We would like to thank anonymous reviewers 1 for the helpful comments and suggestions. In line with the reviewer comments and suggestions, and in line with new publications that were published while this paper was under review, we modify and revised the manuscript. Below are all the comments (in bold) followed by the replies. The parts that are in italic are corrections that are included in the revised version of the paper.

Sincerely,
Karin Ardon-Dryer

Response to Reviewer 1

The work by Ardon-Dryer at al present a large database of PM2.5 mass concentrations collected by a low-cost sensor, the Purple Air PA-II unit across eight locations in the USA. At each location, there were measurements from multiple PA-II units, and the closest air quality monitoring station (AQMS) data was also acquired. The authors have performed a large amount of analysis on this dataset comparing the PA-II to reference instrumentation. However, it was not clear to me what the scientific novelty of the paper was, as there have been a number of papers already that evaluated the Purple Air sensor, as mentioned in the introduction. The authors state that the aim was to 'examine how PA-II units perform under atmospheric conditions when exposed to a variety of pollutants and PM2.5 concentrations', yet this is a rather vague aim, that this dataset may not be suitable to answer. This is a great dataset that could be used to investigate a number of interesting questions regarding low-cost sensors and their calibration and suitability for large scale deployment. I feel that this paper suffers from a lack of focus and could be improved if the authors articulated and addressed more novel, detailed and specific aims and objectives. This leads to another area that could be improved, as most of the analysis is rather descriptive and lacking in depth. In my opinion, it is not enough to just present the regression analysis for all the PA-II units (i.e. r2, slopes etc) to the AQMS instruments. For example, there could have been more analysis on why there was a large range in observed r2 between all the unit and the AQMS? Was there any common factors for units that had a poor or good correlation with the AQMS? Did the actual reference instrument at the AQMS site affect the correlation (e.g. between FRM and FEM instruments)? I would have also like to have seen more focus on the observed slope

between PA-II and the AQMS, as this is a better indicator of the accuracy of the PA-II than the correlation co-efficient (r2).

We appreciate the reviewer's comment, based on the review comment we clarify our aims in the manuscripts.

The following information was added to the manuscript

This study aims to examine how each PA-II unit performs under atmospheric conditions when exposed to a variety of pollutants and $PM_{2.5}$ concentrations (PM with an aerodynamic diameter smaller than 2.5 μm), when at a distance from the reference sensor. We examine how PA-II units perform in comparison to other PA-II units and Environmental Protection Agency (EPA) Air Quality Monitoring Stations (AQMSs) that are not co-located with them.

This study aims to examine how PA-II units perform under atmospheric conditions when exposed to a variety of pollutants and $PM_{2.5}$ concentrations. For the scope of this study, we chose to focus only on regions that contain at least one pair of co-located PA-II and AQMS units. Corrections of $PM_{2.5}$ values for co-located PA-II and AQMS units, based on MLR, were performed and applied to all the other PA-II units in that region. Comparison of $PM_{2.5}$ measurements taken by all units in each region, AQMSs and PA-II units (when $PM_{2.5}$ values were measured or corrected) are presented. The presented comparisons were done for both the entire study period and for specific events that we wanted to examine in greater detail.

Regarding the comments about factors for units that had a poor or good correlation with the AQMS, In our original manuscript most of the PA-II units that had a low correlation with the AQMS units also suffer from low correlation with the other PA-II units, we decided to remove these units as we believe they are outliers. In the current dataset, there are only two units that had a lower correlation with the AQMSs, but these two units were borderline for our PA-II outlier test. Without these two units, most of the R^2 values will be >0.6. The evaluation test between the PA-II units will help identify PA-II units that are not performing well. A reduction in performance can occur over time or due to exposure to events with high PM, as described in Sayahi et al. (2019). This information was added to the manuscript.

The following information was added to the manuscript

Overall, almost all the PA-II units had high correlation values when compared with the other PA-IIs or AQMSs in their region. Two PA-II units, SL-PA-6 and SL-PA-8 had low R^2 values with the AQMS, they also had a relatively low correlation with the other PA-II units. It is feasible, that if stricter rules for identifying outlier PA-II units were in use, these two units would have been considered as such and subsequently removed from the data set.

