Revised manuscript. Changes are shown in red. For clarity old figures are not retained in this version, but the figure caption will indicate "New".

An evaluation of the U.S. EPA's correction equation for PurpleAir Sensor data in smoke, dust, and wintertime urban pollution events

Daniel A. Jaffe[1,2], Colleen Miller[1,4], Katie Thompson[1], Brandon Finley[1], Manna Nelson[3], James Ouimette[5], Elisabeth Andrews[6,7]

[1]University of Washington Bothell, School of STEM, Bothell, WA 98011, USA

[2]University of Washington Seattle, Department of Atmospheric Sciences, Seattle, WA 98195, USA

[3]Seattle University, Department of Civil and Environmental Engineering, Seattle, WA 98122, USA

[4]Now at University of California-Irvine, Department of Chemistry, Irvine, CA 92697, USA

[5]Santa Rosa, CA 95409, USA

[6]NOAA Global Monitoring Laboratory, Boulder, CO 80305, USA

[7]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, USA

Correspondence: Daniel A. Jaffe (djaffe@uw.edu)

Abstract

PurpleAir Sensors (PASs) are low-cost tools to measure fine particulate matter (PM) concentrations and are now widely used, especially in regions with few regulatory monitors. However, the raw PAS data have significant biases, so the sensors must be calibrated to generate accurate data. The U.S. EPA recently developed a national correction equation and hasve integrated corrected PAS data onto its AirNow website. This integration results in much better spatial coverage for $PM_{2.5}$ (particulate matter with diameters less than 2.5 µm) across the U.S. The goal of our study is to evaluate the EPA correction equation for three different types of aerosols: typical urban wintertime aerosol, smoke from biomass burning, and mineral dust.

We identified 50 individual pollution events, each having a peak hourly $PM_{2.5}$ concentration of at least 47 µg m^{-3} and a minimum of 3 hours over 40 µg m^{-3}, and characterized the primary aerosol type as either typical urban, smoke, or long-range transported dust. For each event, we paired an PAS sampling outside air with a nearby regulatory $PM_{2.5}$ monitor to evaluate the agreement. All 50 events show statistically significant correlations (R values between 0.71–1.00) between the hourly PAS and regulatory data, but with varying slopes. We then corrected the PAS data using either the correction equation from Barkjohn et al. (2021) or a new equation that is now being used by the U.S. EPA for the AirNow Fire and Smoke Mmaps (U.S. EPA, 2022b). Both equations do a good job at correcting the data for smoke and typical pollution events, but with some differences. Using the Barkjohn et al. (2021) equation,

we find ~~Using the standard EPA correction for the typical urban and smoke aerosols, we find average~~ mean slopes of 1.00 and 0.99 for urban and smoke aerosol events, respectively, for the corrected data versus the regulatory data. ~~and 0.99, respectively. This means that the standard EPA correction is highly effective at generating accurate data for these aerosol types.~~ For heavy smoke events, we find a small change in the slope at very high $PM_{2.5}$ concentrations (>600 µg m^{-3}), suggesting a ~20 % under-estimate in the corrected PAS data at these extremely high concentrations. Using the new EPA equation, we find slopes of 0.95 and 0.88 for urban and smoke events, respectively, indicating a slight underestimate in $PM_{2.5}$ using this equation, especially for smoke events~~.~~ For dust events, while the PAS and regulatory data still show significant correlations, the PAS data using ~~the~~ either ~~standard~~ correction equation underestimates the true $PM_{2.5}$ by a factor of 5-6.

We also examined several years of co-located regulatory and PAS data from a site near Owens Lake, CA, which experiences high concentrations of $PM_{2.5}$ due to both smoke and locally emitted dust. For this site, we find similar results as above; the PAS-corrected data are accurate in smoke~~,~~ but are too low by a factor of 5-6 in dust. Using these data, we also find that the ratios of PAS-measured PM_{10} to PM_1 mass and 0.3 µm to 5 µm particle counts are significantly different for dust compared to smoke. Using this difference, ~~Given this ability of the PAS data to identify dust aerosols,~~ we propose a modified correction equation~~algorithm~~ that ~~significantly~~ improves the PAS data for some dust events, but further work is needed to improve this algorithm.

Introduction

Low-cost air sensors are becoming a ubiquitous way for the general public to measure local air quality. There are now thousands of these sensors publicly reporting data in real time to the PurpleAir map (map.purpleair.com~~https://purpleair.com/map~~). As one example, there are ~~now~~ more than 700 active PurpleAir sensors (PASs) in the Puget Sound region of Washington State (from Tacoma to Everett), compared to ~15 regulatory monitors in the same area. This provides an enormous increase in spatial information on $PM_{2.5}$ (particulate matter with diameters less than 2.5 µm). However, there are no clear performance standards for accuracy or precision of low-cost sensors. Several studies have examined the performance of low-cost sensors, including the PAS (Singer and Delp, 2018; Li et al., 2020; Ardon-Dryer et al., 2020; Manibusan and Mainelis, 2020; Tryner et al., 2020). The PAS uses the Plantower PMS5003 laser sensor ~~particle counter~~ to count particles that scatter light in the optical range (particles greater than about 0.2 µm in diameter). Most outdoor PASs include two identical PMS5003 sensors that can be compared to enhance quality control. The PAS data can be downloaded with two "conversion factors", CF=1 or CF=Atm. The two $PM_{2.5}$ values are nearly identical until 25 µg m^{-3}, but above this value the CF=1 will be greater. The exact algorithm used by the PAS to convert the Plantower data to mass concentration using either the CF=1 or CF=Atm factors has not been published (Ouimette et al., 2022).

Tryner et al. (2020) evaluated three low-cost ~~particulate matter~~PM sensors, including the PMS5003, by exposing them to five different types of aerosols in the laboratory. They found that the ratios of PMS5003-reported to filter-

derived $PM_{2.5}$ mass concentrations were inversely proportional to mass median diameter (MMD). Wood smoke had the smallest MMD, 0.42 µm; its PMS5003 $PM_{2.5}$ mass had a mean that was 2.5 times the filter-derived mass. Conversely, oil mist had the largest MMD at 2.9 µm; its PMS5003 $PM_{2.5}$ averaged only 0.23 times the filter-derived $PM_{2.5}$. These lab results are consistent with the physical-optical model developed for the PMS5003 by Ouimette et al. (2022). The model predicted that the PMS5003 response decreases relative to an ideal nephelometer by about 70–90 % for particle diameters ≥1.0 µm. This is a result of using a laser that is polarized, the angular truncation of the scattered light, and particle losses (e.g., due to aspiration) before reaching the laser. Their model predicted that the PMS5003 would underestimate $PM_{2.5}$ for dust particles by approximately 70–90 %, depending on the coarse particle size distribution.

The Plantower sensor reports PM mass concentrations in three bins (PM_1, $PM_{2.5}$, and PM_{10}) and particle counts in 6 size bins (>0.3, >0.5, >1, >2.5, >5, and >10 µm), presumably based on the pulse height of the scattered radiation, although the exact procedure is not documented by Purple Air. The PAS also reports temperature, relative humidity (RH) and pressure. A number of field and laboratory studies have found that the PMS5003 size distributions are not correct. Several studies have reported that the PMS5003 tends to create an invariant normalized size distribution, independent of the actual size distribution and concentration (Tryner et al., 2020; He et al., 2020; Kuula et al., 2020; Ouimette et al., 2022). However, the PMS5003 normalized size fractions above 1 µm increased by a factor of 2–5 in one high-$PM_{2.5}$ windblown dust episode observed at Keeler, California (Ouimette et al., 2022). So, at present, there remains some ambiguity over how the PAS reported $PM_{2.5}$ mass concentrations and particle counts respond to different aerosol types.

Aerosol size distributions can vary considerably depending on the source type. Previous studies have shown that the aerosol size distributions for smoke events are similar to the distributions in typical urban pollution events, with geometric mean diameter of around 0.2–0.3 µm (Kleeman et al., 1999; Laing et al., 2016). The mass ~~fraction~~ ratio of $PM_{2.5}/PM_{10}$ for smoke, 0.55–0.75, is also similar to that for urban pollution ~~smoke and urban pollution is also similar at 0.55–0.75~~ (Xu et al., 2017). Dust events are known to have size distributions that are shifted towards larger particles, compared to typical urban and smoke aerosols. Jiang et al. (2018) report an average $PM_{2.5}/PM_{10}$ ratio of 0.1 for dust events in China. Sugimoto et al. (2016), suggest a value of 0.35 for the $PM_{2.5}/PM_{10}$ ratio in dust, similar to the values reported by Tong et al. (2012). In addition, aerosol particles from some cooking methods, such as barbeque, may also have a size distribution that is shifted to larger sizes (Kleeman et al., 1999; Song et al., 2018). If this is correct, then this may have implications for using PAS data to examine indoor air quality.

The South Coast Air Quality Management District (South Coast AQMD) has completed a rigorous evaluation of a variety of sensors, including the PAS (South Coast AQMD, 2015). This evaluation has shown that the PAS gave precise results, showed little response to temperature or humidity, and had relatively small variations between units. The U.S. Environmental Protection Agency (EPA) also provides information about these sensors via its "Air Sensor Toolbox for citizen scientists, researchers, and developers" portal (U.S. EPA, 2022a). All of these evaluations have demonstrated that the raw PAS measurements are precise, but biased high compared to regulatory $PM_{2.5}$ measurements. Several groups have developed correction equations for the PAS measurements. The Lane

Regional Air Protection Agency (LRAPA), the University of Utah, and the EPA have empirical corrections for $PM_{2.5}$ and these can be implemented directly on the PurpleAir website (PurpleAir, 2022~~1~~). Barkjohn et al.~~,~~ (2021~~0~~) (hereafter referred to as Barkjohn 2021) conducted~~did~~ a comprehensive evaluation of PAS $PM_{2.5}$ data against regulatory $PM_{2.5}$ data and developed a U.S.-~~-~~wide correction equation starting from PAS raw data (CF=1) and using the RH as measured by the PAS:~~:~~

Corrected PAS $PM_{2.5}$ = raw PAS $PM_{2.5}$ data (CF=1) * 0.52 – RH * 0.085 + 5.71 (1)

~~where CF is a 'correction factor' and RH is the ambient relative humidity.~~ The LRAPA and the Barkjohn corrections are in close agreement, whereas the University of Utah correction gives somewhat higher values. While the Barkjohn 2021 algorithim (equation 1) was initially used by the EPA, ~~The EPA~~they have recently developed a new correction algorith~~i~~m ~~which~~that it is now being used ~~into~~ incorporated ~~PAS data using the Barkjohn correction into in the~~for the ~~its~~ national Fire and Smoke Map (Barkjohn et al., 2022).~~—~~ This algorith~~i~~m differs significantly from the earlier Barkjohn ~~et al~~ 2021 relationship in that it starts ~~from~~ from the PAS data with CF=Atm and involves a more complex, 5-part piece-wise regression, with weighting to smooth the transitions between segments.

