and approved by Colorado State University in accordance with its conflict of interest policies.

ACKNOWLEDGEMENTS

The authors thank Mollie Phillips, Gabe Neymark, and Casey Quinn for help with ASPEN box development and field data collection.

REFERENCES

Analytical Methods Committee, 1994. Is my calibration linear? The Analyst 119, 2363–2366. https://doi.org/10.1039/an9941902363

Austin, E., Novosselov, I., Seto, E., Yost, M.G., 2015. Laboratory Evaluation of the Shinyei PPD42NS Low-Cost Particulate Matter Sensor. PLOS ONE 10, e0137789. https://doi.org/10.1371/journal.pone.0137789

Bai, L., Huang, L., Wang, Z., Ying, Q., Zheng, J., Shi, X., Hu, J., 2019. Long-term field Evaluation of Low-cost Particulate Matter Sensors in Nanjing. Aerosol Air Qual. Res. https://doi.org/10.4209/aaqr.2018.11.0424

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R.H.B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J., 2019. Package “lme4”: Linear Mixed-Effects Models using “Eigen” and S4 [WWW Document]. The Comprehensive R Archive Network. URL https://CRAN.R-project.org/package=lme4 (accessed 6.10.19).

Chakrabarti, B., Fine, P.M., Delfino, R., Sioutas, C., 2004. Performance evaluation of the active-flow personal DataRAM PM2.5 mass monitor (Thermo Anderson pDR-1200) designed for continuous personal exposure measurements. Atmospheric Environment 38, 3329–3340. https://doi.org/10.1016/j.atmosenv.2004.03.007

Crilley, L.R., Shaw, M., Pound, R., Kramer, L.J., Price, R., Young, S., Lewis, A.C., Pope, F.D., 2018. Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring. Atmos. Meas. Tech. 11, 709–720. https://doi.org/10.5194/amt-11-709-2018

English, P.B., Olmedo, L., Bejarano, E., Lugo, H., Murillo, E., Seto, E., Wong, M., King, G., Wilkie, A., Meltzer, D., Carvlin, G., Jerrett, M., Northcross, A., 2017. The Imperial County Community Air Monitoring Network: A Model for Community-based Environmental Monitoring for Public Health Action. Environmental Health Perspectives 125. https://doi.org/10.1289/EHP1772

Feinberg, S., Williams, R., Hagler, G.S.W., Rickard, J., Brown, R., Garver, D., Harshfield, G., Stauffer, P., Mattson, E., Judge, R., Garvey, S., 2018. Long-term evaluation of air sensor technology under ambient conditions in Denver, Colorado. Atmospheric Measurement Techniques 11, 4605–4615. https://doi.org/10.5194/amt-11-4605-2018

Gao, M., Cao, J., Seto, E., 2015. A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of PM2.5 in Xi’an, China. Environmental Pollution 199, 56–65. https://doi.org/10.1016/j.envpol.2015.01.013

Gillooly, S.E., Zhou, Y., Vallarino, J., Chu, M.T., Michanowicz, D.R., Levy, J.I., Adamkiewicz, G., 2019. Development of an in-home, real-time air pollutant sensor platform and implications for community use. Environmental Pollution 244, 440–450. https://doi.org/10.1016/j.envpol.2018.10.064

Holstius, D.M., Pillarisetti, A., Smith, K.R., Seto, E., 2014. Field calibrations of a low-cost aerosol sensor at a regulatory monitoring site in California. Atmospheric Measurement Techniques 7, 1121–1131. https://doi.org/10.5194/amt-7-1121-2014

Jayaratne, R., Liu, X., Thai, P., Dunbabin, M., Morawska, L., 2018. 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. https://doi.org/10.5194/amt-11-4883-2018

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., Buckley, K., 2016. Community Air Sensor Network (CAIRSENSE) project: evaluation of low-costsensor performance in a suburban environment in the southeastern UnitedStates. Atmos. Meas. Tech. 9, 5281–5292. https://doi.org/10.5194/amt-9-5281-2016

Kelly, K.E., Whitaker, J., Petty, A., Widmer, C., Dybwad, A., Sleeth, D., Martin, R., Butterfield, A., 2017. Ambient and laboratory evaluation of a low-cost particulate matter sensor. Environmental Pollution 221, 491–500. https://doi.org/10.1016/j.envpol.2016.12.039

Levy Zamora, M., Xiong, F., Gentner, D., Kerkez, B., Kohrman-Glaser, J., Koehler, K., 2019. Field and Laboratory Evaluations of the Low-Cost Plantower Particulate Matter Sensor. Environmental Science & Technology 53, 838–849. https://doi.org/10.1021/acs.est.8b05174

Lüdecke, D., 2019. Package “sjstats”: Collection of Convenient Functions for Common Statistical Computations [WWW Document]. The Comprehensive R Archive Network. URL https://CRAN.R-project.org/package=sjstats (accessed 6.1.19).