As for the comment **Did the actual reference instrument at the AQMS site affect the correlation (e.g. between FRM and FEM instruments).**

All the AQMSs that were used in this work were of FEM type, their selection was based on the distance that was used in previous works (e.g. Bi et al., 2020). Therefore, we could not evaluate that difference (FEM vs FRM).

Regarding the reviewer comments ***that like to have seen more focus on the observed slope between PA-II and the AQMS, as this is a better indicator of the accuracy of the PA-II than the correlation co-efficient (r2).*** Information on the slope was added to the manuscripts for all comparisons as shown in Table 1, Table S3, and Fig. 7.

One of the key issues with this dataset, as acknowledged by the authors in section 3.3.3, was that the PA-II units were not co-located with each other or the AQMS and could therefore diminish the ability to compare the PA-II to reference instruments. Unfortunately, in my opinion the authors did not adequately address this issue. It would have been interesting if a more in-depth analysis of how the PA-II relationship with reference instrument varied as function of distance, as this would be of great interest to the community.

The goal of this paper was to observe $PM_{2.5}$ measurements using the PA-II units, these units installed by citizens are for the most found in residential locations across the United States, therefore only a handful are co-located with an AQMS, and in fact many of the regions in which PA-IIs are deployed do not have even a single reference unit. Previous works have examined the efficiency of the PA-II unit by comparing it to a co-located AQMS or in laboratory conditions. For

the purpose of this study we defined a co-located pair as a PA-II that is up to 1.1 km away from an AQMS, the selection of this distance is based on the work of Bi et al. (2020). A major addition to this revision is the implementation of a data correction process that was applied to the PA-II measurements. This correction process was well documents by Bi et al. (2020) and Magi et al. (2020) for both PA-II units and by Malings et al. (2020)for other low cost sensors.

As for the impact of the distance of the units, we did not find that the distance between the units impacted the behavior and comparison of the unit, yet we only evaluate a distance of up to 5 km from an AQMS and up to 10 km between PA-IIs, units which will be far away may have a different impact, but evaluating that would be beyond the scope of this work.

The paper is well written and clearly presented but the large volume of data presented did make it difficult to follow at times. For example, the tables are too big, and could do with either being separated by city, or only the pertinent information being included.

We modified the provided tables and information in the text. Each table now represents a single region and does not include more than two parameters.

In addition to the above, number of more detailed comments are given below Abstract: When you state that the units had good agreement it is important to back this up with numbers, such as giving the slopes, r2 etc. This generally true throughout the paper.

Information on the comparisons between units was added to the manuscript per region. Values of the R-squared (R^2), root mean square error (RMSE), mean absolute error (MAE) as well as the best fit information, including the slope, are provided in the revised manscript.

Based on the reviewer comments the following information was added to the manuscript
*In most cases, the AQMSs and the PA-II units were found to be in good agreement (*75% of the comparisons had a $R^2 > 0.8$)

Page 2, line 63: In addition, the authors could reference Crilley et al 2018 and Di Antonio et al 2018 for possible solutions to the RH effect on low-cost PM sensors.

Both references were added to the manuscript

Previous studies suggested that part of the problem with the PA-II unit results from the optical particle counter being impacted by changes of RH (Crilley et al., 2018; Malings et al., 2020; Magi et al., 2020). .... deliquescent or hygroscopic growth of particles, mainly under high RH conditions, can lead to higher reported PM concentrations (Di Antonio, 2018; Jayaratne et al., 2018; Bi et al., 2020), which ends as an overestimate of the PM compared to the reference units.

Page 5, line 152: this paragraph could instead be presented as a table. Furthermore, it may also help the reader if you were to give the AQMS and PA-11 units more accessible names. For example, the Pittsburgh AQMS could be P-AQ-1 and 2, and the PA-II units, P-PA-1, 2, 3 etc

Per the reviewer's comments, the entire paragraph was removed from the manuscript. Also, all unit's names are now represented by location and instrument code as well as running ID number, as specified in Table S1.