~~(AirNow, 2021).~~ For our analysis, we will refer to the~~se as the Barkjohn 2021 and~~ new algorithm as~~the~~ "new EPA" ~~correction equations.~~ ~~as the "standard correction equation"~~.~~—~~ Note that the PAS data can be downloaded with either ~~either as raw data (~~CF=1 or CF=Atm~~). or with various calibration factors applied (e.g., LRAPA, EPA, etc).~~ ~~—~~ In the present analysis, we start from raw data with CF=1 (for Barkjohn 2021) or CF=Atm (for the new EPA correction ~~or Barkjohn et al 2022~~).~~—~~ Figure S1 compares the raw CF=Atm data with the new EPA correction algorithm, and Figure S2 compares the Bark~~s~~john 2021 and the new EPA correction for the data used in Part I of this analysis.~~.~~

Because many PAS devices are now installed around the world, both outside and inside~~, homes and workplaces,~~ they can experience a wide range of aerosol types. Thus, it is essential to understand the accuracy and precision of the PAS for various aerosol events, which could differ based on the particle size distribution or other aerosol characteristics. In this study, we evaluated the ~~standard correction equation (~~Barkjohn ~~et al., 2020~~2021 correction ~~)~~ and the new EPA correction for 50 different aerosol pollution events, encompassing typical urban aerosols, as well as smoke and dust aerosols. Our goals are:

1. Evaluate the accuracy of ~~the standard~~both correction equations for each aerosol type;
2. Examine whether the correction changes at very high $PM_{2.5}$ levels (e.g., >~~-~~250 µg m^{-3});
3. Identify whether the PAS data can provide an indication of the aerosol type and, if so, whether this information can be used to improve the correction algorithm.

Below we first describe data treatment and events and aerosol type identification. Then we report on results comparing regulatory and PAS observations for different aerosol types for 50 short-~~-~~term pollution events.~~—~~ We also use a longer time series from a single site (Keeler, CA) that experiences frequent high dust and smoke pollution episodes. Our results demonstrate that the PAS sensors can give accurate $PM_{2.5}$ data in urban pollution and smoke, but more work is needed to develop an improved correction for dust aerosols.

Methods

Part I: 50 paired sites

For this analysis we identified 50 short term pollution events where the aerosols could be clearly characterized as either typical urban, smoke or dust. For these events, PAS data were downloaded for each sensor from the PurpleAir website (map.purpleair.com). The raw data (CF=1 and CF=Atm) were downloaded as hourly averages. Regulatory PM data for the nearest monitoring site were downloaded from the EPA "Air Data" website (https://www.epa.gov/outdoor-air-quality-data) or the AirNow-Tech website (airnowtech.org), except for data from the monitoring site at Portland Cully Helensview School in Portland, OR (AQS Id 410512011), which were downloaded from the Oregon Department of Environmental Quality website (https://www.oregon.gov/deq/aq).

For each paired PAS-regulatory site, we identified an intense pollution event that had an hourly peak $PM_{2.5}$ value at the regulatory site with hourly $PM_{2.5}$ values >40 µg m^{-3} for at least 3 hours. We also required that there be a good correlation between the regulatory and PAS data. For the 50 events we analyzed, the correlation coefficients between the regulatory and PAS corrected data ranged from 0.77 to 0.996. For each pollution event, we identified the most likely source of elevated aerosols: typical urban wintertime pollution, biomass burning smoke, or dust. Table 1 summarizes the method used to characterize each pollution event. Table 2 gives a summary of the events and SI Table S1 provides show details on each of the 50 individual events, including PAS site, regulatory site ID, event dates, and distance between the two sites. The average distance between the PAS and regulatory sites was 5.4 km, with a range of 0-15 km. As shown in Figure S3, there is not a significant relationship between the correlation coefficients and distance between sites. Table S1 shows each set of paired sites and the start and end times for each event.

Typical urban pollution events were identified for the non-wildfire season (winter months) and with no evidence of smoke or dust. The PM sources for those events reflect typical urban, wintertime pollution (vehicles, power plants, industry, and residential wood combustion), and the $PM_{2.5}$ mass is dominated by particles with diameters in the range of 0.30–0.60 µm (Zhang et al., 2010; Herner et al., 2005). The typical urban pollution events had peak hourly $PM_{2.5}$ values at the regulatory sites of 47–259 µg m^{-3}.

Smoke events were identified by elevated $PM_{2.5}$ during the summer fire season and confirmed using the Hazard Mapping System (HMS) Fire and Smoke Product (Rolph et al., 2009; Kaulfus et al., 2017). The HMS product is derived from multiple satellite images and updated multiple times each day. Details on the HMS product are given in the references above. The HMS imagery was obtained via the AirNow-Tech website. The smoke events had the highest peak $PM_{2.5}$ values at the regulatory sites with peak hourly values of 60–713 µg m^{-3}.

Dust events were identified by examining large-scale spatial patterns of $PM_{2.5}$, media reports, and the measured $PM_{10}/PM_{2.5}$ ratios from regulatory sites, if available. In Part I of our analysis, all 6 dust events occurred during the well-known June 2020 Saharan dust cloud that was transported to the U.S. and impacted surface concentrations across the sSouthern U.S. (Francis et al., 2020; Euphrasie-ClotildeLovely et al., 2021; Pu and Jin, 2021). This event

brought huge amounts of dust to the southern U.S. and resulted in daily average $PM_{2.5}$ concentrations of 60–103 µg m^{-3} at many locations. The six dust events included in our analysis had peak hourly $PM_{2.5}$ values at the regulatory sites of 52–72 µg m^{-3}. Figure S43~~1~~ shows the impact of this dust on $PM_{2.5}$ across the southeastern U.S.

In total, we identified 50 events as either typical urban, smoke, or dust, lasting from 24 to 528 hours. We verified that each had an operating PAS and a nearby regulatory monitoring site. For typical urban pollution, 16 cases were identified, with the majority (13) ~~being~~ located in California~~,~~ and the remainder in Utah. We identified 28 smoke cases, with locations in Alaska, California, Idaho, Oregon, and Washington. Six dust cases were identified, with locations throughout the southeast U.S. Of the 50 events identified, 17 have co-located regulatory PM_{10} data (3 urban, 8 smoke, and 6 dust). The event times were chosen to incorporate a few hours of low concentrations before and after the highest $PM_{2.5}$ concentrations to improve correlations. The corrections on these low PAS values can sometimes yield negative values at high RH. If corrected PAS values were less than 2 µg m^{-3}, these values were excluded from the calculation of correlation with the regulatory measurements.

Data quality control

The data were quality controlled and screened using four~~ive~~ criteria:~~:~~

1. Since most PASs contain~~s~~ two sensors, A and B, we compared mass concentrations ~~data~~ from the two sensors and the data were ~~only~~ used only if the values were within 30 %. Most values are much closer than this, with an average difference of 10 % across all events considered (4.6 % for the Keeler, CA, PAS data).
2. The PAS raw A and B values were averaged and excluded if less than 1 µg m^{-3}.
3. The PAS values were corrected using the ~~standard~~ Barkjohn 2021 correction and ~~only~~ included only if greater than 2 µg m^{-3}. ~~Note that at high relative humidity, negative values can occur.~~
4. Regulatory $PM_{2.5}$ data must be greater than 1 µg m^{-3} (there were a number of 0 and negative values ~~are~~ in the EPA's $PM_{2.5}$ data records).

In total, these steps removed approximately 10 % of the available data. After screening, the PAS data were corrected using the Barkjohn et al (2021) algorithm and the new EPA

~~both the EPA~~ correction algorithms ~~(equation 1)~~. We evaluated ~~the~~ both sets of ~~raw and~~ corrected PAS data using the same linear relationship using standard linear regression:~~:~~

Regulatory data = Slope * PAS data (~~raw or~~ corrected) + Intercept (2)

We also compared the slopes with reduced major axis regression (RMA) and found essentially no difference in the results. Generally, the intercepts were small (a few µg m^{-3}), so we can interpret the slopes as giving the overall indication of agreement between the two ~~bias to the~~ datasets. A slope near 1 with a zero intercept would indicate no bias. A slope <1 indicates that the corrected PAS data ~~(raw or corrected)~~ are biased high compared to the regulatory data, a slope >1 indicates the corrected PAS data ~~results (raw or corrected)~~ are biased low compared to the regulatory data.

Part II–: -Keeler, CA, site.

