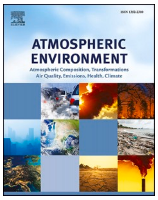
Atmospheric Environment 256 (2021) 118432

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv



Calibration of low-cost PurpleAir outdoor monitors using an improved method of calculating $PM_{2.5}$

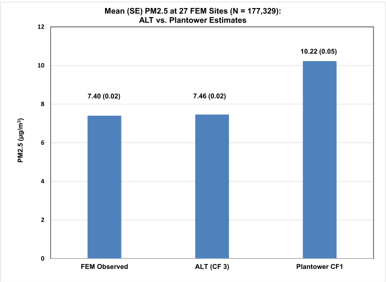
Lance Wallace [a,*], Jianzhao Bi [b], Wayne R. Ott [c], Jeremy Sarnat [d], Yang Liu [d]

[a] *Independent Researcher, 428 Woodley Way, Santa Rosa, CA, 95409, USA*
[b] *University of Washington, USA*
[c] *Stanford University, USA*
[d] *Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, 30322, USA*

HIGHLIGHTS

- 33 PurpleAir monitors were calibrated against 27 EPA regulatory PM2.5 monitors.
- A reproducible method to estimate PM2.5 from PurpleAir particle counts is presented.
- The method is superior to the undisclosed algorithm of the Plantower sensor.
- The method has improved accuracy and precision, with a lower limit of detection.
- A different calibration factor for PM2.5 aerosols from wildfires was indicated.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Low-cost monitors
PurpleAir
$PM_{2.5}$
Calibration factor (CF)

ABSTRACT

$PM_{2.5}$ hourly average measurements from 33 outdoor PurpleAir particle monitors were compared with hourly measurements from 27 nearby US EPA Air Quality System (AQS) stations employing Federal Equivalent Method (FEM) monitors in California over an 18-month (77-week) period. A transparent and reproducible alternative method (ALT) of calculating $PM_{2.5}$ from the particle numbers in three size categories was used in place of the estimates provided by Plantower, the manufacturer of the sensors used in PurpleAir monitors. The ALT method was superior in several ways (better precision, lower limit of detection, improved size distribution) compared to Plantower's CF1 or ATM data series. PurpleAir monitors were strongly correlated with the nearby US EPA Air Quality System AQS stations. A calibration factor (CF) ranging between 2.9 and 3.1 was empirically derived for the PurpleAir estimates using the ALT method. This value was based on comparing the average value of 177,329 PurpleAir measurements to the value calculated from the FEM stations. The monitoring period included about 13 weeks showing very high outdoor values due to several major fires covering several hundred thousand acres. The CF during these 13 weeks averaged 2.39, whereas the CF for the remaining 64 weeks averaged about 3.21, suggesting a different response to the smoke from wildfires compared with normal ambient fine particulate matter ($PM_{2.5}$). The standard Plantower CF1 data series overestimated the FEM values by about 40%, in agreement with several other studies.

---

* Corresponding author.
E-mail address: lwallace73@gmail.com (L. Wallace).

https://doi.org/10.1016/j.atmosenv.2021.118432
Received 10 March 2021; Received in revised form 13 April 2021; Accepted 19 April 2021

1. Introduction

Low-cost air quality monitors reporting particulate matter (PM) concentrations directly to the Internet are now being widely used in the US and in Asia (Williams et al., 2014; US EPA, 2017). Multiple manufacturers are producing tiny electronic sensors for use in these monitors. The sensors often employ lasers to scatter light off the particles into detectors. The scattered light is analyzed, generally using Mie scattering theory, to estimate particle size, number, and mass.

Among the most widely used monitors, with >16,000 devices in operation, is PurpleAir (https://www2.purpleair.com/). PurpleAir monitors use one or two identical sensors (Plantower Model PMS 5003) manufactured by the company Plantower, (http://www.plantower.com/en/). This sensor is constructed using a laser providing light at an approximate 650 nm wavelength that is scattered throughout a 90° viewing area (He et al., 2020). The sensor measures particle numbers per deciliter for 6 overlapping size categories: >0.3 μm, >0.5 μm, >1 μm, >2.5 μm, >5 μm and >10 μm (Plantower User Manual, 2016). From these particle number measurements, Plantower uses a proprietary algorithm to estimate various PM size fractions, including PM_1, $PM_{2.5}$, and PM_{10} concentrations. Two data series are produced (CF1 and ATM) for these three particle mass concentration estimates. Plantower provides no information on the calibration aerosol used to correct its sensor retrieval signal.

Our aim in this study is to present the first calibration of PurpleAir monitors using all monitors in the State of California within 500 m of official US EPA Air Quality System (AQS) Federal Equivalent Method (FEM) stations. We also provide here an existing widely used transparent and reproducible independent alternative (ALT) approach to calculating $PM_{2.5}$ from the particle number estimates provided by the PurpleAir monitors. This approach is widely used by manufacturers of optical particle monitors capable of collecting simultaneous data on multiple particle size categories (e.g., the TSI Model 3330 (https://tsi.com/products/particle-sizers/particle-size-spectrometers/optical-particle-sizer-3330/)). Since this approach is central to our study, we devote an Appendix to explaining the approach in detail and comparing it with the $PM_{2.5}$ data provided by PurpleAir. In particular, we compare results on precision, bias, limit of detection, coefficient of variation, and size distribution using multiple months-long datasets collected in two locations.

1.1. *Previous studies employing PurpleAir monitors*

PurpleAir monitors have been evaluated in several studies. The South Coast Air Quality Management District, part of the California Air Resources Board (CARB), operates a program called AQ-SPEC that has issued reports on about a dozen low-cost monitors (http://www.aqmd.gov/aq-spec). Their field evaluation of PurpleAir monitors took place from February to April 2016 (AQ-SPEC, 2016). This test found generally good agreement with two reference monitors for $PM_{2.5}$ (R^2 = 78% and 90%), but poor agreement for PM_{10} (R^2 = 34% and 45%). The PurpleAir instrument overestimated $PM_{2.5}$ concentrations by about 40% compared to the two reference instruments. The report does not specify whether the CF1 or CF ATM series was used. Similarly, the Lawrence Berkeley National Laboratory (LBNL) tested 7 monitors for ability to measure particles from a large number of common indoor sources such as cooking and cleaning (Singer and Delp, 2018). Seven low-cost PM monitors were tested using a variety of common indoor sources, including candles, incense, and cooking of various types. Four monitors, including PurpleAir, were described as "quantitative" (within a factor of 2 compared to two reference instruments) for most of the aerosol types measured. Indeed, the investigators commented that the low-cost monitors were about as accurate as the research reference instruments. A presentation based on the same study concluded that the PurpleAir instruments were "quantitatively much better than the others" Walker (2018). The CF1 data series was used (Delp and Singer, 2020).

State and local air quality agencies have tested the PurpleAir monitors for possible use in their air quality monitoring programs, including Lane County, OR and Salt Lake County, Utah (Kelly et al., 2017). Sayahi et al. (2019) reported a limit of detection (LOD) of about 6 μg/m^3.

A long-term comparison of four sensors, including the two Plantower models PMS 5003 and PMS 7003, took place at two schools near Southhampton, England (Bulot et al., 2019) between March 13, 2018, and February 28, 2019. The two Plantower models showed excellent agreement with each other (Spearman coefficient of 0.98). They were also well correlated with the background reference station data, with a bias ranging between 1.15 and 1.22 at one school, and similar values at the second school.

Zheng et al. (2018) compared three Plantower Model PMS 3003 sensors to a reference monitor for 30 days at a US EPA facility in Research Triangle Park, NC. Correlation across the three sensors was very high (R^2 = 97%). Correlation with the reference monitor was moderate (R^2 = 66%) for both 1-min and 1-h averaging periods, but a semiempirical correction for RH improved the correlation to 93% for both averaging times. The RH values at the site averaged 64% ± 22%. The authors concluded that the 3003 sensors were highly precise and capable of measuring $PM_{2.5}$ to within 10%.

Levy Zamora et al., (2019) evaluated three Plantower PMS A003 sensors when exposed to eight particulate matter (PM) sources (i.e., incense, oleic acid, NaCl, talcum powder, cooking emissions, and monodispersed polystyrene latex spheres under controlled laboratory conditions and also residential air and ambient outdoor air in Baltimore, MD). Overall accuracy for $PM_{2.5}$ ranged between 86% and 93% for incense, cooking, residential air and ambient outdoor air, and precision ranged between 9% and 12% for these four sources. The authors remark that performance was good for both PM_1 and $PM_{2.5}$ but poor for larger particles.

Three important studies appeared in early 2020. He et al. (2020) tested the response of three Plantower PMS5003 sensors to various aerosols of known composition and size. They developed the first estimate of the transfer function governing the response of the Plantower optical sensor to particles of different sizes. The authors found that particles well outside the boundaries of the size channels contributed to the sensor response. For example, 0.2 μm monodisperse particles were readily detectable in the smallest channel (>0.3 μm). More concerning, 1.0 μm monodisperse particles provided the strongest signal in the smallest size channel rather than the channels more directly related to that size. This important finding adds considerable uncertainty to the size classifications employed by the Plantower sensors, and to the mass calculations based on those sizes, including our own $PM_{2.5}$ estimates.