Magi, B.I., Cupini, C., Francis, J., Green, M., Hauser, C., 2019. Evaluation of PM2.5 measured in an urban setting using a low-cost optical particle counter and a Federal Equivalent Method Beta Attenuation Monitor. Aerosol Science and Technology 1–13. https://doi.org/10.1080/02786826.2019.1619915

Malings, C., Tanzer, R., Hauryliuk, A., Saha, P.K., Robinson, A.L., Presto, A.A., Subramanian, R., 2019. Fine Particle Mass Monitoring with Low-Cost Sensors: Corrections and Long-Term Performance Evaluation. Aerosol Science and Technology 1–40. https://doi.org/10.1080/02786826.2019.1623863

McGraw, K.O., Wong, S.P., 1996. Forming inferences about some intraclass correlation coefficients. Psychological Methods 1, 30–46. https://doi.org/10.1037/1082-989X.1.1.30

Nees, M., 1993. Quality Assurance Handbook for Air Pollution Measurement Systems: Volume I — A Field Guide to Environmental Quality Assurance (No. EPA-600/R-94/038a). U.S. Environmental Protection Agency, Research Triangle Park, NC.

Nelson, K., 2016. Dirty Air Monitors: Are they telling us the whole truth?

Reimann, C., Filzmoser, P., Garrett, R.G., Dutter, R., 2008. Statistical Data Analysis Explained: Applied Environmental Statistics with R. John Wiley & Sons, Chichester, England.

Rickenbacker, H., Brown, F., Bilec, M., 2019. Creating environmental consciousness in underserved communities: Implementation and outcomes of community-based environmental justice and air pollution research. Sustainable Cities and Society 47, 101473. https://doi.org/10.1016/j.scs.2019.101473

Sardar, S.B., Fine, P.M., Sioutas, C., 2005. Seasonal and spatial variability of the size-resolved chemical composition of particulate matter ($PM_{10}$) in the Los Angeles Basin. Journal of Geophysical Research 110. https://doi.org/10.1029/2004JD004627

Sayahi, T., Butterfield, A., Kelly, K.E., 2019. Long-term field evaluation of the Plantower PMS low-cost particulate matter sensors. Environmental Pollution 245, 932–940. https://doi.org/10.1016/j.envpol.2018.11.065

Singer, B.C., Delp, W.W., 2018. Response of consumer and research grade indoor air quality monitors to residential sources of fine particles. Indoor Air 28, 624–639. https://doi.org/10.1111/ina.12463

Sioutas, C., Kim, S., Chang, M., Terrell, L.L., Gong, H., 2000. Field evaluation of a modified DataRAM MIE scattering monitor for real-time PM2.5 mass concentration measurements. Atmospheric Environment 34, 4829–4838. https://doi.org/10.1016/S1352-2310(00)00244-2

Sousan, S., Koehler, K., Hallett, L., Peters, T.M., 2017. Evaluation of consumer monitors to measure particulate matter. Journal of Aerosol Science 107, 123–133. https://doi.org/10.1016/j.jaerosci.2017.02.013

Spielvogel, J., Hartstock, S., Grimm, H., 2009. New methods and standards for fine dust. J. Phys.: Conf. Ser. 170, 012024. https://doi.org/10.1088/1742-6596/170/1/012024

Therneau, T., 2018. Package “deming”: Deming, Thiel-Sen and Passing-Bablock Regression [WWW Document]. The Comprehensive R Archive Network. URL https://CRAN.R-project.org/package=deming (accessed 7.11.18).

Tryner, J., Good, N., Wilson, A., Clark, M.L., Peel, J.L., Volckens, J., 2019a. Variation in gravimetric correction factors for nephelometer-derived estimates of personal exposure to PM2.5. Environmental Pollution 250, 251–261. https://doi.org/10.1016/j.envpol.2019.03.121

Tryner, J., L’Orange, C., Mehaffy, J., Miller-Lionberg, D., Hofstetter, J.C., Wilson, A., Volckens, J., 2019b. Dataset associated with “Laboratory evaluation of low-cost PurpleAir PM monitors and in-field correction using co-located portable filter samplers” [WWW Document]. https://doi.org/10.25675/10217/195445

Tryner, J., Quinn, C., Windom, B.C., Volckens, J., 2019c. Design and evaluation of a portable $PM_{2.5}$ monitor featuring a low-cost sensor in line with an active filter sampler. Environ. Sci.: Processes Impacts 21, 1403–1415. https://doi.org/10.1039/C9EM00234K

Volckens, J., Quinn, C., Leith, D., Mehaffy, J., Henry, C.S., Miller-Lionberg, D., 2017. Development and evaluation of an ultrasonic personal aerosol sampler. Indoor Air 27, 409–416. https://doi.org/10.1111/ina.12318

Wang, Y., Li, J., Jing, H., Zhang, Q., Jiang, J., Biswas, P., 2015. Laboratory Evaluation and Calibration of Three Low-Cost Particle Sensors for Particulate Matter Measurement. Aerosol Science and Technology 49, 1063–1077. https://doi.org/10.1080/02786826.2015.1100710

Zheng, T., Bergin, M.H., Johnson, K.K., Tripathi, S.N., Shirodkar, S., Landis, M.S., Sutaria, R., Carlson, D.E., 2018. Field evaluation of low-cost particulate matter sensors in high and low concentration environments. Atmospheric Measurement Techniques Discussions 1–40. https://doi.org/10.5194/amt-2018-111

Zikova, N., Masiol, M., Chalupa, D., Rich, D., Ferro, A., Hopke, P., 2017. Estimating Hourly Concentrations of PM2.5 across a Metropolitan Area Using Low-Cost Particle Monitors. Sensors 17, 1922. https://doi.org/10.3390/s17081922