To further understand the nature of the PAS response to dust aerosol, we also used data from Keeler, CA, near Owens Lake. Owens Lake is now a dry lake bedlakebed due to diversion of its primary water source, the Owens River, to Los Angeles. As a result, the dry lake bedlakebed is now one of the largest sources of dust in North America (Cahill et al., 1996; GilletteGillette et al., 1997), and the region experiences many significant dust events each year. With increasing drought, it appears that the dust flux from Owens Lake is increasing (Borlina and Rennó, 2017). We use regulatory $PM_{2.5}$ and PM_{10} data– from February 2019–May 2022 from a site in Keeler CA and a California Air Resources Board (CARB) site located in Keeler, CA, as well as a nearby PAS site–. Both the regulatory and PAS instruments are operated and maintained by the Great Basin Unified Air Pollution Control District (GBUAPCD, Chris Howard, personal communication, Dec. 2022) and the regulatory data were obtained from their data archive (https://www.gbuapcd.org/). The Rregulatory $PM_{2.5}$ and PM_{10} data arefor the Keeler site are from the Great Basin Unified Air Pollution Control District website (https://www.gbuapcd.org/) measured using a Thermo Fischer model 1400a TEOM with a Thermo Fischer model 8500C conditioning system.— Other information about the site is given in Table S2. . According to Based on the coordinates given in each data for the Keeler data, the PAS and regulatory sensors are within 30 meters of each other. The regulatory measurements are made with a Thermo Fischer model 1400a TEOM with a Thermo Fischer model 8500C conditioning system. For the Keeler PAS data, as in Part I, we use the mean of channels A and B, which have a mean difference of 4.6 %. More details on the Keeler site are given in Table S2. For the Keeler, CA analysis, we did not specifically identify events types. Instead, we consider only hours where the Keeler regulatory $PM_{2.5}$ >25 µg m^{-3}, which provides 1366 hours of data 3.3 year period. We also restrict the analysis of the Keeler data to hours where regulatory PM_{10} exceeds $PM_{2.5}$ by at least 0.5 µg m^{-3} and where simultaneous regulatory and PAS data are available. This yields 1257 hours of data with mean $PM_{2.5}$ and PM_{10} concentrations from the regulatory monitors of 59 and 118 µg m^{-3}, respectively.

Results

Part I: Event analysis

Figure 1 shows time series plots of two- example eventsFigure 1 (# 44 and # 45). The top plot in Figure 1a shows PAS and regulatory data during a major smoke event in Washington State during July–August 2021. The regulatory $PM_{2.5}$ exceeded 200 µg m^{-3} at this site. This figure demonstrates shows the well-known overestimation of $PM_{2.5}$ concentrations measured by the PAS during smoke events and also shows thatthat the Barkjohn 2021 standard EPA correction yields excellent bias correction of the data. The new EPA also improves the fit, compared to the raw data, but appears to yield a positive bias at the highest concentrations (200–250 µg m^{-3}). Regression was calculated twice for each event: (i) using ordinary linear regression (OLR) and (ii) using reduced major axis regression (RMA). In practice, there was rather little difference in the results, given that the correlation coefficients

were fairly high. Both are reported in Table S1. The bottom plot in For the smoke event shown in Figure 1 (event #44), the OLR slope is 0.81 and the RMA slope is 0.86. Figure 1b also shows data from a dust period in June 2020 (event # 45). In contrast to the smoke event, both correction algorithms are significantly under-predicting the regulatory values and there is essentially no difference between the two correction schemes. raw PA data are already well below the regulatory data and the standard correction only makes the bias greater. While there is still a good correlation between the regulatory and PAS data (R value of 0.82), the OLR slope is 6.766.43, indicating that boththe standard correction equations are significantly is under-estimating the true concentrations by a factor of 6 or more (note that this figure shows the PA data on the right axis). Table S3 shows the results for each of the 50 individual events. Table 2 summarizes the results for all 50 events. summarizes the relationship and correlation slopes between the corrected PAS data and the regulatory measurements for the 50 events and for the three different aerosol types. The results are consistent with Figure 1 in that urban pollution and smoke events are reasonably corrected by either the Barkjohn 2021 or new EPA algorithms, whereas dust events are not. There are some Table 2 also lists p-values, where p-values <0.05 indicate a statistical difference at 95% confidence.differences between the two correction equations, which we discuss below.

Table 3 and Table 4 summarizes the results by aerosol type and includes all hourly data for each identified aerosol type. Table 3 uses the Barkjohn 2021 correction, whereas Table 4 shows results using the new EPA correction. For urban, smoke, and dust aerosols, the slope of regulatory $PM_{2.5}$ versus the PAS corrected data with the Barkjohn 2021 algorithm were 1.02, 1.082, 1.08 and 4.988, respectively, using all hourly data of each type (Table 3). Using the new EPA correction, thesse slopes were 0.95, 0.81, and 4.99, respectively (Table 4). Thesise slopes indicates that both correction algorithms the standard equation yields excellent bias correction for typical urban and smoke events, but theyat it generates a large negative (low) bias for dust events. Using either the standard EPA correction on the PAS data during dust events gives values that are low by approximately a factor of 5 6.

Tables 2, 3 and 4 suggest that the new EPA algorithm has slightly lower slopes, especially for the smoke events. For example, the mean slope for smoke events (shown in Table 2) is 0.99 for the Barkjohn 2021 correction, vs 0.88 for the new EPA correction. Similarly, using hourly data for smoke influenced periods the slopes are 1.08 using the Barkjohn 2021 correction (Table 3) vs 0.81 using the new EPA correction (Table 4). We also want to examine whether there is evidence that the PAS data respond differently at very high PM concentrations. Figure 2 shows the mean bias using the hourly data with both correctioned algorithms versus slopes versus the regulatory peak $PM_{2.5}$. This plot includes only concentration for the data during the urban and, smoke and dust eevents. The bias is strongly negative using the Barkjohn 2021 correction at very high $PM_{2.5}$, greater than about 300 µg m^{-3}. At medium high $PM_{2.5}$ concentrations, such as 150 300 µg m^{-3}, the new EPA correction shows a positive bias, which is consistent with the results shown in Figure 1a and Tables 2-4. Thus we conclude that the new EPA correction improves the bias at very high concentrations (>300 µg m^{-3}), but introduces a modest bias at moderately high pollution levelsWhile the slopes for urban and smoke events are generally close to 1, there is a small, but statistically

~~significant trend in the slope with maximum $PM_{2.5}$ concentrations for the urban and smoke cases.~~ (150–300 µg m^{-3}), compared to the Barkjohn 2021 algorithm.

~~The results suggest that at the highest concentrations (600-700 µg m^{-3}), the standard PAS correction is reporting data that are 20-30% low compared to the regulatory.~~

We show above that the PAS data, ~~using the~~for ~~standard correction~~ both corrections, are~~is~~ substantially under-reporting $PM_{2.5}$ concentrations during dust events. The next question is whether the PAS data can give some information about dust events (i.e., the presence/absence of dust), despite significant issues with the reported size distribution (Ouimette et al., 2022). To address this question, we calculated the slopes of the PM_1 to PM_{10} ~~-~~mass ratios and the 0.3 µm to 5 µm particle counts ratio, ~~both~~ using the PAS data for each event.~~—~~ The results are reported by event type in Table 2. The results show that the PAS reports a greater fraction of coarse mass and proportionally more larger particles, compared to the 0.3 µm particles, in the dust aerosols, compared to urban or smoke aerosols. Both the PM_1/PM_{10} mass ratio and the 0.3 µm to 5 µm particle counts ratio increases in the order dust < smoke < urban. These differences are statistically significant ($P<0.05$) for urban versus dust~~,~~ using a two sample, two tailed t-test and assuming unequal variance. ~~but not for smoke versus dust at the 95 % confidence level.~~ ~~T~~These relationships will be explored further below in Part II of this analysis.

We ~~have~~ also looked at the coarse aerosol fraction (CAF) for these events using both ~~-~~the regulatory and PAS data. We define the CAF as:

$$\mathbf{CAF = (PM_{10} - PM_{2.5}) / PM_{10}} \quad \mathbf{(3)}$$

Out of the 50 events considered, 17 have both regulatory $PM_{2.5}$ and PM_{10}. Figure 3 shows the CAF, calculated using both the regulatory data and the PAS raw data for all hours for the 17 events with both $PM_{2.5}$ and PM_{10} data. For the PAS data, we use the raw values for ~~both~~ $PM_{2.5}$ and PM_{10}, since there are no known correction algorithms for the PM_{10} data. Several things are apparent in Figure 3. First, the CAF values using the regulatory data are much higher than CAF values obtained from the PAS data. Nonetheless, both the regulatory and PAS data show the expected pattern of higher CAF in dust compared to the other aerosol types.~~—~~ In addition, the number of data points is much higher for the PAS, due to the relative sparsity of regulatory PM_{10} data.~~—~~ We note that these relationships change very little if the PAS data are restricted to the same times as the regulatory data.

Part II~~—~~: ~~-~~Keeler, CA, analysis

~~Here~~ In Part II we use the multi-year dataset from the Owens Lake/Keeler, CA, site.~~—~~ The hourly data cover a period of a little more than 3 years (February~~.~~ 2019~~–~~May 2022).~~—~~ We focus exclusively on hours with regulatory $PM_{2.5}$ >~~-~~2.5 µg m^{-3}, which yields 1257 hours, after our quality control described above. Table S2 has more details on both the regulatory and PAS sites in Keeler, CA.

Figure 4 shows a histogram of the CAF based on the regulatory data. There is a clear bimodal distribution, indicating two very different aerosol types during these pollution events.~~—~~ Given that Keeler is ca 150 km from the

urban areas of Bakersfield and Fresno, CA, this aerosol is likely either dust generated from Owens Lake, or smoke from the many California wildfires during 2019–2022. For the points with CAF<0.5 (n=1013 hours), the vast majority (99 %) occurred in August–October 2020 or August–September 2021, both times when large fires were burning in central CA California and influencing air quality across the region. Thus, it is reasonable to conclude that those hours with CAF<0.5 are predominantly wildfire smoke (1013 hours), and those with CAF>0.7 (n=xxxx200 hours) are predominantly dust. In contrast to the smoke data, the dust data tend to occur in the winter and spring periods. There are a relatively smaller number of points (n=xxxx44 hours) with 0.5 7< >CAF < >0.75 and these appear to have a mixed character, as shown below.

Table 54 and Figures 5 and 6 5 show results grouped by the CAF calculated using the regulatory data. Tables shows that fFor all values of CAF below 0.5, there areis a similar ratios of PM_1/PM_{10}behavior. and 0.3 µm/0.5 µm counts. For this group, the PAS $PM_{2.5}$ with the Barkjohn 2021 with standard correction shows a good fit to the regulatory $PM_{2.5}$. There is also a fairly consistent set of PAS measured PM_1/PM_{10} and 0.3 µm to 5 µm count ratios. For the values with CAF>0.7, there is similar consistency in the PAS–measured ratios (PM_1/PM_{10} –and 0.3 µm/5 µm counts), but for this group the PAS Barkjohn 2021standard correction significantly under-estimates the regulatory concentrations. For the group with CAF between <0.5 and CAF>0.7, the primary aerosol type has a mixed character, likely including boths are smoke and dust., respectively. The middle group (0.7>5<CAF>0.75) appears to have a more mixed character.