Simultaneously, a major study used Plantower PMS A003 sensors in seven cities to predict the $PM_{2.5}$ measurements made by Federal Reference Methods (Zusman et al., 2020). This study did not use the ATM or CF1 PM values, but rather the particle number counts. The authors chose to use the original overlapping definitions (i.e., >0.3 μm, >0.5 μm …) after finding that they (slightly) outperformed the non-overlapping size categories. The model used the first five of the six categories, choosing not to use the >10 μm category (even though all particles in that category also appear in the first five categories). Also entered were temperature and humidity measurements by the low-cost monitors. Very good results were obtained across the seven cities with 10-fold cross-validation ($PM_{2.5}$ precision r = 0.99, accuracy R^2 = 0.96, RMSE 1.15 μg/m^3).

The third study (Bi et al., 2020) employed all outdoor PurpleAir monitors (2,090) operating in California in the year 2018 to provide additional $PM_{2.5}$ measurements in the geographic areas covered by 138 official air monitoring stations. The density of sensors allowed creation of a grid 1 km to a side (almost 500,000 cells). 5.84 million hourly measurements were obtained. Geographic calibration of the instruments with the reference monitors resulted in reducing the overall bias from 1.9 to ~0 μg/m^3, and reducing the residual error by 36%. After correction for the 17% bias and down-weighting the PurpleAir values by about a factor of 5 to account for lower-quality measurements, the authors found a useful improvement in estimating $PM_{2.5}$ concentrations throughout the state.

Wallace et al. (2020) co-located two PurpleAir monitors with two research-grade monitor types (Piezobalance and SidePak) in 124 experiments measuring $PM_{2.5}$ from vaping marijuana liquid. The PurpleAir instruments had a precision and coefficient of variation (COV) that was similar to the research-grade instruments, and both had been calibrated by comparison with gravimetric instruments. That study used the alternative (ALT) method presented in the Appendix below to estimate $PM_{2.5}$. The calibration factor for PurpleAir was determined to be 3.0.

Bi et al. (2021) examined 91 pairs of indoor/outdoor PurpleAir monitors within 500 m of each other over a 20-month period ending June 2020. Using LOWESS local regression, they determined infiltration factors for all indoor sites, with a mean value of 0.26 (IQR 0.15–0.34). The resulting exposure error (difference between total indoor exposure and exposure due to particles of ambient origin) was plotted as a function of outdoor concentration. The peak of the exposure error occurred at low outdoor concentrations <5 μg/m^3 and dropped to nearly zero at outdoor concentrations near 25–30 μg/m^3.

These and other studies dealing with the Plantower sensor are briefly described in Table 1.

Table 1
Studies evaluating Plantower sensors.

Reference	Year	Data series	Model PMS-	N	Location	Time span	Comment
Chen	2017	CF1 or ATM	3003 5003	N/A	Taiwan	N/A	General outline of system with data archiving. Two short-term case studies.
Wang, K.	2018	CF1 or ATM	7003	17	Shanghai, China	7 days	Compared with TEOM: Outdoor R^2 0.72–0.78, mean RSD 21%; indoor R^2 0.95–0.96, mean RSD 16%.
Zheng	2018	CF1	3003	7	Durham NC; Kanpur, India	NC: 90 days; India 45 days	1-h mean errors of 200% in Durham, but 35%–46% in India, indicating improved performance at high concentrations. Following empirical RH correction, estimates are within ~10%.
Becnel	2019	CF1 or ATM	3003	50	Salt Lake County, Utah	6 months	Individual calibration improved RMSE by 1.8 times. Good agreement (88% R^2) with reference monitors
Bulot	2019	CF1 or ATM	5003 7003	6	Southhampton, UK	1 year	Moderate to good correlation: $0.61 < r < 0.88$.
Francis	2019	ATM, Particle #	5003	N/A	Sabah, Malaysia	1 week	Use of particle number to estimate mass. PM_1 & $PM_{2.5}$ $R^2 = 0.82$ & 0.88. PM_{10} unreliable.
He	2019	Particle #	5003	6	Clarkson Univ. NY		Theoretical analysis of transfer function. All size channels include response to sizes outside their boundaries.
Kaduwela	2019	Particle #	7003	1	Albany, CA	2 weeks	School study including wildfire days (15X increase in particle number indoors).
Levy Zamora	2019	CF1 or ATM	A003	3	Baltimore, MD	1 month, 10 days	Accuracy 87%–96%, precision 9%–10% for incense, cooking, and residential indoor. Underestimates with precision 10–24% for NaCl, talcum powder, oleic acid. Overestimate of ambient data by 1.6–2.4X, reduced to 0.9–1.4X after correction for RH.
Magi	2019	ATM	5003	1	Charlotte, NC	16 months	Multiple linear regression with T, RH, and BAM improved accuracy by 27%–57%. 15% of data < LOD of 5 μg/m^3.
Malings	2019	CF1 or ATM	5003	20	Pittsburgh, PA	17 months	Multiple segmented linear regression with T and RH (10 parameters). Median correlation with BAM of 0.73; mean absolute error 2.5 μg/m^3; bias −0.14 μg/m^3
Masic	2019	CF1 or ATM	5003	2	Sarajevo, Bosnia	16 months	R^2 93% and 96%. Bias 37% and 31%
Tryner	2019	CF1 and ATM	5003	8	Fort Collins, CO	1 week	Gravimetric correction reduced bias and RSD. CF1 values 50% higher than ATM values at high $PM_{2.5}$ concentrations
Bi	2020	CF1 or ATM	5003	2090	California	1 year	Overestimate of 13/11.1 = 17%, based on 137,000 1-h measurements. Calibration reduced bias to ~0. Most important parameter in estimating $PM_{2.5}$ was the $PM_{2.5}/PM_{10}$ ratio.
Wang, Z.	2020	CF1 or ATM	1003 3003 5003	9	Berkeley, CA	1 month	3 different monitors incorporating the three models listed. All performed reasonably well (within a factor of two) with multiple different aerosols containing particles >0.25 μm.
Zusman	2020	Particle #	A003	80	Seattle, 6 other cities	30–95 weeks	$PM_{2.5}$ precision $r = 0.99$, accuracy $R^2 = 0.96$, RMSE 1.15 μg/m^3. Region-specific models in 6 other cities R^2 0.74–0.95, RMSE 0.84–2.46 μg/m^3. Seattle model applied to other 6 cities: R^2 0.67–0.84, RMSE 1.67–3.41 μg/m^3.
Wallace	2020	Particle #	5003	4	Santa Rosa, CA Redwood city, CA	14 months (124 experiments)	Indoor measurements of aerosol from vaping marijuana liquids. PurpleAir monitors were collocated with research-grade monitors. Precision and COV were comparable across all monitor types. Calibration factor estimate was about 3.0.
Bi	2021	Particle # (ALT)	5003	91 indoor-outdoor pairs	California	20 months	Infiltration factors and indoor-generated exposures were estimated for pairs of PurpleAir indoor-outdoor monitors within 500 m distance. Exposure errors were calculated as a function of outdoor concentrations

2. Methods and materials

A database consisting of $PM_{2.5}$ measurements from PurpleAir monitors near EPA Air Quality System (AQS) monitors employing Federal Equivalent Methods (FEM) to provide hourly estimates was prepared. Across the entire state of California, 33 PurpleAir monitors were within 500 m of 27 EPA monitors. 500 m was used as a cutoff in an earlier study of PurpleAir monitors (Bi et al., 2020) based on low and slowly-changing correlations between PurpleAir and AQS stations as a function of distance. Data from November 2018 to April 2020 (18 months) was downloaded from the PurpleAir website (PurpleAir.com) and the EPA website (https://aqs.epa.gov/aqsweb/airdata/download_files.html). The 2-min average PurpleAir measurements were averaged over hourly periods to match the EPA measurements. The hourly PurpleAir averages were required to have valid reading for at least 20 of the 30 possible

2-min measurements to be included in the analysis. Only PurpleAir sites with the PA-II monitors employing two independent Plantower PMS 5003 sensors were included in the analysis. The $PM_{2.5}$ estimates from the two sensors within each monitor were required to agree within 30% of each other, corresponding to a precision [abs $(A\text{-}B)/(A + B)$] of 0.130. About 90% of all PurpleAir 1-h averages met this condition. Only monitors with at least 3 weeks of data were included. A small number of measurements with abnormal temperature and relative humidity readings were also removed. The final database had 177,329 matched PurpleAir and EPA hourly $PM_{2.5}$ estimates.

A crucial aspect of our approach is our use of a transparent and reproducible alternative method (ALT) to estimate the $PM_{2.5}$ concentrations. The ALT method is based on the number of particles per deciliter reported by the PMS 5003 sensors in the PurpleAir instrument for the three size categories less than 2.5 μm in diameter. The method makes no use of either the CF1 or ATM data series calculated according to a proprietary and undisclosed algorithm by the Plantower manufacturers of the sensors. The method is fully explained in the Appendix to this article. The practical improvements allowed by the method in precision, limit of detection, and distribution shape are also contained in the Appendix, using measurements by two co-located PurpleAir indoor monitors in each of two homes, one in Santa Rosa, CA, and the other in Redwood City, CA.

3. Results and discussion

The means (standard errors) for the AQS measurements and our $PM_{2.5}$ ALT estimates were 7.40 (SE 0.03) and 2.49 (0.01) $\mu g/m^3$, respectively, leading to an overall average calibration factor (CF) of $7.40/2.49 = 2.98$ (0.01). An overall average calibration factor can also be calculated for the Plantower CF1 data series. The same requirement of the sensors to agree within 30% of their CF1 estimates was applied, reducing the number of accepted measurements *for this comparison only* to 154,900. The mean and standard error for the AQS measurements and the CF1 estimates were 7.98 (0.02) and 11.14 (0.03), respectively, leading to a calibration factor for the CF1 series of 0.716 (0.003). This finding of about a 40% (1/0.716) overestimate for the CF1 data series is in very good agreement with the finding of the AQ-SPEC program and the paper by Gupta et al. (2018), each of which found about a 35%–40% overestimate of $PM_{2.5}$ for the PurpleAir monitors.