For simplifications, each region was defined by two letters to represent its name (DE for Denver, SF for San Francisco, VA for Vallejo, and SL for Salt Lake City). Also, each unit type received a two letter code (AQ for AQMS and PA for PA-II). Each unit received a number instead of an ID, as shown in Table S1.

Section 2.4: more info is needed on the data analysis, what sort of regression analysis did you do? In what computer program? Which AQMS station did you use, the closest or each one for a given city?

In the manuscript, we describe that we used Multivariate linear regression (MLR) models between the $PM_{2.5}$ values of the co-located PA-II and AQMS with T and RH. In addition, all the analyses were performed using Matlab and Excel. This information was added to the manuscript

To evaluate the similarities and differences between the PA-II units and the AQMSs and other PA-II units, a set of calculations and comparisons was performed using Matlab and Excel

Regarding the AQMS database, we downloaded the entire data set of the hourly $PM_{2.5}$ for all AQMS units that were active during the study period. Using a distance calculation, we were able to identify regions with multiple PA-II units as well as at least one AQMS. We added this information to the manuscript

Hourly measurements of $PM_{2.5}$ (FRM/FEM Mass code - 88101 file) from all AQMSs collected by the EPA from January 1, 2017, to December 31, 2018, were selected from the EPA website (https://aqs.epa.gov/api).

By using the JSON file for the PA-II units and the 88101 file for the AQMS, we calculated the distances between all the units to identify regions with multiple PA-II units (a minimum of five units) and at least one AQMS. At least one AQMS unit needed to be at a distance of 1.1 km from at least one PA-II unit (defined as a co-located pair, a similar range used by Bi et al., 2020). All the units in these regions needed to be active during the designated time period of January 1, 2017, to December 31, 2018. In each region PA-II units needed to be less than 5 km from at least one AQMS unit and up to 10 km from each other.

Four different regions containing a total of seven different AQMSs (all FEM type) and 46 different PA-II units were identified:

Section 2.6: I do not see the point of calculating the AQI when the point of this article is to compare the measurements between the PA-II and reference instruments. If they report the same concentration, wont they give the same AQI? I think you should just focus on reported concentrations.

AQI information was removed from the manuscript per the reviewer's comment.

Section 3.1.1: If Fig 2 is on page 29, then this is not a distribution but a time series of reported concentrations. A distribution to me implies a histogram, please correct the naming. Also why did the AQMS report higher PM2.5 concentrations at Berkley, Ogden, Linden and Salt Lake City compared to all the PA-II units during the first half of 2018? Understanding why the relationship changed is important for knowing the parameters that affect the PA-II measurements.

This plot is now presented as a time series per the reviewer's comment.

Time series of daily $PM_{2.5}$ values for each unit at each of the four regions are presented in Fig. 3.

Some of the regions mentioned in the reviewer comment have been removed from the manuscript as they were missing a co-located AQMS. For the remaining regions, the higher AQMS measurements are attributed to what seems to be a connection with days that have low RH values resulting in lower $PM_{2.5}$ values being measured by the PA-II units. We also believe that chemical analysis during for these times could help understand the difference between the AQMS and PA-II, unfortunately, such analysis will be beyond the scope of this study.

In some cases, the AQMS measured higher $PM_{2.5}$ daily values compared to the PA-II units, mainly at days with low $PM_{2.5}$ values, as seen in April - June 2018 in Vallejo (Fig. 3C) and Salt Lake City (Fig. 3D). These differences were observed mainly in days with low RH values (Fig. S3).