Figure 5 shows a plot of the regulatory $PM_{2.5}$ versus PAS $PM_{2.5}$ with the Barkjohn 2021standard correction, sorted by these three groups (CAF<0.5, 0.57<>CAF<>0.7,5 and CAF>0.7). For the smoke aerosols, the PAS with standard the Barkjohn 2021 correction shows a slope of 0.99 and an R^2 of 0.92, whereas for the dust aerosols, the slope is 5.6, similar to the slopes shown in Table 2 (5.5) and Table 3 (5.0). Thus, we conclude that for dust aerosols the PAS standardBarkjohn corrected PAS values show a 5–6x– underestimate of the $PM_{2.5}$ regulatory values. The mixed aerosols show behavior that is more difficult to characterize, with some showing more similarity to dust and others to smoke.

Figure S5 and S6, show the ratios of PM_1/PM_{10} and the 0.3 µm/5 µm counts and the ratios of PM_1/PM_{10}, as measured by the PAS, versus the CAF, and Table 54 also shows a summary of the data segregated by CAF. Both ratios of PM_1/PM_{10} and the 0.3 µm/5 µm counts, show clear differentiation for the low CAF aerosols compared to the high CAF aerosols. So these that the ratios of PM_1/PM_{10} and the 0.3 µm/5µm counts, as measured by the PAS, vary by CAF. Theseis unitless ratios provides a tool that can identify dust aerosols, so that and apply a separate correction can be applied.. We explored both using both the ratio of the PM_1 to /PM_{10} –mass concentrations and the ratio of 0.3 µm to /5µm counts– as a tools to identify $PM_{2.5}$ aerosol that is dominated by dust.ratios and found better separation using the count ratios. Figure S5 and Table 54 shows that using a ratio of the 0.3 µm to /5µm counts– ratio of somewhere between 150–250 will provide the best separation of dust and mixed aerosols. By examination of various plots of regulatory $PM_{2.5}$ versus corrected PAS $PM_{2.5}$ for the Keeler, CA, data, we found an optimum value of 190. The value of 5.6 comes from the slope of the dust aerosols in Figure 5. So, this leads to a new correction equation that depends on PAS–measured values:

If PAS 0.3 µm / 5 µm > 190, use Barkjohn 2021 ~~standard~~ correction~~;~~

If PAS 0.3 µm / 5 µm < 190, use Barkjohn~~standard~~ 2021 correction * 5.6 (4)

In ~~equation~~Eq. (4), we use the Barkjohn 2021 correction, but in practice there is little difference in the results regardless of whether this or the new EPA correction is used. Figure 6 shows a plot of the Keeler, CA, regulatory $PM_{2.5}$ versus PAS $PM_{2.5}$ with ~~the new correction a~~equation ~~4~~ Eq. (4) applied. There is very little change to the smoke data as most of these points have PAS measured 0.3 µm/5 µm counts >190. For the dust aerosols, the majority of the data points are now much closer to the regulatory values. The mean bias for the points with CAF >0.7 is now 1.3 µg m^{-3} compared with 51.4 µg m^{-3} for the dust data using the ~~standard~~ Barkjohn 2021 correction. Figures S7~~6~~4 and S8~~7~~5 show how the choice of 0.3 µm/5 µm ratio impacts the analysis. Using a higher threshold in ~~equation~~ Eq. (4) results in identifying some points (smoke) with corrected $PM_{2.5}$ values that are substantially too high. Using a lower threshold in ~~equation~~ Eq. (4) results in missing some dust points and, for those points, generating PAS corrected $PM_{2.5}$ values that are too low. While using a value of 190 in Eq. (4)~~equation 4~~ does miss a small number of dust points, it appears to be the best balance in finding and correcting the dust data points for this location. Finally, Figure S9 shows regulatory $PM_{2.5}$ versus PAS $PM_{2.5}$ with the new EPA correction separated by CAF. The results are nearly identical to Figure 5, showing that both the Barkjohn 2021 and new EPA correction algorithms have similar behavior with dust aerosols.

~~Eq. (4)~~

Equation (4) was developed based on data from one site (Keeler) that has strong dust and smoke occurrence and with the sensors in close proximity (30m) ~~(Keeler)~~. We apply Eq. (4) ~~equation 4~~ to the 50 ~~dust~~ events from ~~at~~ different sites identified in Part I and find a wider range of~~in~~ results. ~~T~~Table S3~~4~~ summarizes the results for each event. Out of the 6 dust events, 4 show moderate improvements with slopes of 0.46 0.72. However, ~~f~~for some smoke events (e.g., 38, 39, and 40), the slopes are dramatically lower, in the range of 0.17 0.26, which indicates that the PAS corrected with the dust algorithm (equation 4) are overestimating the regulatory data by a large amount. This occurs due to the fact that during these smoke events some hours have a ratio of the 0.3 µm to 5 µm counts of >190 and thus get multiplied by 5.6. ~~S~~So, while the new dust algorithm does appear to improve PAS-corrected data in dust events at a single controlled site that is operated by an air quality agency, it does not provide a useful correction for the bulk of publicly operated sensors. ~~N~~Nonetheless, the fact that the PAS data indicate changes in the observed ratios of PM_1/PM_{10} ~~a~~and the 0.3 ~~µ~~µm/ 5 µm counts during mineral dust events indicates that the PAS data do provide some useful information on dust and that more work to identify a suitable correction algorithm for dust is warranted.

~~a significant improvement in the mean bias (6.1 µg m^{-3} using equation 4, in contrast to 24.2 µg m^{-3} using the standard correction). However the requirement that the A and B channel 0.3 and 5 5 µm counts agree within 30% reduces the number of data points by more than half. This largely reflects the low aerosol numbers in the 5 µm channel and therefore relatively large variability, due to the factors identified in Ouimette et al. (2022). It is possible that there is a better dust correction algorithm that could integrate more of the PAS measured parameters. Given~~

~~that the PAS substantially underestimates $PM_{2.5}$ concentrations during dust events using the standard correction, we propose that further evaluation of this new algorithm is warranted.~~

Conclusion

PASs are now ubiquitous around the world and far outnumber the more accurate, regulatory-grade instruments for $PM_{2.5}$. These low-cost sensor data are proving to be highly valuable for a variety of analyses, but especially for improving our understanding of the spatial distribution of $PM_{2.5}$. However, to use these data, it is essential to understand the measurements. Using the Barkjohn 2021 and new EPA~~'s~~ correction algorithm for PAS data, we find that the sensors give reasonably accurate results for $PM_{2.5}$ for typical urban wintertime pollution and smoke events, but give concentrations that are a factor 5–6 too low for dust events. The Barkjohn 2021 algorithm yields a negative bias at very high $PM_{2.5}$ concentrations (>300 µg m^{-3}), whereas the new EPA algorithm yields a positive bias at moderate $PM_{2.5}$ concentrations (150–300 µg m^{-3}). Both algorithms underestimate $PM_{2.5}$ during dust events by a factor of 5–6. Using the PAS ratios of PM_{10} to PM_1 mass concentrations and 0.3 µm to 5 µm counts, we find that there are significant differences in these ratios for smoke and dust at a site with frequent incursions of both aerosol types. Using this result, we propose a new PAS correction algorithm that significantly improves the correction for dust aerosols and does not change the results for smoke aerosols, but only at this one site. Applying this equation to a broader array of sites ~~(Part I)~~, we find significant problems with the proposed dust algorithm—it improves PAS $PM_{2.5}$ estimates in some dust cases~~,~~ but worsens PAS $PM_{2.5}$ estimates for some smoke events ~~(equation 4)~~. Nonetheless, our analysis ~~. While this new equation needs further evaluation, it~~ demonstrates that it may be possible to develop an improved PAS correction algorithm ~~could be developed~~ that ~~which w~~could identify dust and provide a better estimate of the $PM_{2.5}$ concentrations when dust is present.

Data availability. All data used in this analysis are publicly available. Most regulatory data were obtained from the EPA Air Data site (https://www.epa.gov/outdoor-air-quality-data) and the AirNow Tech site (https://www.airnowtech.org/). Data for the Keeler, CA, site were from the Great Basin Unified Air Pollution Control District (https://www.gbuapcd.org/). Data for the Cully Helensview School in Portland, OR, were~~as~~ downloaded from the Oregon Department of Environmental Quality website (https://www.oregon.gov/deq/aq). PurpleAir data were from the PurpleAir site (http://map.purpleair.com)

Author contributions. DJ designed the study, developed the analysis protocols, and wrote the manuscript. CM, KT, and ~~N~~MN conducted data analysis. B~~.~~F, J~~.~~O~~.~~, and E~~.~~A reviewed the manuscript and provided comments on the analysis.

Competing interests. The authors have no competing interests to declare.

Acknowledgements

We wish to acknowledge and thank the many individuals that have made their PAS data freely available for scientific analysis. Partial support for this work came from the UW Bothell SRCP Seed Grant Program. MN was

supported by an internship from the Confederated Tribes of the Colville Reservation, which was funded by an EPA Environmental Justice grant. EA was supported by the NOAA Cooperative Agreement with CIRES, NA17OAR4320101.

References

~~**AirNow. Fire and Smoke Map (https://fire.airnow.gov/).** Accessed Dec. 13th, 2022.~~

Ardon-Dryer, K., Dryer, Y., Williams, J. N., and Moghimi, N.: Measurements of $PM_{2.5}$ with PurpleAir under atmospheric conditions, Atmos. Meas. Tech., 13, 5441–5458, https://doi.org/10.5194/amt-13-5441-2020, 2020.

Barkjohn, K. K., Gantt, B., and Clements, A. L.: Development and application of a United States-wide correction for PM2.5 data collected with the PurpleAir sensor, Atmos. Meas. Tech., 14, 4617–4637, https://doi.org/10.5194/amt-14-4617-2021, 2021.

Barkjohn, K., Holder, A., Clements, C., ~~and S.~~ Frederick, S., and Evans, R.: Sensor data cleaning and correction: Application on the AirNow Fire and Smoke Map, https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=353088&Lab=CEMM, ~~accessed Dec.~~last access: 13th December ~~,~~2022.