Individual calibration factors for each PurpleAir monitor can also be calculated as the ratio of the mean AQS to the mean ALT $PM_{2.5}$ measurement (Table S1). Four of the 33 monitors had fewer than 2 months of data and were not analyzed further. The 29 remaining monitors had at least 1800 h (75 days) to 10,770 h (>400 days) of hourly average measurements. The median CF was 3.13 (Interquartile range (ICR) = 2.5–3.9). The range across the 29 individual sites was 2.2–6.0.

A third approach to estimate a calibration factor used correlations across each of the 33 sites. 31 of the sites had strong Spearman correlation coefficients exceeding 0.50, and 16 of those sites had Spearman correlation coefficients >0.7. Averaging across these 16 "best" sites gave a mean CF of 3.14 (SE 0.18).

A fourth approach used regression of the ALT hourly values on the AQS values. Outliers more than a factor 5 away from the overall FEM/ALT ratio were deleted, leaving 154,900 hourly values in the regression. The resulting equation was ALT $PM_{2.5} = -0.10$ (SE 0.0066) + 0.3409 (SE 0.0006) FEM (adj. $R^2 = 0.679$). This produced a CF estimate of 2.93 (0.01).

The four approaches produced an average CF estimate of 3.05 (SE = 0.05).

3.1. *Effect of wildfires*

The period of monitoring included three wildfires, each exceeding 100,000 acres burned. Each fire affected air quality readings for several weeks up to two months. Over the 77 weeks of the study, 13 weeks affected by the wildfires had greater than 50% increases in the anomalies (departures from the 77-week mean) for both the PurpleAir and FEM monitors (Fig. 1). There was a clear change in the FEM/ALT ratio during these weeks. For those 13 weeks the CF was 2.39 (0.06) compared to 3.21 (0.07) in the remaining 64 weeks. The difference was significant ($p < 0.001$) and likely due to the different composition, refractive index, or reduced density of the wildfire smoke from the normal $PM_{2.5}$ background level. Since not all FEM sites were affected by the fires, the difference is likely to be even greater than the approximate 25% decline observed. Delp and Singer (2020) studied the Camp wildfire of November 8–21, 2018 in California using 53 PurpleAir monitors within 11 km of 12 Northern California AQS sites. They found a Smoke

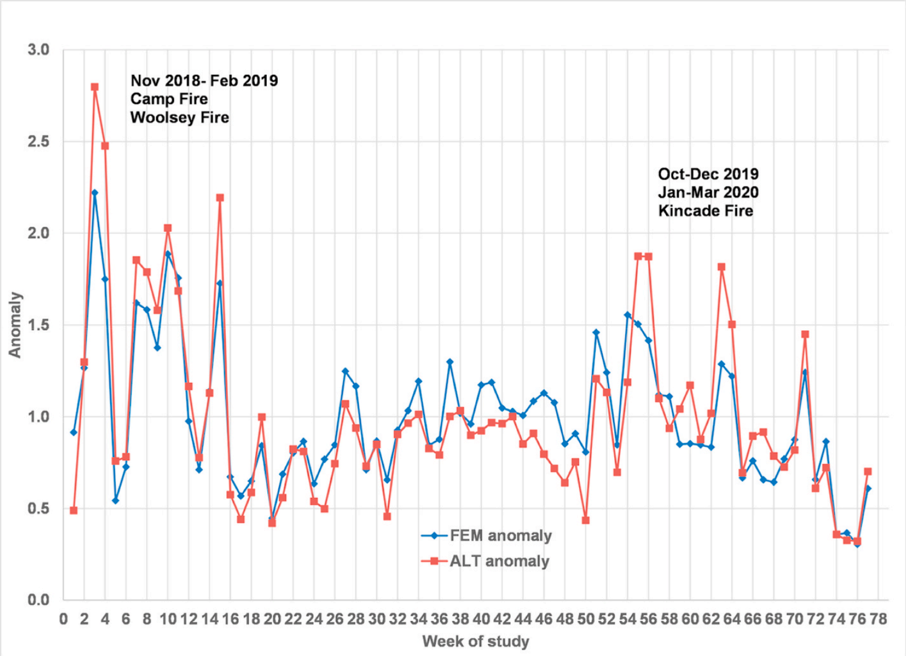


Fig. 1. Variation of $PM_{2.5}$ anomalies (ratios to 77-week mean values) associated with wildfires. The 13 weeks in which wildfires raised $PM_{2.5}$ readings by 50% or more had a lower average calibration factor (2.39) than the 64 weeks without wildfires (3.21).

Adjustment Factor of 0.42 for the PurpleAir monitors used. This reduction is in the same direction but larger than our value of 25% based on all monitors including those very distant and probably unaffected by the wildfires. The overall CF from these 77 weekly averages was 3.09 (SE 0.09), within the values observed for the hourly and site averages.

Summarizing, five different ways of determining the overall CF using hourly, weekly, or site-specific approaches gave CF values in the range of 2.9–3.1.

3.2. *Effect of RH*

Under conditions of increasing relative humidity (RH), inorganic atmospheric particles other than dust undergo deliquescence, a sudden increase in diameter and mass due to absorption of water (Seinfeld and Pandis, 2nd edition). The RH at which this occurs varies according to chemical composition, but is often about 60%–85% RH for common atmospheric salts such as KCl, Na_2SO_4, $(NH_4)_2SO_4$, etc. For H_2SO_4, however, the molecule grows in diameter and mass with no threshold RH required. Most of these molecules also exhibit hysteresis, maintaining the water in the particle as RH declines, even well below the RH value at which the particle first took on water. For optical particle monitors, then, an increase in RH will produce an increase in particle size and a corresponding increase in the apparent mass. Since many of these increases will occur at fairly high RH values near 60%, we would not expect a strictly linear effect, but a convex upward shape. We can test this by observing the effect, if any, of RH on the FEM monitors. Many of these are beta attenuation monitors (BAM), in which the sampler inlet is heated to reduce moisture deposition in the pipeline, and thus these FEM monitors would not be expected to be influenced strongly by RH. In fact, not much effect can be seen: a linear regression is $PM_{2.5}$ (FEM) = 7.186 (0.050) + 0.0037 × RH (R^2 adj. = 0.000054), so the entire range is from 7.19 at RH = 0% to 7.56 at RH = 100%. We can then consider the ratio of the ALT to the FEM $PM_{2.5}$ estimates (Fig. 2) using distance-weighted least squares, since we expect nonlinear effects to appear at relatively high RH. Indeed, the graph shows an inflection point at about 50% RH when the slope increases until about 80% RH and then stays nearly constant until 100% RH. This would be expected if some particles were reaching their deliquescence limit at these higher RH values. The curve shows a small effect of RH even at the lowest values, a finding also of Zheng et al. (2018). A complicating factor is that the RH values recorded by the PurpleAir monitor are considerably lower than the RH in the surrounding air. A mean decrease of about 15% was noted for two indoor PurpleAir monitors studied over a period of several months, along with a corresponding increase of about 8 °C due to the near-infrared lasers.

Although it might be possible to use the curve shown in Fig. 2 above to calculate the CF as a function of RH, the extreme variability of the 1-h $PM_{2.5}$ values suggests that not much gain would be expected from using the exact relationship. Also, any extended data set (say a day or more) will be likely to experience substantial diurnal swings in RH. The average diurnal variation for the 33 PurpleAir sites is shown in Fig. 3.

Given the range of 30%–55% in expected diurnal variation (the diurnal variation curve would move upwards and downwards over the seasons but continue to include a substantial range on most days), using the overall calibration factor of about 3 would provide a reasonable estimate of $PM_{2.5}$ in many cases.

3.3. *Comparison of alternate (ALT) $PM_{2.5}$ estimates to plantower CF1 system*

The ALT and CF1 hourly $PM_{2.5}$ estimates are very highly correlated (Fig. 4). The CF1 estimates are about 4.5 times the ALT estimates (N = 159,400); R^2 = 98%). This suggests that the unknown algorithm adopted by Plantower may be closely related to our approach using the particle numbers in the three smallest size categories. Based on this close agreement of CF1 and ALT measurements, we can calculate that for the 2.93–3.15 range of the calculated CF for the ALT system, the calculated CF for the CF1 system would lie between 0.65 and 0.72. This corresponds to the approximate 40% overestimate of the CF1 series as found in other studies (AQ-SPEC (2016); Kelly et al., 2017; Sayahi et al., 2019).