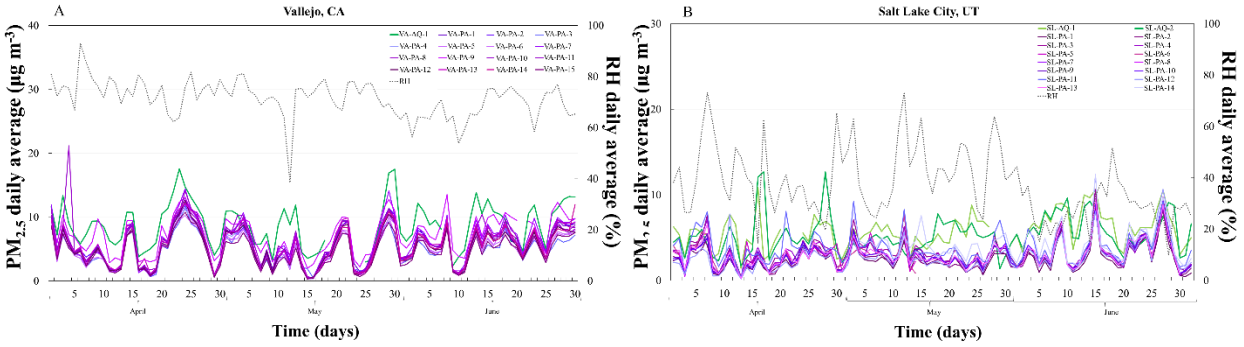

Figure S3: *Time series of daily $PM_{2.5}$ measurements from the AQMS and PA-II units in Vallejo (A), and Salt Lake City (B) during April-May 2018. Measurements from AQMS are represented by the green lines and the PA-II units are indicated by purple lines. Relative Humidity values represented by the gray dotted line.*

Page 8, line 236: the authors state "These high correlation values and relatively low RMSE indicate that although the PA-II units and the AQMS are not co-located, they still tend to behave in a similar way." Why do think this was the case?

This sentence was removed from the current manuscript.

Page 8, line 242: I do not understand what you mean by instrument efficiency?

This sentence was removed from the current manuscript.

Page 8, line 250: why did you subset the data below 40 ug m-3?

This analysis was removed from the current manuscript. Originally, we set 40 ug m-3 as the maximum point for the study as the work of Sayahi et al. (2019) suggests that above it the PA-II measurements are impacted by the high PM concentrations. Meaning at lower PM concentration we will find a better correlation between the PA-II and the AQMS.

Page 12, line 364. In the previous paragraph you state that RH is a more important parameter than temperature when considering potential artefacts for the PA-II, so why compare to temperature?

This sentence was removed from the current manuscript.

The original sentence was based on findings from several papers. We originally compared the temperature in order to prove our theory that temperature is not as important. However, during the time that our original manuscript was under review several new papers were published which in turn made us make extensive changes in our manuscript regarding the impact of RH and T. We added a hunidogram and a plot that investigate the impact of T in each of the co-located units (AQMS with PA-II, Fig S4). Some of the PA-II units might be impacted by both T and RH, this information was added to the manuscript

Calculations of the ratio between the measured $PM_{2.5}$ from the PA-II to the AQMS as a function of T and RH, known as a hunidogram, were performed (Fig S4). Some of the PA-II units seem to be impacted by T and RH more than others; these units also had relatively low R^2 values with the AQMS unit, as in the case of DE-PA-6 in Denver (Fig. S4A).