Borlina, C. S., and Rennó, N. O.: The impact of a severe drought on dust lifting in California's Owens Lake area, Sci. Rep., 7, 1784, https://doi.org/10.1038/s41598-017-01829-7, 2017.

Cahill, T. A., Gill, T. E., Reid, J. S., Gearhart, E. A., and Gillette, D. A.: Saltating particles, playa crusts and dust aerosols at Owens (Dry) Lake, California, Earth ~~Surface~~ Surf. Process. ~~es and~~ Landforms, ~~v.~~ 21, 7, p. 621–639, https://doi.org/10.1002/(SICI)1096-9837(199607)21:7<621::AID-ESP661>3.0.CO;2-E, ~~1~~1996.

Euphrasie-Clotilde, L., Plocoste, T., and Brute, F.-N.: Particle size analysis of African dust haze over the last 20 years: a focus on the extreme event of June 2020, Atmos., ~~phere~~ 12, 502, https://doi.org/10.3390/atmos12040502, 2021.

Francis, D., Fonseca, R., Nelli, N., Cuesta, J., Weston, M., Evan, A., and~~&~~ Temimi, M.: The atmospheric drivers of the major Saharan dust storm in June 2020. ~~Geophysical~~ Geophys. ~~Research~~ Res. ~~Letters~~Lett., 47, e2020GL090102, https://doi.org/10.1029/2020GL090102, 2020.

~~Gillette, D., D. Fryrear, J. B. Xiao, P. Stockton, D. Ono, P. Helm, T. Gill, T. Ley, in press a, Large-scale variability of wind erosion mass flux rates at Owens Lake. Part I. Vertical profiles of horizontal fluxes of wind eroded particles (d>50 micrometers), J.Geophys.Res. 102, 25977-25987, https://doi.org/10.1029/97JD00961, 1997.~~

Gillette, D. A., Fryrear, D. W., Xiao, J. B., Stockton, P., Ono, D., Helm, P. J., Gill, T. E., and Ley, T.: Large-scale variability of wind erosion mass flux rates at Owens Lake: 1. Vertical profiles of horizontal mass fluxes of wind-eroded particles with diameter greater than 50 µm, J. Geophys. Res., 102, ~~(~~ D22~~)~~, 25977–25987, https://doi.org/~~doi:~~10.1029/97JD00961, 1997.

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

Herner, J., Aw, J., Gao, O., Chang, D. P., and Kleeman, M.: Size and composition distribution of airborne particulate matter in Northern California: 1-particulate mass, carbon, and water-soluble ions., Air & Waste Manage. Assoc., 55, 30–51, https://doi.org/10.1080/10473289.2005.10464600, 2005.

Jiang, N., Dong, Z., Xu, Y., Yu, F., Yin, S., Zhang, R., and Tang., X.: Characterization of PM_{10} and $PM_{2.5}$ source profiles of fugitive dust in Zhengzhou, China., Aerosol Air Qual. Res., 18, 314–329, https://doi.org/10.4209/aaqr.2017.04.0132, 2018.

Kaulfus, A. S., Nair, U., Jaffe, D. A., Christopher, S. A., and Goodrick, S.: Biomass burning smoke climatology of the United States: Implications for particulate matter air quality, Environmental Environ. Science &Sci. Technol.,ogy 50, 11731–11741, https://doi.org/doi: 10.1021/acs.est.7b03292, 2017.

Kleeman, M. J., Schauer, J. J., and Cass, G. R.: Size and composition distribution of fine particulate matter emitted from wood burning, meat charbroiling, and cigarettes., Environ. Sci. Technol., Environmental Science & Technology 33, 3516–3523, https://doi.org/DOI: 10.1021/es981277q, 1999.

Kuula, J., Mäkelä, T., Aurela, M., Teinilä, K., Varjonen, S., González, Ó., and Timonen, H.: Laboratory evaluation of particle-size selectivity of optical low-cost particulate matter sensors, Atmos. Meas. Tech., 13, 2413–2423, https://doi.org/10.5194/amt-13-2413-2020, 2020.

Li, J., Mattewal, S. K., Patel, S., Biswas, P.: Evaluation of Nine Low-cost-sensor-based Particulate Matter Monitors. *Aerosol Air Qual. Res.*, 20, 254–270, https://doi.org/10.4209/aaqr.2018.12.0485, 2020.

Laing, J., Jaffe, D., and Hee, J.: Physical and optical properties of aged biomass burning aerosol from wildfires in Siberia and the wWestern USA at the Mt. Bachelor Observatory., Atmos. Chem. Phys., 16, 15185–15197, https://doi.org/10.5194/acp-16-15185-2016, 2016.

Li, J., Mattewal, S. K., Patel, S., and Biswas, P.: Evaluation of nine low-cost-sensor-based particulate matter monitors. Aerosol Air Qual. Res., 20, : 254–270, https://doi.org/10.4209/aaqr.2018.12.0485, 2020.

Euphrasie-Clotilde, L.; Plocoste, T.; Brute, F. N.: Particle Size Analysis of African Dust Haze over the Last 20 Years: A Focus on the Extreme Event of June 2020. Atmosphere 12, 502. https://doi.org/10.3390/atmos12040502, 2021.

Manibusan, S., and Mainelis, G.: Performance of four consumer-grade air pollution measurement devices in different residences. Aerosol Air Qual. Res., 20, 217–230, https://doi.org/10.4209/aaqr.2019.01.0045, 2020.

Ouimette, J. R., Malm, W. C., Schichtel, B. A., Sheridan, P. J., Andrews, E., Ogren, J. A., and Arnott, W. P.: Evaluating the PurpleAir monitor as an aerosol light scattering instrument, Atmos. Meas. Tech., 15, 655–676, https://doi.org/10.5194/amt-15-655-2022, 2022.

Pu, B., and Jin, Q.: A record-breaking trans-Atlantic African dust plume associated with atmospheric circulation extremes in June 2020, B. ulletin of the Am.erican Meteorol.ogical Soc.iety, 102, 7, E1340–E1356Apr 2021., https://doi.org/10.1175/BAMS-D-21-0014.1, 2021.

PurpleAir: .PurpleAir Map, https://map.purpleair.com/, last access 18 December 2022.

Rolph G.D., *et al.*: Description and verification of the NOAA Smoke Forecasting System: the 2007 fire season. *Weather Forecast.* 24, 361–378, https://doi.org/10.1175/2008WAF2222165.1, 2009.Rolph, G. D., Draxler, R. R., Stein, A. F., Taylor, A., Ruminski, M. G., Kondragunta, S., Zeng, J., Huang, H. C., Manikin, G., McQueen, J. T., and Davidson, P. M.: Description and verification of the NOAA Smoke Forecasting System: the 2007 fire season, Weather Forecast., 24, 361–378, https://doi.org/10.1175/2008waf2222165.1, 2009.

Singer, B. C., and Delp, W. W.: Response of consumer and research grade indoor air quality monitors to residential sources of fine particles,. Indoor Air, 28, (4), 629–639, https://doi.org/10.1111/ina.12463, 2018.

Song, Y., Sun, L., Wang, X. et al. Pollution characteristics of particulate matters emitted from outdoor barbecue cooking in urban Jinan in eastern China. Front. Environ. Sci. Eng. 12, 14, https://doi.org/10.1007/s11783-018-1024-0, 2018.Song, Y., Sun, L., Wang, X., Zhang, Y., Wang, H., Li, R., Xue, L., Chen, J., and Wang, W.: Pollution characteristics of particulate matters emitted from outdoor barbecue cooking in urban Jinan in eastern China, Front. Environ. Sci. Eng., 12, 14, https://doi.org/10.1007/s11783-018-1024-0, 2018.

South Coast Air Quality Management District (South Coast AQMD). (2015).: AQ-SPEC Air Quality Sensor Performance Evaluation Center,. https://www.aqmd.gov/docs/default-source/aq-spec/summary/purpleair-pa-ii---summary-report.pdf?sfvrsn-16, 2015, last access: 12 December 2022..

Sugimoto, N., Shimizu, A., Matsui, I., and Nishikawa, M.: A method for estimating the fraction of mineral dust in particulate matter using PM2:5-to-PM10 ratios, Particuology, 28, 114–120, https://doi.org/10.1016/j.partic.2015.09.005, 2016.

Tong, D. Q., Dan, M., Wang, T., and Lee, P.: Long-term dust climatology in the western United States reconstructed from routine aerosol ground monitoring, Atmos. Chem. Phys., 12, 5189–5205, https://doi.org/10.5194/acp-12-5189-2012, 2012.

Tryner, J., Mehaffy, J., Miller-Lionberg, D., and Volckens, J.: Effects of aerosol type and simulated aging on performance of low-cost PM sensors, J. Aerosol Sci., 150, 105654, https://doi.org/10.1016/j.jaerosci.2020.105654, 2020.

U.S. Environmental Protection Agency (U.S. EPA 2022):. Air Sensor Toolbox. https://www.epa.gov/air-sensor-toolbox/, accessed last access: Sept. 14th September, 2022a.

U.S. Environmental Protection Agency (U.S. EPA 2022):. AirNow Fire and Smoke Map, https://fire.airnow.gov/, accessedlast access: December 18th December, 2022b.

Xu, G., Jiao, L., Zhang, B., Zhao, S., Yuan, M., Gu, Y., Liu, J., and Tang, X.: Spatial and temporal variability of the $PM_{2.5}/PM_{10}$ ratio in Wuhan, cCentral China,. Aerosol Air Qual. Res., 17,: 741––751,. https://doi.org/10.4209/aaqr.2016.09.0406, 2017.

Zhang, Y., Liu, P., Liu, X., Pun, B., Seigneur, C., Jacobson, M., and Wang, W.: Fine scale modeling of wintertime aerosol mass, number, and size distributions in central California. J. Geophys. Res.-Atmos., 115, https://doi.org/10.1029/2009JD012950, 2010.

535

540

Tables/Figures

Table 1. Methodology for identification of pollution events for 50 cases in Part I.