3.4. *Precision and coefficient of variation*

As explained in the Appendix, the ALT system was compared to the

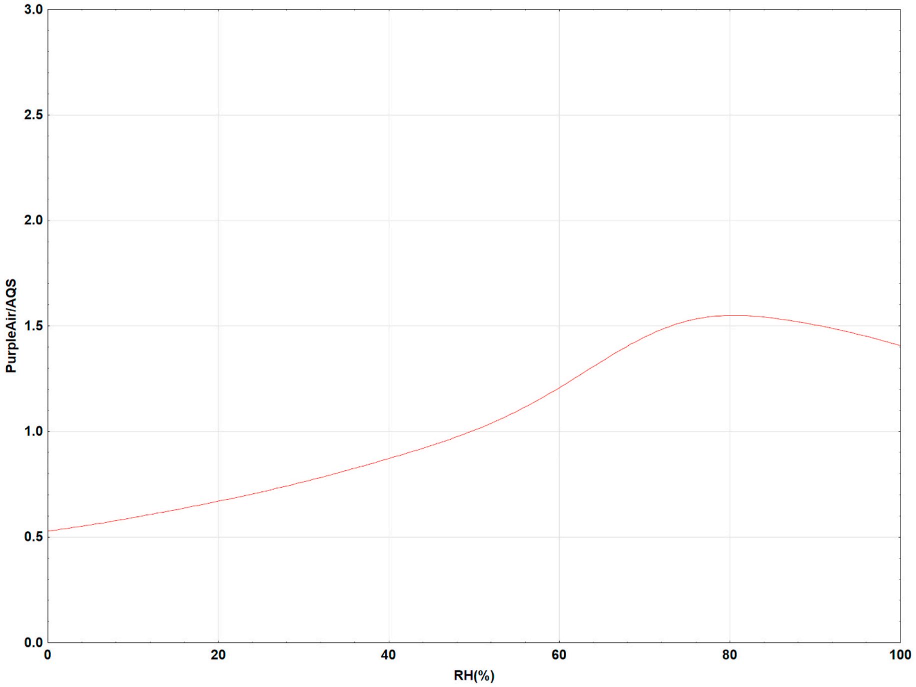


Fig. 2. Ratio of PurpleAir (CF 3) ALT $PM_{2.5}$ estimates to FEM $PM_{2.5}$ as a function of RH, using distance-weighted least squares regression. There is a slow nearly linear increase from 0% to about 50%–60% but then a faster increase until leveling off at about 80%.

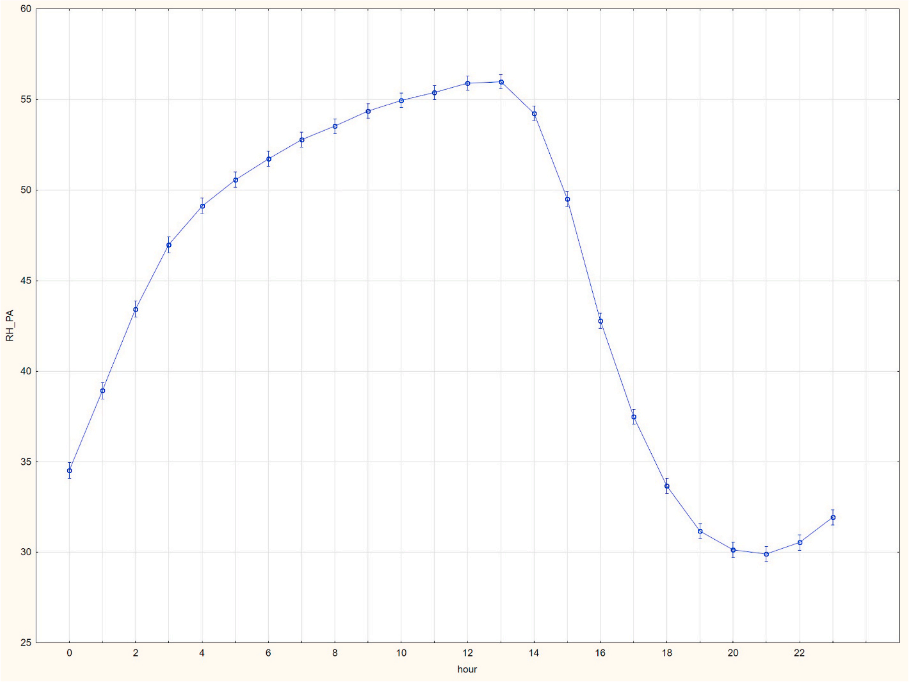


Fig. 3. Diurnal variation for PurpleAir (PA) RH values averaged across each hour of the day. (N = 174,169).

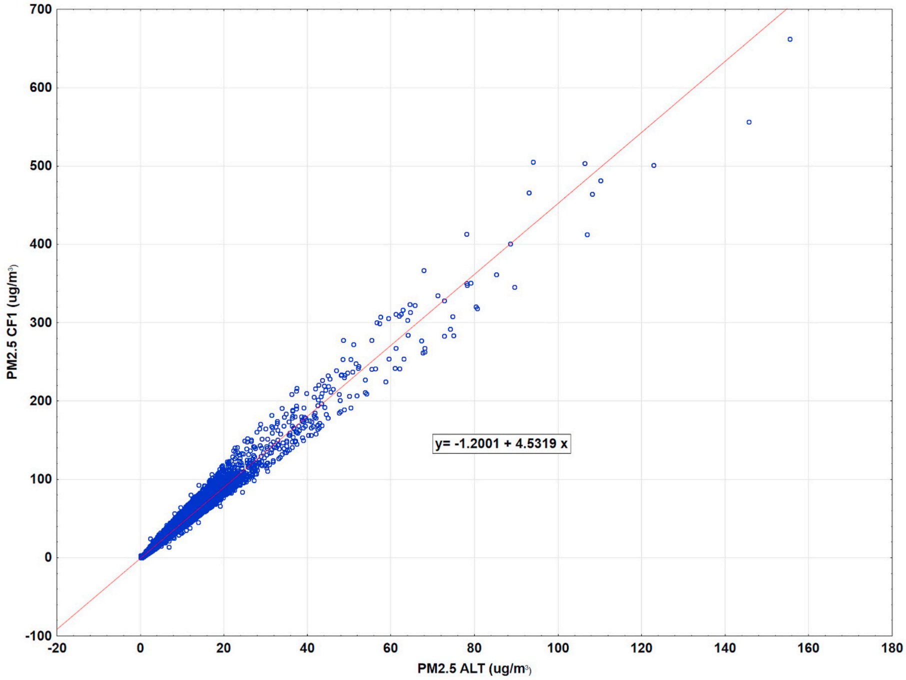


Fig. 4. Scatterplot of $PM_{2.5}$ estimates for the CF1 and ALT systems.

CF1 system in experiments carried out using duplicate collocated PurpleAir monitors in two homes in Santa Rosa and Redwood City. The data were collected between Jan 10, 2019, and March 26, 2020, a total of about 350,000 measurements over 433 days. During this time, mean precision in the ALT series was excellent at 4%–6%, compared to 7%–14% for the CF1 series. Over 358 days in Santa Rosa when only normal indoor activities were carried out, the COV for the ALT series improved by 26%; the corresponding value over 85 days in Redwood City was 22%.

3.5. Size distribution of $PM_{2.5}$ concentrations

The CF 1 (and CF ATM) series distort the distribution of the PM_1 and $PM_{2.5}$ mass concentrations due to cutting off low concentrations and replacing them with zero. In our example dataset from Santa Rosa, for one PurpleAir monitor there were 353,551 observations, but 73,266 (20.7%) reported $PM_{2.5}$ values of zero for the CF1 data series (see Appendix, Figures A1 and A2). This distorts the distribution to the point that no fit could be determined by the software program (Statistica) using the CF1 distribution, although a log-normal distribution gave a good fit to the ALT distribution.

Table 2
Comparison of the CF1 and ALT series LODs for $PM_{2.5}$ in two locations, with fraction of observations exceeding the LOD.

	Santa Rosa		Redwood City	
	CF1	ALT	CF1	ALT
LOD (μg/m^3)	1.77	1.15	2.55	1.23
fraction > LOD	0.13	0.48	0.51	0.90

3.6. Limits of detection (LOD)

To calculate an LOD for a continuous monitor such as the PurpleAir instrument, we use a method developed in Wallace et al. (2010). In brief, the method searches for the lowest concentration above which mean concentrations exceed their standard deviations by a factor of 3 more than 95% of the time. The LOD was improved by 30%–50% for the ALT system compared to the CF1 system (Table 2). (The ALT system LODs were multiplied by the CF1/ALT ratio of 4.5 to make the comparisons shown.)

3.7. Limitations of the study

No information was available on the siting of the outdoor PurpleAir monitors. The manufacturer recommends mounting the instrument away from vents and foliage; mounting high enough off any nearby surface to avoid splashing water or snow; and mounting it with the open surface pointing down. The finding by He et al. (2020) of imperfect assignment of particles to the size categories affects our $PM_{2.5}$ estimates directly. The finding by Zusman et al. (2020) that using the overlapping observations gave slightly better results than using the non-overlapping size categories, is further evidence of problems in assigning particles to their proper size category. These findings add uncertainty of unknown magnitude to our estimates.

The FEM monitors are subject to occasional quality assurance procedures carried out by EPA (https://www.epa.gov/outdoor-air-quality-data/pm25-continuous-monitor-comparability-assessments). The program provides gravimetric measurements (Federal Reference Methods, or FRM) to compare to the FEM measurements. Ten of the 27 FEM sites included in this study were visited in recent years (See Excel file in Supplementary Information). Only 3 sites, all in Fresno, had biases <10%. Six of the seven remaining sites had positive biases exceeding 20% (and up to 38%) during the years (2018–2020) overlapping our study. The remaining site had both high positive and high negative bias in adjacent years. Overall, however, the assessment program has roughly equal numbers of positive and negative bias of the FEM sites.

We compared daily $PM_{2.5}$ averages at FEM sites to nearby (500 m) FRM sites. Only 16 of our 27 FEM sites were associated with FRM sites. There were 2290 days of data recorded. The mean (median) FEM/FRM ratios ranged from 0.74 to 1.55 (0.64–1.67). A linear regression of FEM on FRM had a slope of 0.89 and an intercept of 1.56 with an R^2 of 85%. The FEM error was <20% for only 8 of 16 cases, and >30% for 4 cases. The errors had a slight tendency toward overestimates, with 10 of 16 FEM/FRM ratios >1 and a mean (median) error of 1.11 (1.15). These errors clearly contribute to the "noise" observed across the comparisons of PurpleAir monitors to FEM monitors. However, we are unable to calculate an overall bias, if any, based on only 16 of the 27 FEM sites in our database.

The temperature values reported by the PurpleAir monitors are influenced by the heated internal region due to the near-infrared lasers. In a long period of comparison of an indoor monitor in a Santa Rosa home, the increase in temperature averaged 8 °F. A corresponding average decrease of 15% in RH was also measured. The particles entering the monitors thus undergo a sharp change in both T and RH occurring across a few cm. Depending on how rapidly the particles adjust to the changing environmental conditions, the effects associated with the recorded RH may actually relate to some mixture of ambient RH and the internal RH of the monitor.

4. Conclusions

We have applied an alternative method (ALT) to calculate $PM_{2.5}$ estimates in a transparent and reproducible way rather than relying on the unknown algorithm provided by Plantower for their PMS 5003 sensors used in PurpleAir monitors. We have tested the results using 433 days of monitoring in two homes and demonstrate that the ALT method has superior performance with respect to precision, coefficient of variation, limit of detection, and less distorted size distribution.

To calibrate PurpleAir outdoor monitors using the ALT method of estimating $PM_{2.5}$, we compared hourly averages of 33 PurpleAir monitors to the hourly averages of 27 official FEM sites within 500 m distance over 20 months (77 weeks). Several different ways of computing the calibration factor (CF) resulted in a range of estimates from 2.93 to 3.15, with standard errors ranging from 0.01 to 0.18. This value near 3.0 for 27 outdoor PurpleAir monitors was also the estimated CF for 4 *indoor* PurpleAir monitors collocated with research-grade monitors over a year-long study (Wallace et al., 2020).

The time period included major wildfires affecting observations over 13 weeks, and we detected a difference in the CF during the 13 weeks from a value of 2.78 to a value of 3.12 in the remaining 64 weeks. We also observed a nonlinear effect of relative humidity (RH), with the PurpleAir results compared to FEM values increasing slowly at levels from 0 to 50%, more rapidly at levels from 50% to 75%, then remaining constant or slightly declining up to 100%. Although this behavior could conceivably be incorporated into a model, considering the wide range of uncertainty and particularly the fact that any period of monitoring exceeding a day will encounter swings in RH over ranges of about 25%–30%, employing the central estimate near 3 would be sufficient in many cases.

We also found a CF ranging between 0.65 and 0.72 for the CF1 data series provided by PurpleAir, corresponding to an approximate 40% overestimate of $PM_{2.5}$ values using the CF1 series. Overestimates of PurpleAir monitors in the 30%–40% range have been found in other studies.

Our study is the first to provide a $PM_{2.5}$ calibration factor for PurpleAir monitors based on 18 months of hourly observations by 33 monitors near 27 FEM sites. We have also shown that an existing alternative (ALT) method of calculating $PM_{2.5}$ outperforms the CF1 and ATM data series offered by PurpleAir. These results are presently being incorporated into the PurpleAir database as a correction factor "overlay" available to all users.

Funding

Funding for the portion of this study concerned with calibration of the Purpleair monitors was provided by the Multi-Angle Imager for Aerosols (MAIA) science team and the Multi-angle Imaging SpectroRadiometer (MISR) science team, both at the Jet Propulsion Laboratory (JPL), California Institute of Technology, led by D. Diner (Subcontract # 1588347 and 1363692). No funding was provided for the study described in the Appendix. The work of J. Sarnat and Y. Liu was supported by the MAIA science team at the JPL, California Institute of Technology (Subcontract #1588347), and the National Institute of Environmental Health Sciences of the National Institutes of Health (Award # R01-ES027892). The content is solely the responsibility of the authors and does not necessarily represent the official views of NASA or NIH.

CRediT authorship contribution statement

Lance Wallace: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft. **Jianzhao Bi:** Methodology, Software, Validation, Investigation, Formal analysis, Writing – review & editing. **Wayne R. Ott:** Methodology, Software, Validation, Investigation, Formal analysis, Writing – review & editing. **Jeremy Sarnat:** Methodology, Validation, Investigation, Formal analysis, Writing –

review & editing. **Yang Liu:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Adrian Dybwad of PurpleAir corrected a pair of misapprehensions in an earlier version of this paper. The authors appreciate the commitment of PurpleAir to creating a completely open and accessible database of thousands of these sensors.

Appendix

Since our study uses an alternative (ALT) method of calculating $PM_{2.5}$, this Appendix presents an explanation of the method and a comparison with the corresponding $PM_{2.5}$ estimates supplied by PurpleAir. These concentrations are provided with no explanation from the sensor manufacturer (http://www.plantower.com/en/) of how they are calculated. The alternative method presented here is a standard method employed in many optical particle monitors with multiple particle size measurement capabilities. It is transparent and reproducible by any user. It depends on the particle numbers in three size categories as reported by PurpleAir. The basic approach is to select an intermediate particle diameter within each size category, calculate the number of particles and the associated particle volume (assuming sphericity) and resulting particle mass (assuming an arbitrary density). The resulting estimates of particle mass (e.g., $PM_{2.5}$) can then be compared to those supplied by Plantower.

The PurpleAir data include two data series: CF1 and ATM. The CF1 series is described as “for laboratory use” and the ATM for “atmospheric” conditions. Many of the above studies (Table 1) employ the PM values reported in either the CF1 or ATM data series. However, we find serious deficiencies in both of these data series. For example, both series report multiple values of zero for typical indoor and outdoor measurements; yet in our studies of scores of sites with hundreds of thousands of observations there are never cases when the particle numbers in the 0.3–0.5 μm or 0.5–1 μm size categories are zero. The values of zero are apparently applied by the unknown Plantower algorithm to concentrations below some arbitrary cutoff number.

Another deficiency is the relation between the CF1 and ATM data series. A two-week period between 4/25/19 and 5/3/19 during which both high and low concentrations of $PM_{2.5}$ were produced provided data comparing the two series at a wide range of concentrations. For $PM_{2.5}$, it is clear that for low concentrations (less than about 28 $\mu g/m^3$), CF1$\equiv$ ATM (Fig. S1). A breakpoint occurs near 28 $\mu g/m^3$, after which the ATM concentrations increase more slowly than the CF1 concentrations (CF1/ATM ratio increases). Another breakpoint appears at 78 μg/m3, at which point the CF1 value is close to 50% higher than the ATM value, where it remains for all values > 78 $\mu g/m^3$. A comparison of the ALT series with the CF1 series shows a linear relationship, but a comparison with the ATM series is nonlinear (Fig. S2).

Because of these deficiencies, we examined whether an alternative (ALT) method of estimating particle mass concentrations could perform better than the CF1 or ATM series.

Outline of alternative (ALT) method

The approach adopted here is identical to that recommended by manufacturers of optical particle monitors capable of collecting simultaneous data on multiple particle size categories (e.g., the TSI Model 3330 (https://tsi.com/products/particle-sizers/particle-size-spectrometers/optical-particle-sizer-3330/)).

1. Calculate the number *N* of particles in each of the size categories 0.3–0.5 μm, 0.5–1 μm, 1–2.5 μm, 2.5–5 μm and 5–10 μm. Note: Since the Plantower output provides the total number of particles >0.3 μm, the total number >0.5 μm, etc., one must subtract the total number >0.5 μm from the total number >0.3 μm to arrive at the number of particles in the 0.3–0.5 μm size category. Similar subtractions will result in five size categories under 10 μm: 0.3–0.5 μm, 0.5–1 μm, 1–2.5 μm, 2.5–5 μm, 5–10 μm.
2. Select an average diameter *D* for each of the five size categories up to 10 μm. For example, the average particle diameter *D* in the 0.3–0.5 μm size category must be between 0.3 and 0.5 μm, so we can approximate it by the midpoint (0.4 μm) or by the geometric mean of the size boundaries (0.387 μm). Several manufacturers use the geometric mean so we use it throughout this work.
3. Calculate the total particle volume *V* in each size category. This is given by multiplying the number of particles *N* times the volume per particle:

$$V = N\pi D^3 / 6$$

4. Multiply by a density ρ to arrive at an estimate of the mass concentration *M* in each size category. We choose ρ to be the density of water, a choice that is also made by the Model 3330 manufacturers:

$$M = \rho V$$

5. Add the appropriate size categories to estimate PM_1, $PM_{2.5}$ and PM_{10}. PM_1 is the sum of the mass concentrations in the two smallest size categories: 0.3–0.5 μm and 0.5–1 μm. $PM_{2.5}$ is the sum of the mass concentrations in the two smallest size categories plus the next largest size category 1–2.5 μm. PM_{10} adds the mass in the two largest size categories (2.5–5 μm and 5–10 μm) to the three smallest size categories.
6. Calculate a calibration factor (CF) for the aerosol mixture being monitored by comparison to gravimetric studies of the same mixture or to research-grade monitors that have themselves been calibrated using gravimetric measurements.

Comparison of ALT method to plantower CF1 method

To compare results of the ALT method with the PM mass concentrations reported by Plantower, we consider measurements by two collocated PurpleAir monitors in each of two homes, one in Santa Rosa, CA, and the other in Redwood City, CA.