Event	Method of Identification	$PM_{2.5}/PM_{10}$ (if available)
Typical urban	One-hour regulatory $PM_{2.5}$ measurements exceeded 47 µg m^{-3} during non-wildfire season with no known presence of smoke or dust.	>0.5
Smoke	One-hour regulatory $PM_{2.5}$ measurements exceeded 47 µg m^{-3} in the presence of smoke as indicated on the NOAA Hazard Mapping System-Fire and Smoke Product.	>0.5
Dust	One-hour regulatory $PM_{2.5}$ measurements exceeded 47 µg m^{-3} during a known dust event.	<0.5

NEW Table 2. Peak regulatory $PM_{2.5}$, mean slope and R^2 results from analysis of regulatory and PAS data, with Barkjohn 2021 correction and new EPA correction, for 50 individual pollution events (Part I dataset). N gives number of events of each type, SD is standard deviation. R^2 is the mean value for all events of that type. Also shown are the average slopes by aerosol type for the PM_1 versus PM and 0.3 0.3 µm versus 5 µm counts regressions, correlations both of which are. All slopes are unitless.

	Average peak regulatory $PM_{2.5}$ (µg m^{-3})	Average slope (R^2) using Barkjohn 2021 correction	Average slope (R^2) using new EPA correction	Average slope of raw PAS PM_1 versus PM_{10} mass concentrations	Average slope of raw PAS 0.3 µm versus 5 µm counts
Urban-avg (N=16)	85.15	1.00 (0.88)	0.95 (0.88)	0.56	72732.6
SDd	56.69	0.11	0.15	0.10	42640.8
Smoke-avg (N=28)	280.32	0.99 (0.93)	0.88 (0.92)	0.445	402376.7
SDd	226.28	0.18	0.13	0.104	2658.8
Dust-avg (N=6)	59.76	5.54 (0.85)	5.53 (0.85)	0.29	1332.6
SDd	7.91	1.13	1.10	0.08	76.6

NEW Table 3. Relationship between hourly regulatory $PM_{2.5}$ and corrected PAS $PM_{2.5}$ with Barkjohn 2021 ~~standard~~ algorithm. Data are included for all simultaneous measurements for the 50 identified events in Part I. (N gives number of hours of data of each type.)~~.~~ .

	Mean r~~R~~egulatory $PM_{2.5}$ ($\mu g\ m^{-3}$)	Mean corrected PAS $PM_{2.5}$ ($\mu g\ m^{-3}$)	Slope for r~~R~~egulatory versus PAS w/Barkjohn 2021 correction (R^2)	Intercept ($\mu g\ m^{-3}$)	RMSE* ($\mu g\ m^{-3}$)	Mean bias ($\mu g\ m^{-3}$)
Urban (N=966)	33.9	28.7	1.02 (0.793)	4.60	10.9	-5.2
Smoke (N=6536)	66.4	66.0	1.08 (0.866)	-4.68	36.0	-0.4
Dust (N=240)	30.5	6.4	4.98 (0.661)	-1.09	27.9	-24.1

*Root mean squared error

NEW Table 4. Relationship between hourly regulatory $PM_{2.5}$ and corrected PAS $PM_{2.5}$ with new EPA ~~standard~~ algorithm. Data are included for all simultaneous measurements for the 50 identified events in Part I. (N gives number of hours of data of each type.)~~.~~

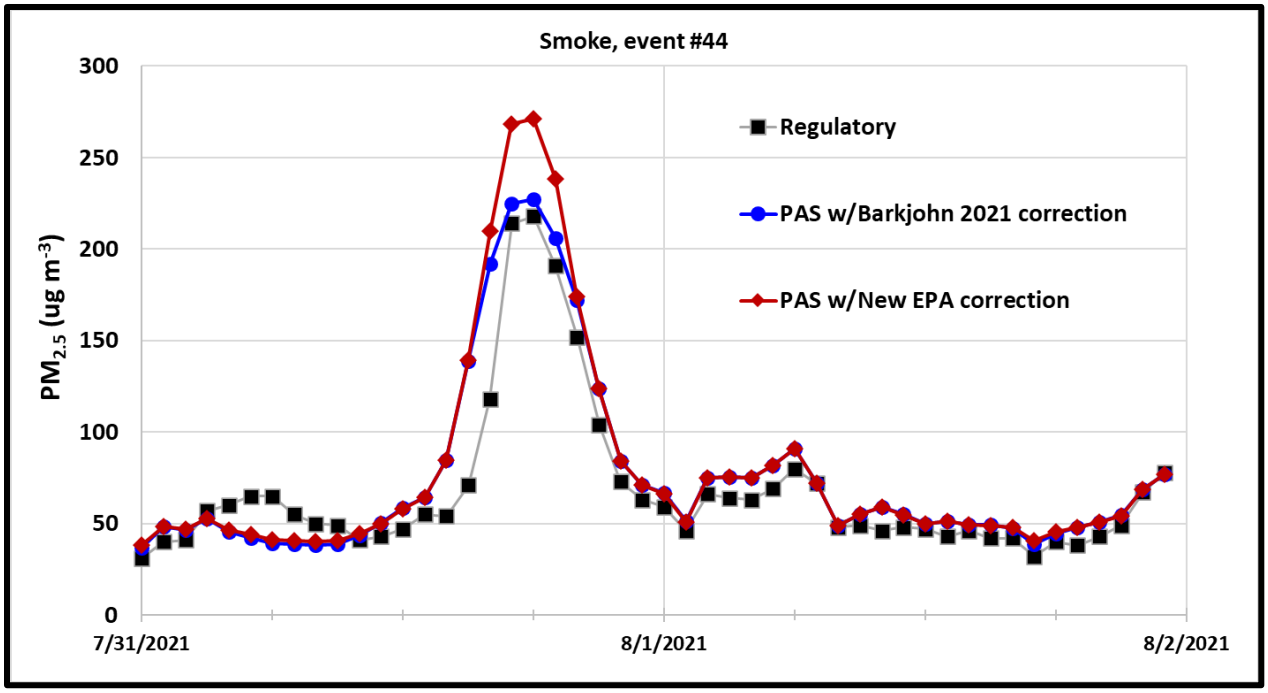
	Mean r~~R~~egulatory $PM_{2.5}$ ($\mu g\ m^{-3}$)	Mean corrected PAS $PM_{2.5}$ ~~w/corr~~ ($\mu g\ m^{-3}$)	Slope for r~~R~~egulatory versus PAS w/new EPA correction (R^2)	Intercept ($\mu g\ m^{-3}$)	RMSE* ($\mu g\ m^{-3}$)	Mean bias ($\mu g\ m^{-3}$)
Urban (N=966)	33.9	30.3	0.950 (0.744)	4.90	11.1	-3.6
Smoke (N=6536)	66.4	77.3	0.807 (0.858)	5.56	43.2	11.0
Dust (N=240)	30.5	6.4	4.99 (0.664)	-1.22	27.9	-24.1

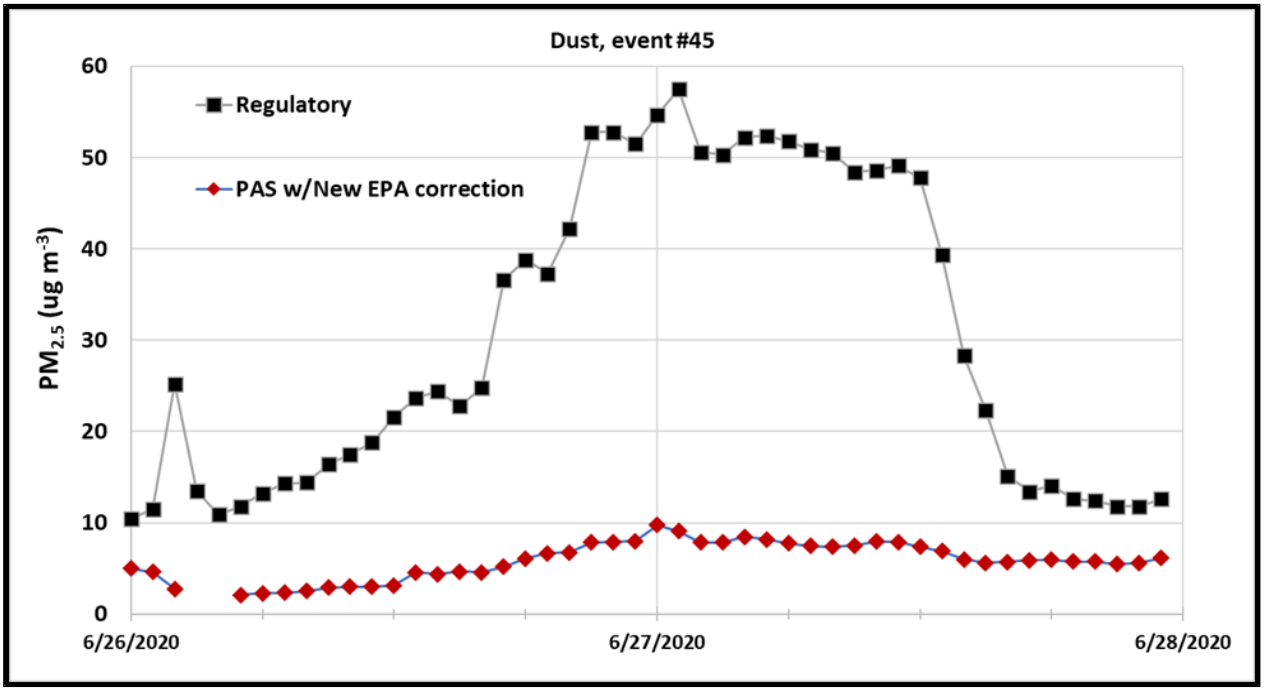
*Root mean squared error

Table 54.— Mean regulatory (reg) $PM_{2.5}$, PAS $PM_{2.5}$ (with Barkjohn 2021 correction and with proposed dust correction~~s~~), ratio of PAS PM_1/PM_{10} concentration, and ratio of PAS 0.3 to 5 µm counts by coarse aerosol fraction (CAF)~~CAF~~ bins. –The CAF bins are centered on the indicated value.