In both homes, the PurpleAir monitors each contained two PA-II sensors, model PMS 5003, plus a sensor measuring temperature, relative humidity (RH) and atmospheric pressure (Bosch Model BME 280) (https://www.bosch-sensortec.com/products/environmental-sensors/humidity-sensors-bme280/). In the Santa Rosa home, the monitors were placed on the front of a dresser 33" (83 cm) high 0.5 m from the wall in a 30 m^3 room in a single-story 385 m^3 private residence housing two people. The house employs forced air with a central fan on at all times. There is a return air filter with electret fibers. Blower door tests have been performed on the home twice, and air exchange measurements have been made multiple times, with results indicating a tight home with average outdoor air exchange rates of 0.15 (SD 0.05) h^{-1}.

The data were collected between Jan 10, 2019 and March 26, 2020, a total of about 350,000 measurements over 433 days. The PurpleAir instruments took samples every 80 s for the first 6 months and every 2 min thereafter. On most days ($N = 379$) the room's door was open to the rest of the house. On those days, normal household activities such as cooking were carried out. Neither resident was a smoker. On 55 days, the room was shut off from the rest of the house, the floor register was sealed, the central fan was turned off, and experiments were run using various sources of fine and ultrafine particles such as candles, laboratory hot plates, toaster ovens, and other typical particle sources. A full description of the experiments is contained in Wallace et al. (2020). The experiments typically elevated the $PM_{2.5}$ concentrations to levels above 1 mg/m^3. These high concentrations were permitted to decay over the next few hours before the room was opened up again.

In the Redwood City home, the two monitors were set up in a 43 m^3 bedroom. The 2-story house has a volume of 430 m^3. There is a gas furnace with forced-air ventilation, but no air conditioner. There is one resident nonsmoker. The monitors collected mostly 2-min average data from May 21, 2019 through April 15, 2020. A 3-month dataset (July 23, 2019 through October 22, 2019) was chosen for analysis. This period included 7 days during which experiments on various particle sources were carried out in the test room, which was closed off from the rest of the house during those days only. A full description of the experiments at this second home is provided in a companion paper Ott et al., 2021. These periods of high concentrations were analyzed separately from the remainder of the data (85 days).

Data Analysis. Raw data were examined for duplicates. Duplicates of every measurement occurred between March and November of each year between 2017 and 2020, a period coinciding with daylight savings time. This error affects all downloaded reports of PurpleAir concentrations during these months (up to April 10, 2020). The PurpleAir organization is working on correcting the error. Our study has eliminated all duplicates.

All data were examined for outliers, such as large negative values, and they were removed. These errors were rare; for one sensor, there were 3 negative values for $PM_{2.5}$ out of 353,551 measurements, an error rate of about 10^{-5}.

The initial comparison of the ALT data series with the CF1 and ATM data series resulted in a strong linear relationship ($R^2 > 0.999$) with the CF1 data series, but a nonlinear relationship with the ATM series. Therefore, the following analysis compares our ALT series with the CF1 series only.

Results are presented for PM_1 and $PM_{2.5}$. Although we have also considered PM_{10}, the relationship between the ALT and CF1 series is not linear. Also, the LOD is often high, such that the majority of PM_{10} concentrations as measured by the CF1 data series are below the LOD. We will not, therefore, present results for PM_{10}.

For both PM_1 and $PM_{2.5}$, we can compare our ALT values to the CF1 values considering bias, precision, coefficient of variation (COV), size distribution, and limit of detection (LOD).

Precision and bias

Bias across the four independent sensors was calculated with respect to the average of the four values. The bias was then corrected and the bias-corrected precision was calculated using the equation.

$$Precision = abs(A-T)/(A+T)$$

where A is a measurement from one sensor and T is the average of the four sensors.

Over the 433 days measured at the Santa Rosa site, mean precision in the ALT series was excellent, ranging across the four sensors from 3% to 6% for PM_1 and from 4% to 6% for $PM_{2.5}$ (Table A1). Corresponding values for the CF1 series are 9%–28% and 7%–14%. Bias relative to the mean of the four sensors in the ALT series ranged from 0.95 to 1.06 for PM_1 and from 0.93 to 1.06 for $PM_{2.5}$. Corresponding values for the CF1 series are 0.86–1.12 and 0.96 to 1.03.

Table A1

Precision and bias for the ALT series compared to the CF1 series for PM_1 **and $PM_{2.5}$ measured during 433 days in Santa Rosa. Sensors A and B in monitors 1 and 2 are labeled as 1A, 1B, 2A, and 2B.**

PM size	Data series	1A	1B	2A	2B
Precision					
PM_1	ALT	0.03	0.05	0.04	0.06
	CF1	0.11	0.10	0.09	0.28
$PM_{2.5}$	ALT	0.05	0.06	0.04	0.06
	CF1	0.07	0.09	0.12	0.14
Bias					
PM_1	ALT	1.06	0.99	0.95	1.00
	CF1	1.12	1.01	1.01	0.86
$PM_{2.5}$	ALT	1.01	1.00	1.06	0.93
	CF1	1.02	0.96	1.03	0.99

For the 3-month period from July 23 to October 22, 2019 at the Redwood City location, two monitors were col-located. In this case, the mean precision for the CF1 series was 9.6% compared to 2.86% for the ALT Series. Another way of looking at the difference is that 98% and 99% of values in the ALT Series had precision better than 10% and 15%, respectively, compared with 73% and 83% in the CF1 series.

Coefficient of variation (COV)

This is a measure of the amount of variation within a dataset. It is calculated as the Relative Standard Deviation (RSD), the SD divided by the mean. The smaller the COV, the more dependable will be the estimate of the mean.

Since the aerosol mixtures from the experiments at the two sites may vary considerably from the normal indoor aerosol, we compare the $PM_{2.5}$ results for the experimental and non-experimental days separately (Table A2).

Table A2
Comparison of CF1 to ALT $PM_{2.5}$ data series in Santa Rosa and Redwood City locations: mean, SD, and COV for days with and without major sources of indoor particles.

	Santa Rosa (1/10/19 to 3/26/20)				Redwood City (7/23/19 to 10/22/19)			
	Experimental days		No experiments		Experimental days		No experiments	
	$PM_{2.5}$ CF1	$PM_{2.5}$ ALT	$PM_{2.5}$ CF1	$PM_{2.5}$ ALT	$PM_{2.5}$ CF1	$PM_{2.5}$ ALT	$PM_{2.5}$ CF1	$PM_{2.5}$ ALT
N	54	54	378	378	7	7	85	85
mean	29.20	5.61	3.97	0.91	118.8	24.2	3.30	0.81
SD	19.57	3.68	7.96	1.35	10.4	2.42	2.58	0.49
RSD (COV)	0.67	0.66	2,00	1.48	0.09	0.10	0.78	0.61
COV improvement	2%		26%		−10%		22%	
Ratio (CF1/ALT)	5.2		4.4		4.9		4.1	

In both locations, the COVs of the CF1 and ALT series were not much different on the experimental days (higher concentrations), but for the lower concentrations, the COV was substantially lower (better) for the ALT series, by 26% at the Santa Rosa location and 22% at the Redwood City location.

Calculation of the limit of detection (LOD)

To calculate an LOD for a continuous monitor such as the PurpleAir instrument, we use a method developed in Wallace et al. (2010). To be 95% certain that a reported observation is > 0, we require that the mean value of multiple collocated instruments be > 3 times the standard deviation. However, if the observations are in a series ordered by the mean of the instrument measurements, it can and typically does happen that an observation will meet this criterion, whereas at a higher concentration it does not. Then the LOD is *not* this lower concentration. Therefore, we search for the lowest concentration above which higher concentrations *always* have at least 95% of the calculated mean/SD ratios >3. For a large dataset, this can be done by considering “batches” of, say, 100 observations at a time in the dataset ordered by mean concentrations, and counting the number of cases in which the mean/SD ratio is < 3. As higher concentrations are examined, eventually a batch is found with fewer than 5 such cases. If testing more batches at yet higher concentrations never shows 5 or more such cases, we assume that the earlier batch contains the limit of detection.

Two large datasets were employed to calculate the LOD for PM_1 and $PM_{2.5}$ in the two locations: a Santa Rosa dataset running from Jan 30, 2020, to April 27, 2020, with 63,091 observations and a Redwood City dataset for the 3-month period from July 23 to October 22, 2019, with 65,319 observations. In all calculations, the bias of the four individual sensors compared to the mean was determined and the bias-corrected values were then analyzed to determine the mean and standard deviation of each observation.

A Santa Rosa database running from Jan 30, 2020, to April 27, 2020, with 63,091 observations was used to calculate PM_1 and $PM_{2.5}$ LODs for the CF1 and ALT series (Table A3). For the CF1 series, the LODs for PM_1 and $PM_{2.5}$ were 1.54 $\mu g/m^3$ and 1.77 $\mu g/m^3$, values that are comparable to those reported in other studies (e.g., the $PM_{2.5}$ LODs found for two PMS 5003 sensors (2.62–3.65 $\mu g/m^3$) in two seasons by Sayahi, Butterfield, and Kelly; Sayahi Supplementary Material, 2018). (Note: Since the LOD is determined by ordering all data by concentration and then determining the concentration above which the ratio of mean to standard deviation never exceeds 3 by as much as 5% for the remainder of the data, there is not an obvious approach for determining the uncertainty of the estimate.) Despite the low CF1 LODs found here, only 6% of the PM_1 concentrations and 13% of the $PM_{2.5}$ concentrations exceeded the LOD for the CF1 series. By comparison, the improved LODs for the ALT series resulted in 78% of PM_1 concentrations and 48% of $PM_{2.5}$ concentrations exceeding the LOD. (Note: The results for the ALT series have been multiplied by the factors relating the ALT series to the CF1 series for proper comparison.)