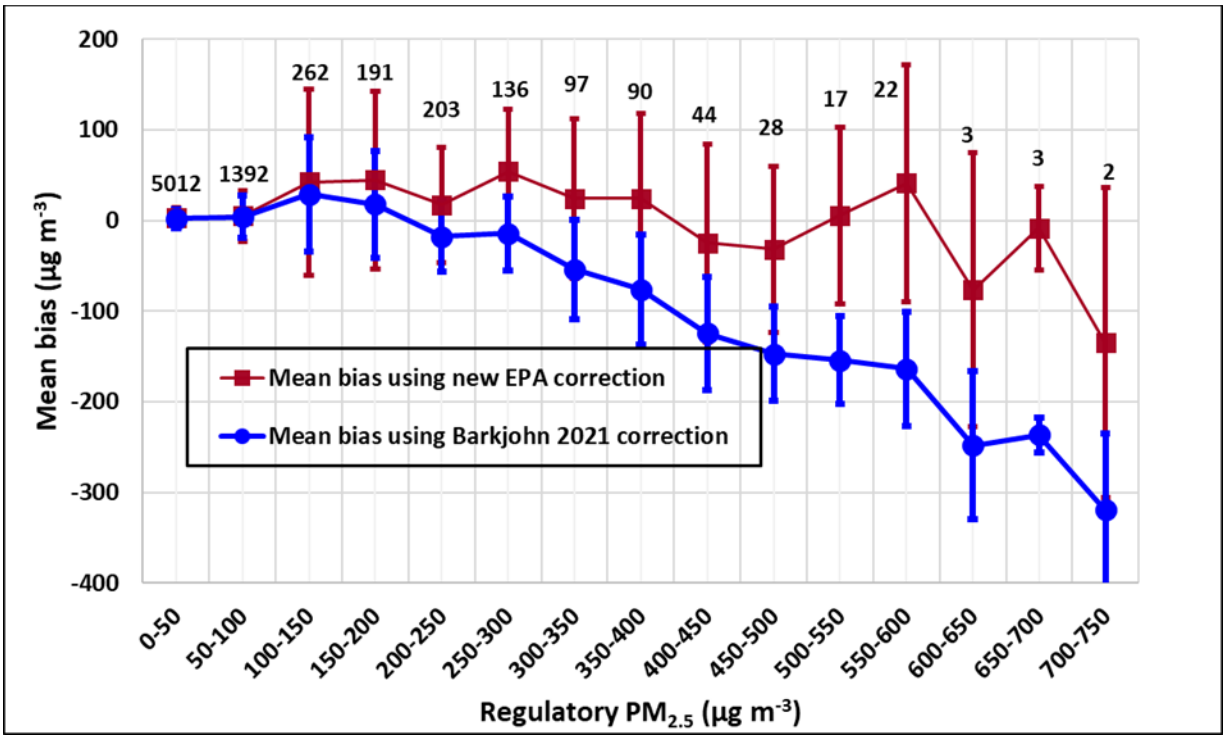
CAF bin midpoint	N (hrs)	Regulatory $PM_{2.5}$ ($\mu g\ m^{-3}$)	PAS $PM_{2.5}$ w/Barkjohn 2021 correction ($\mu g\ m^{-3}$)	PAS $PM_{2.5}$ w/dust correction ($\mu g\ m^{-3}$)	Mean ratio of PAS PM_1/PM_{10}	Mean ratio of PAS 0.3 to 5 µm counts
0.05	260	89.5	91.4	91.4	0.55	730
0.15	334	59.4	61.5	61.5	0.55	697
0.25	231	41.6	43.4	43.4	0.56	723
0.35	131	37.6	38.5	38.5	0.54	623
0.45	57	36.9	37.3	37.3	0.54	611
0.55	14	40.6	25.1	33.0	0.44	474
0.65	30	52.5	16.0	45.7	0.33	249
0.75	104	68.4	13.5	63.8	0.25	151
0.85	86	59.3	11.2	60.7	0.20	105

0.95	10	57.2	12.4	66.1	0.21	111





NEW Figure 1: Time series of hourly regulatory and PAS data $PM_{2.5}$ (raw and corrected) for two events, # 44 (smoke, top) and # 45 (dust, bottom). Time is in UTC. Note that for the dust event (# 45), the two correction schemes give identical results. Details on the sites used for these figures are~~is~~ given in Table~~s~~ S1 and S3. For event 44, the slopes (using Eq.~~equation~~ (2)) comparing ~~using~~ the Barkjohn 2021 and new EPA corrections schemes are 0.81 and 0.70, respectively.~~—~~ For event 45, the slope~~s~~ using the Barkjohn 2021 ~~and new EPA~~ corrections schemes ~~is~~are 6.76~~— and 6.70, respectively~~.



NEW Figure 2: Comparison of mean bias (corrected PAS-regulatory) using the hourly data for smoke and urban pollution events in Ppart I using the Barkjohn 2021 and new EPA correction schemes. Data are binned by regulatory $PM_{2.5}$ in 50 µg m^{-3} bins, as shown on the X axis. The values above the red points are the number of hourly datapoints in each bin, which is the same for both the Barkjohn 2021 and new EPA corrected values. Tthe error bars show one standard deviation within that bin.

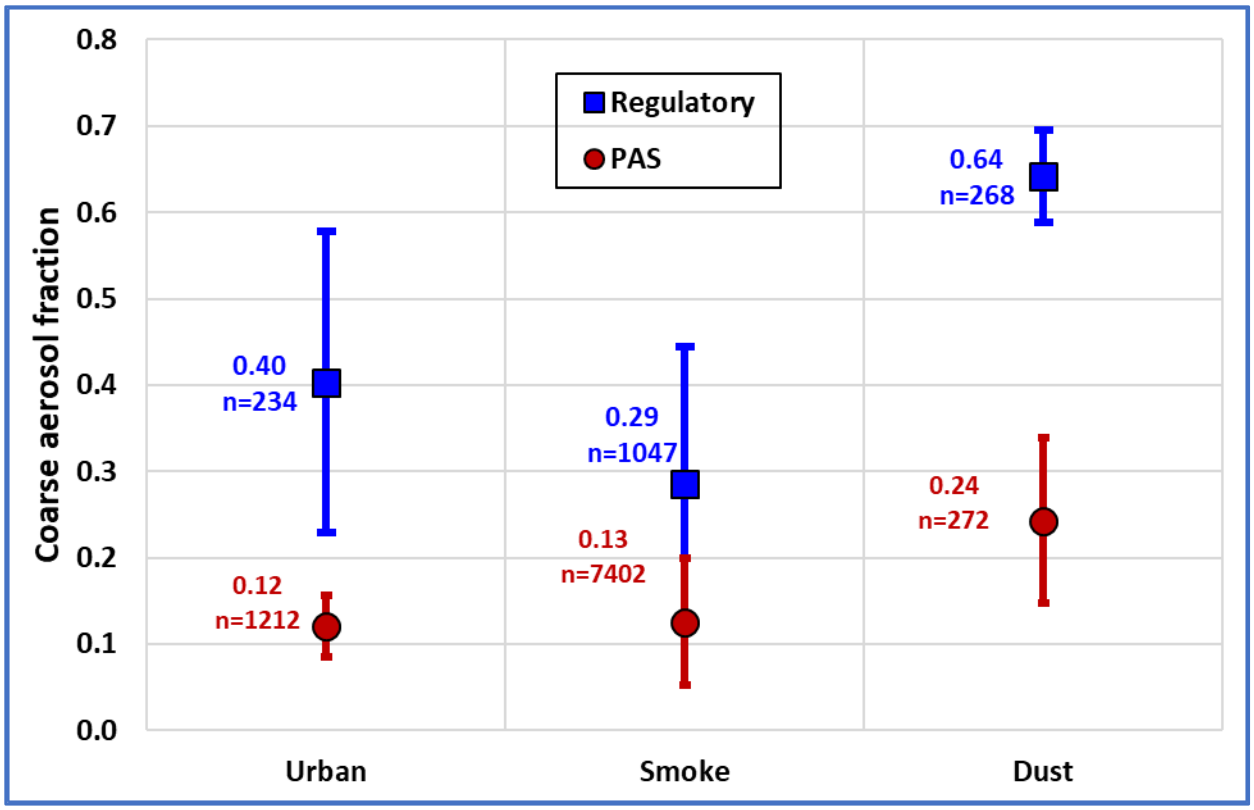


Figure 3: Mean coarse aerosol fraction (CAF) (~~Equation~~ Eq. (3)) calculated using the regulatory data and the PAS raw data for 17 events from the Part I dataset that had both $PM_{2.5}$ and PM_{10}. The values near each point give the mean and number of data points (hours) in each bin.

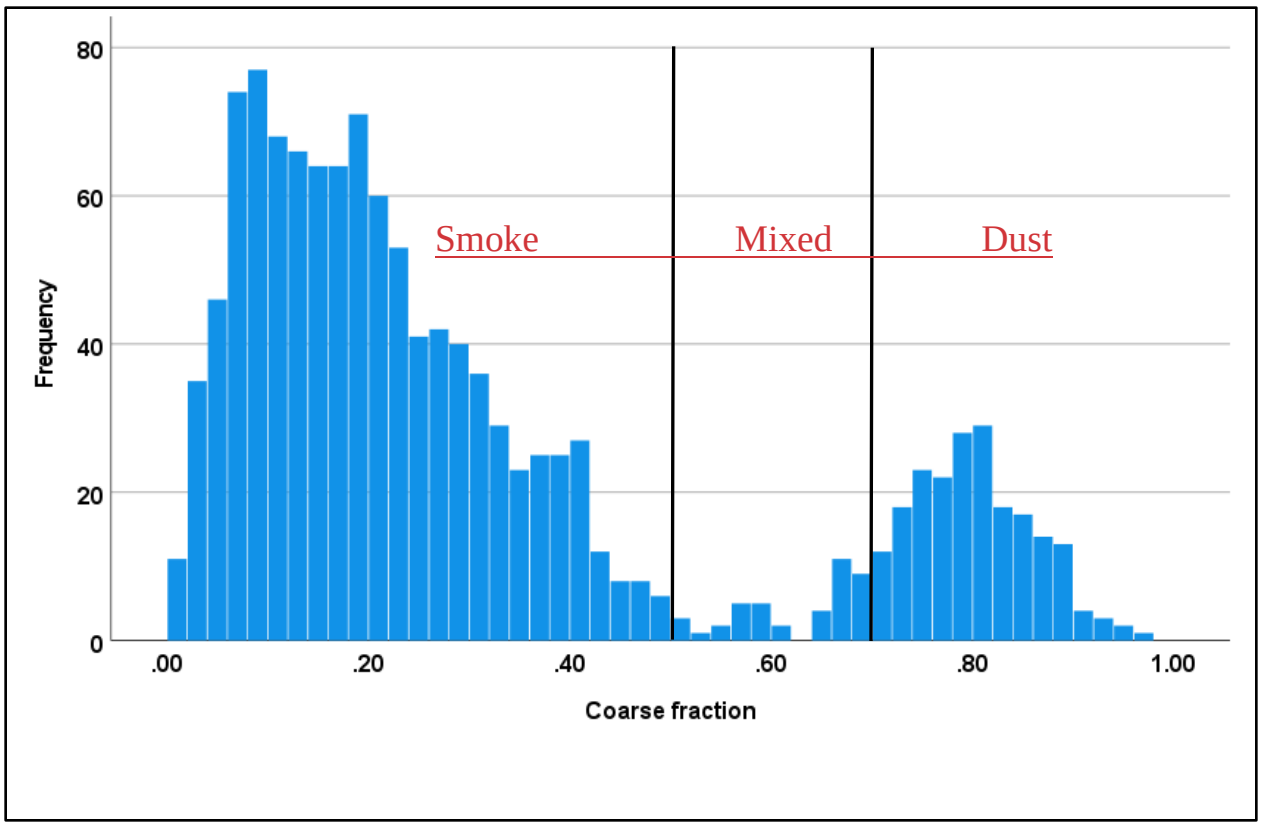


Figure 4. — Histogram of coarse aerosol fraction (CAF) at Keeler, CA, using regulatory $PM_{2.5}$ and PM_{10} data for hours with $PM_{2.5}$ > 25 µg m^{-3}.— We assume that the aerosol is primarily smoke when CAF <0.5, mixed when CAF is between 0.5 and 0.7, and dust for times with CAF>0.7.

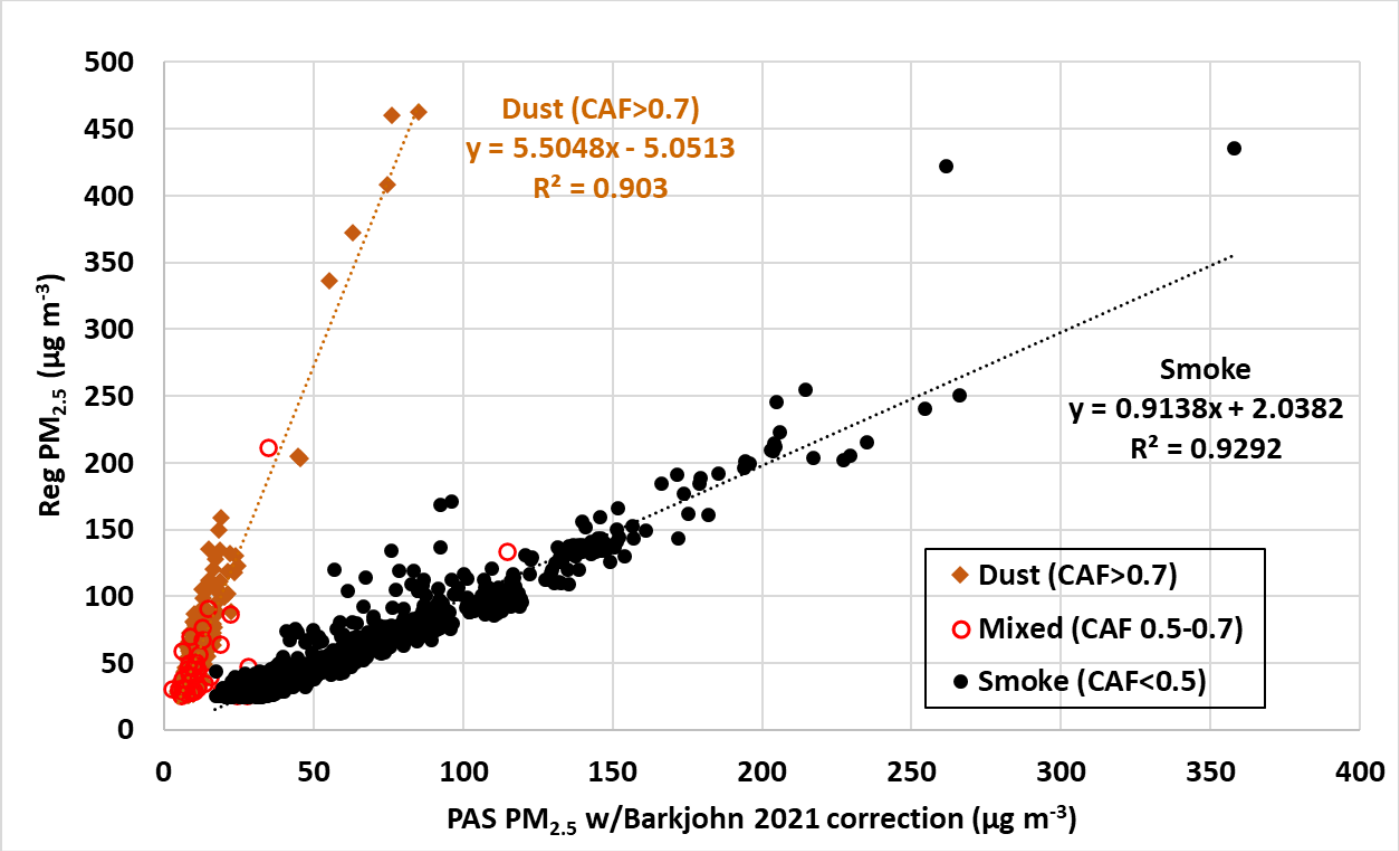


Figure 5. Regulatory $PM_{2.5}$ versus PAS $PM_{2.5}$ with Barkjohn 2021 correction at Keeler, CA, for hours with regulatory $PM_{2.5}$ > 25 µg m^{-3}. The data are separated by the coarse aerosol fraction (CAF)~~CAF~~, as measured by the regulatory data.

620

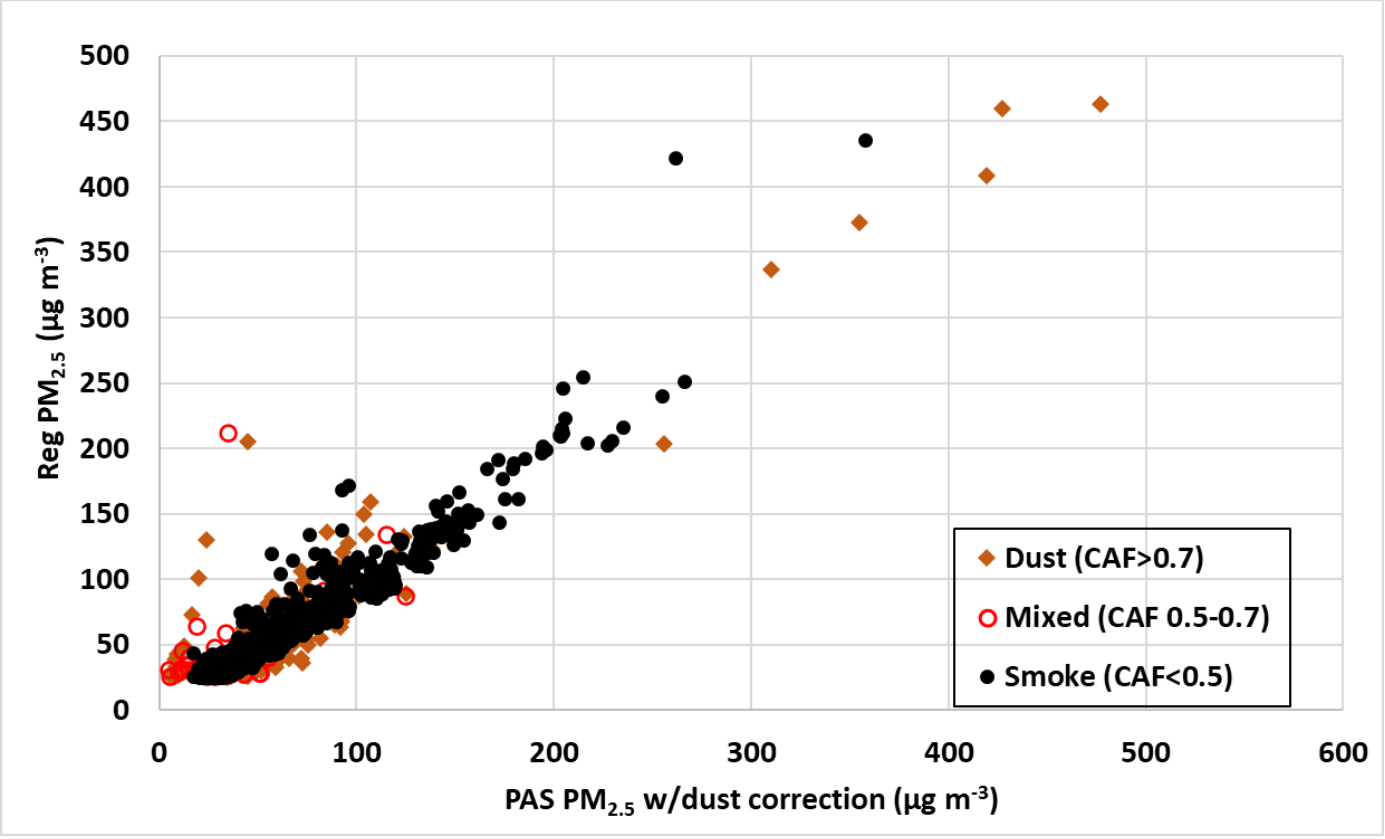


Figure 6. Regulatory $PM_{2.5}$ versus PAS $PM_{2.5}$ with dust correction (equation 4) at Keeler, CA, for hours with regulatory $PM_{2.5}$ > 25 µg m^{-3}. The data are separated by the coarse aerosol fraction (CAF), as measured by the regulatory data.