In the Redwood City location, the 3-month database with 65,309 observations as described above was used. The $PM_{2.5}$ LOD determined for the CF1 series was 2.55 $\mu g/m^3$. The LOD for the ALT series was 1.23 $\mu g/m^3$, about half that found for the CF1 series. For the CF1 data series, there were 32,100 (49.0%) values below the LOD. For the ALT data series, there were only 6900 (10.6%) values below the LOD. The improvement in the LOD was only by a factor of 2, but it led to almost 90% detected vs. less than 50% for the CF1 series. For the PM_1 CF1 series, the LOD was 1.98 $\mu g/m^3$, and 47% of values exceeded the LOD. For the ALT PM_1 series, the LOD was 0.50, and 99.4% of values exceeded the LOD.

Table A3
Comparison of the CF1 and ALT series LODs for PM_1 and $PM_{2.5}$ in two locations, with fraction of observations exceeding the LOD

	Santa Rosa		Redwood City	
	CF1	ALT	CF1	ALT
PM_1 LOD	1.54	0.52	1.98	0.50
fraction > LOD	0.06	0.78	0.47	0.99
$PM_{2.5}$ LOD	1.77	1.15	2.55	1.23
fraction > LOD	0.13	0.48	0.51	0.90

These results are concerning, because outdoor values can be quite low in many parts of the US, and thus the PM_1 and $PM_{2.5}$ estimates in the CF1 data series will be below the LOD (i.e., indistinguishable from zero) for a possibly large fraction of all measurements. This problem can be largely

overcome by calculating the PM_1 and $PM_{2.5}$ values in the ALT method demonstrated here.
Size distribution of PM concentrations

The CF1 and CF ATM series distort the distribution of the PM_1 and $PM_{2.5}$ mass concentrations due to cutting off low concentrations and replacing them with zero. In our example dataset, for one PurpleAir monitor there were 353,551 observations, but 73,266 (20.7%) reported $PM_{2.5}$ values of zero for the CF1 data series (Figure A1). This distorts the distribution to the point that a normal or lognormal fit could not be determined by the software program (Statistica).

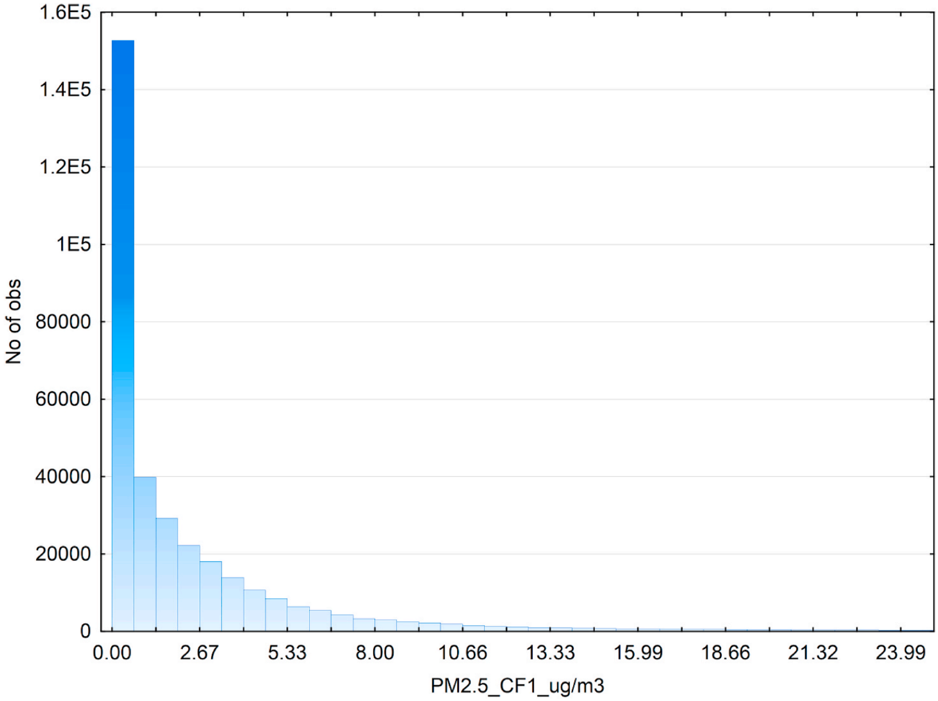


Fig. A1. Histogram of 353,511 reported $PM_{2.5}$ concentrations in the CF1 data series provided by Plantower. "No of obs" = number of observations. The 73,266 reported values of zero distort the distribution so extensively that no fit could be determined.

For the ALT series, the $PM_{2.5}$ size distribution is seen to approximate a lognormal distribution (Figure A2). No values of zero were reported.

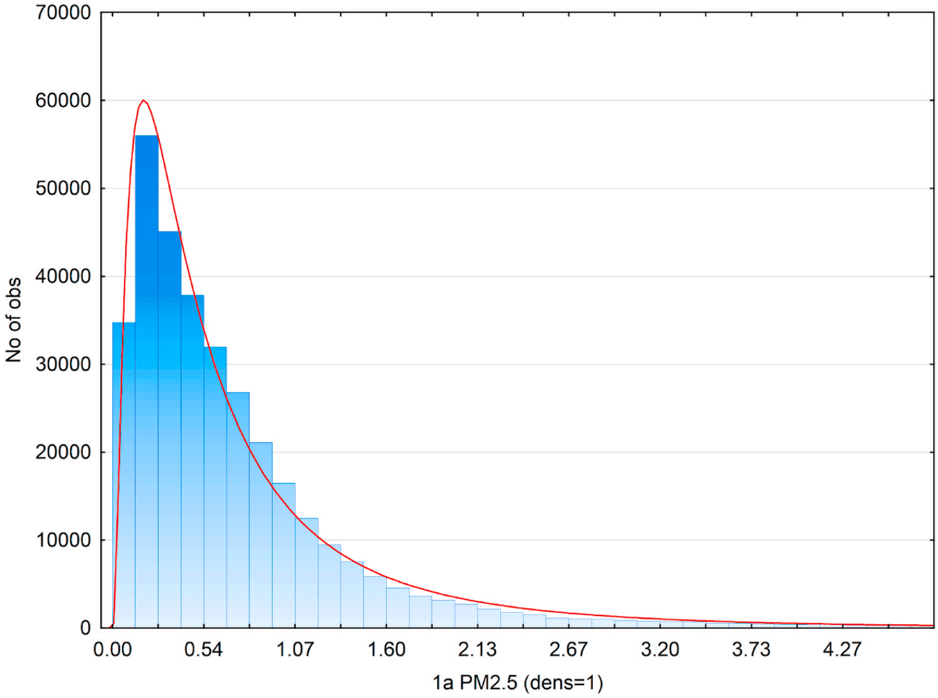


Fig. A2. Histogram of $PM_{2.5}$ concentrations reported by the ALT series. The shape is approximated by a lognormal fit.

Comparison of ALT and CF1 findings for the difference between $PM_{2.5}$ and PM_1

The PurpleAir estimates of PM_1 depend on the smallest two size categories, and the estimates of $PM_{2.5}$ depend on the smallest 3 size categories. This means the *difference* between PM_1 and $PM_{2.5}$ depends only on a single size category (1–2.5 μm). This difference should therefore be a linear function of the number of particles in that single size category. In our own estimate of the difference ($PM_{2.5}$- PM_1 ALT series), it is in fact a perfect ($R^2 = 1$) multiple of the number of particles in this size category. Yet this same difference in the CF1 series is far from a simple constant factor times the number of particles, explaining only 8% of the variation (Fig. S3).

Many of the problems in using the CF1 or ATM data series provided by Plantower stem from the arbitrary decision to substitute zero for measurements falling below some value. For example, this obviously leads to distortions of the mass distributions if a substantial number of measurements fall below this limit. But it also leads to worse precision and higher coefficients of variation, since the clustering of otherwise positive values at zero serve to widen the standard deviation. Statisticians generally oppose the practice of substituting a single value, such as 0, the LOD/2, or the LOD itself because it leads to exactly these problems. A review of many of these studies is found in Helsel (2010). His first conclusion reads: "**In general, do not use substitution**. Journals should consider it a flawed method compared to the others that are available and reject papers that use it …" (emphasis added).

However, this is not the only problem with the CF1 and ATM data series. The example discussed above of the difference between $PM_{2.5}$ and PM_1, which should rest entirely on the mass M of particles in the 1–2.5 μm size category, achieves only an 8% R^2 when plotted against M. The ALT method achieves an R^2 of precisely 1 in this case, as expected from theory. Apparently, the hidden algorithms of the Plantower approach assign values to measurements that in some way depart from using the mass calculated from the numbers of particles in the size categories.

The ALT method works well when calculating PM_1 or $PM_{2.5}$ from the two or three smallest size categories. However, when adding in the two largest size categories (2.5–5 μm and 5–10 μm) to estimate PM_{10}, it finds a very large LOD in both of our selected data series, such that <10% of the data exceeds the LOD. Other studies have also found problems with the Plantower PM_{10} estimates. The SPEC study (AQ-SPEC, 2016) found that the PM_{10} estimates showed a poorer relationship to the reference monitor than the PM_1 or $PM_{2.5}$ estimates. Therefore, we do not attempt to apply the ALT method to estimate PM_{10}.

Comparison of number estimates with research-grade monitors

Plantower number estimates in the three size categories from 0.3 to 2.5 μm were compared to those produced by a collocated reference monitor, the TSI Optical Sizer Model 3330 (Wallace, unpublished data). For the two smallest size categories, the Plantower number estimates were smaller by factors on the order of 8–10. For the sum of numbers in all three categories, the Plantower estimate was also smaller, by about a factor of 4, than the Model 3330 estimate. This observation may partially explain the factor of 4.5 by which the CF1 data series exceeds the ALT data series (using a density of 1).

A calibration factor for *indoor* $PM_{2.5}$ estimates has been recently estimated (Wallace et al., 2020). This was a study of an indoor aerosol using PurpleAir monitors collocated with research-grade SidePak monitors which themselves had been calibrated using gravimetric techniques (Zhao et al., 2020). This study arrived at a calibration factor of 3 for the indoor $PM_{2.5}$ estimates, in good agreement with our present estimate of 2.9–3.1 (SE < 0.2) for the outdoor $PM_{2.5}$ data calibrated using FEM monitors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosenv.2021.118432.

Conflicts of interest/competing interests

No conflict of interest exists.

Availability of data and material

All data available from author LW on request.

Code availability

Some R scripts were written for this study. Contact author JB. Software applications include Excel and Statistica.

References

AQ-SPEC, 2016. http://www.aqmd.gov/docs/default-source/aq-spec/field-evaluations/purpleair—field-evaluation.pdf. (Accessed 19 December 2020).

Becnel, T., Tingey, K., Whitaker, J., Sayahi, T., Le, K., Goffin, P., Butterfield, A., Kelly, K., Gaillardon, P.E., 2019. A distributed low-cost pollution monitoring platform. IEEE Internet of Things Journal 6 (6), 10738–10748.

Bi, J., Wildani, A., Chang, A.H.H., H, H., Liu, Y., 2020. Incorporating low-cost sensor measurements into high-resolution $PM_{2.5}$ modeling at a large spatial scale. Environ. Sci. Technol. (Article ASAP).

Bi, J., Wallace, L.A., Sarnat, J.A., Liu, Y., 2021. Characterizing outdoor infiltration and indoor contribution of PM2.5 with citizen-based low-cost monitoring data. Environ. Pollut. 276, 116763.

Bulot, F.M.J., Johnston, S.J., Basford, P.J., Easton, N.H.C., Apetroaie-Cristea, M., Foster, G.L., Morris, A.K.R., Cox, S.J., Loxham, M., 2019. Long-term field comparison of multiple low-cost particulate matter sensors in an outdoor urban environment. Sci. Rep. 9, 7497. https://doi.org/10.1038/s41598-019-43716-3.

Chen, L.J., Ho, Y.-H., Lee, H.-C., Wu, H.-C., Liu, H.-M., Hsieh, H.-H., Huang, Y.-T., Lung, S.-C.C., 2017. An open framework for participatory $PM_{2.5}$ monitoring in smart cities. IEEE Access 5, 14441–14454. https://doi.org/10.1109/ACCESS.2017.2723919. Accessed. (Accessed 18 December 2020).

Delp, W.W., Singer, B.C., 2020. Wildfire smoke adjustment factors for low-cost and professional $PM_{2.5}$ monitors with optical sensors. Sensors 20, 3685. https://doi.org/10.3390/s20133683.

Francis, A.S., Chee, F.P., Chang, J.H.W., Sentian, J., Dayou, J., Payus, C.M., 2019. Parametric model for estimation of mass concentration based on particle count distribution for ambient air monitoring. J. Phys. Conf. 1358 https://doi.org/10.1088/1742-6596/1358/1/012042, 012042.

Gupta, P., Doraiswamy, P., Levy, R., Pikelnaya, O., Maibach, J., Feenstra, B., et al., 2018. Impact of California fires on local and regional air quality: the role of a low-cost sensor network and satellite observations. GeoHealth 2, 172–181. https://doi.org/10.1029/2018GH000136. Accessed. (Accessed 18 December 2020).

He, M., Kuerbanjiang, N., Dhaniyala, S., 2020. Performance characteristics of the low-cost Plantower PMS optical sensor. Aerosol. Sci. Technol. 54 (2), 232–241. https://doi.org/10.1080/02786826.2019.1696015.

Helsel, D., 2010. Much ado about next to nothing: incorporating nondetects in science. Ann. Occup. Hyg. 54 (3), 257–262.

Kaduwela, A.P., Kaduwela, A.P., Jrade, E., Brusseau, M., Morris, S., Morris, J., Risk, V., 2019. Development of a low-cost air sensor package and indoor air quality monitoring in a California middle school: detection of a distant wildfire. J. Air Waste Manag. Assoc. 69 (9), 1015–1022. https://doi.org/10.1080/10962247.2019.1629362.

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. Environ. Pollut. 221, 491–500.

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. Environ. Sci. Technol. 53 (2), 838–849.

Magi, B.I., Cupini, C., Francis, J., Green, M., Hauser, C., 2019. Evaluation of $PM_{2.5}$ measured in an urban setting using a low-cost optical particle counter and a Federal Equivalent Method Beta Attenuation Monitor. Aerosol. Sci. Technol. 54, 147–159. https://doi.org/10.1080/02786826.2019.1619915.

Malings, C., anzer, 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. Sci. Technol. 54, 160–174. https://doi.org/10.1080/02786826.2019.1623863.

Masic, A., Bibic, D., Pikula, B., 2019. On the Applicability of Low-Cost Sensors for Measurements of Aerosol concentrations, 0452-0456. In: Katalinic, B. (Ed.), Proceedings of the 30th DAAAM International Symposium. Published by DAAAM International, ISBN 978-3-902734-22-8. https://doi.org/10.2507/30th.daaam.proceedings.060, 1726-9679, Vienna, Austria.

Ott, W.R., Zhao, T., Cheng, K-C, Wallace, L.A., Hildemann, L.M., 2021. Measuring indoor fine particle concentrations, emission rates, and decay rates from cannabis use in a residence. Atmos. Environ. 10, 100106. https://doi.org/10.1016/j.aeaoa.2021.100106.

Plantower, 2016. Accessed. https://www.aqmd.gov/docs/default-source/aq-spec/resources-page/plantower-pms5003-manual_v2-3.pdf. (Accessed 17 January 2020).

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

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. Accessed. (Accessed 19 December 2020).

Tryner, J., Quinn, C., Windom, B.C., Volckens, J., 2019. Design and evaluation of a portable $PM_{2.5}$ monitor featuring a low-cost sensor in line with an active filter sampler. Environmental Science Processes and Impacts 21, 1403–1415.

US EPA, 2017. https://www.epa.gov/air-sensor-toolbox/how-use-air-sensors-air-sensor-guidebook.

Walker, 2018. Accessed. http://conference2018.resnet.us/data/energymeetings/presentations/RESNET2018_LBL_LowCostMonitors_walker.pdf. (Accessed 19 December 2020).

Wallace, L.A., Wheeler, A., Kearney, J., Van Ryswyk, K., You, H., Kulka, R., Rasmussen, P., Brook, J., Xu, X., 2010. Validation of continuous particle monitors for personal, indoor, and outdoor exposures. J. Expo. Sci. Environ. Epidemiol. 21, 49–64.

Wallace, L.A., Ott, W.R., Zhao, T., Cheng, K.-C., Hildemann, L., 2020. Secondhand exposure from vaping marijuana: concentrations, emissions, and exposures determined using both research-grade and low-cost monitors. Atmos. Environ. X. https://doi.org/10.1016/j.aeaoa.2020.100093.

Wang, K., Chen, F.E., Au, W., Zhao, Z., Xia, Z.-L., 2019. Evaluating the feasibility of a personal particle exposure monitor in outdoor and indoor microenvironments in Shanghai, China. Int. J. Environ. Health Res. 29, 209–220. https://doi.org/10.1080/09603123.2018.1533531.

Wang, Z., Delp, W.W., Singer, B.C., 2020. Performance of Low-Cost Indoor Air Quality Monitors for $PM_{2.5}$ and PM_{10} from Residential Sources. Building and Environment. https://doi.org/10.1016/j.buildenv.2020.106654. Accessed. (Accessed 19 December 2020).

Williams, R., Vasu, K., Snyder, E., Kaufman, A., Dye, T., Rutter, A., Russell, A., Hafner, H., 2014. Air Sensor Guidebook. U.S. Environmental Protection Agency, Washington, DC. EPA/600/R-14/159 (NTIS PB2015-100610).

Zhao, T., Cheng, K.-C., Ott, W.R., Wallace, L.A., Hildemann, L.M., 2020. Characteristics of Secondhand Cannabis Smoke from Common Smoking Methods: Calibration Factor, Emission Rate, and Particle Removal Rate. Atmospheric Environment. https://doi.org/10.1016/j.atmosenv.2020.117731.

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 11 (8), 4823–4846. https://doi.org/10.5194/amt-11-4823-2018.

Zusman, M., Schumacher, C.S., Gassett, A.J., Spalt, E.W., Austin, E., Larson, T.V., Arvlin, G.C., Seto, E., Kaufman, J.D., Sheppard, L., 2020. Calibration of low-cost particulate matter sensors: model development for a multi-city epidemiological study. Environ. Int. 134, 105329.