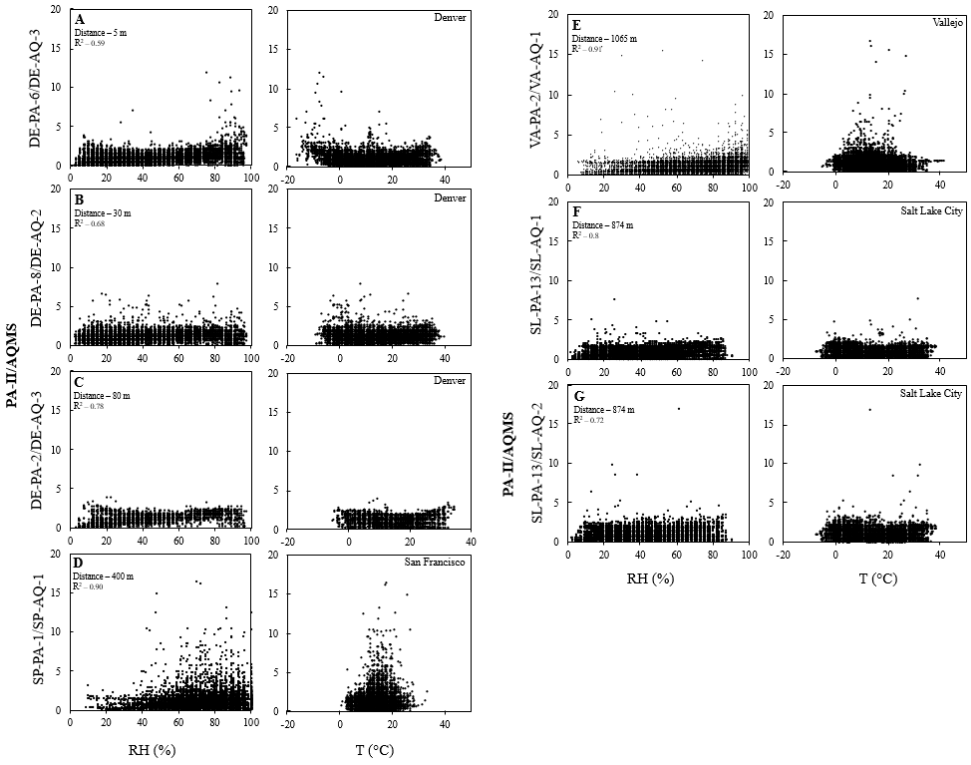

Figure S4: *Ratio between measured $PM_{2.5}$ from PA-II to the AQMS, as a function of temperature and relative humidity (hunidogram) for all collocated PA-II and AQMS pairs. Information on the distance and R^2 values between the two presented in each plot.*

We also added a multivariate linear regression (MLR). The MLR takes into account changes of T and RH. All the PA-II units' measurements were corrected based on the MLR of the co-located PA-II and AQMS, this information was added to the manuscript:

Based on the MLR, the multivariable linear dependence of PA-II $PM_{2.5}$ on AQMS, RH and T created the predictors of PA-II as:

$$PA - II(PM_{2.5}) = A_1 + A_2 AQMS(PM_{2.5}) + A_3 T + A_4 RH$$

(1)

where A_1, A_2, A_3, and A_4 fit coefficients received from the MLR, PA-II ($PM_{2.5}$) and AQMS($PM_{2.5}$) are in units of µg m^{-3}, T is in Celsius, and RH is in percentage. Based on these parameters and fit coefficients, a calculation of the corrected PA-II $PM_{2.5}$ hourly values for each PA-II was performed using the following:

$$PA - II(PM_{2.5}), corrected = \frac{PA-II(PM_{2.5}),uncorected-A_1-A_3T-A_4RH}{A_2}$$

(2)

Details of the coefficients received in the MLR as well as the regression output including R^2, RMSE, MAE, and slope for each correction of $PM_{2.5}$ values in the PA-II units, for each region, can be found Table 1. Figure 3 presents a comparison of the $PM_{2.5}$ values from the uncorrected PA-II unit to the AQMS as well as the PA-II $PM_{2.5}$ values hourly after correction, per region.

Page 12. Line 381: I do not agree with this statement as you have not able to test the precision of the PA-II as they were not co-located. The precision of the PA-II units would be tested by how well each PA unit agree with each other at a given RH, but you have looked for correlation between RH and PA-II reported PM2.5. this does not indicate the precision of the PA-II only if there was a relationship between RH and reported PM2.5 concentrations.

This sentence was removed from the current manuscript.

Page 13, line 418: where the slopes between the PA-II and AQMS instruments affected by distance?

We did not find an impact of the distance on the slop., No impact was observed when the PA-II units were compared to the nearest AQMS in all regions, and no effect was found when comparing the PA-II to each other.

This information was added to the manuscript as text and figure:

Because the AQMS and the PA-II units were not co-located, we wanted to verify whether the distance between all the units affected the R^2, RMSE, MAE and slope values. We compared the $R^{2,}$

RMSE, MAE and slope values received from the comparisons of hourly $PM_{2.5}$ measurements with the corresponding distances between the units (Fig. 7). There was no correlation between the two. Not when the PA-II units were compared to the nearest AQMS units (Fig. 7A), or between the PA-II units (Fig.7B), before or after the corrections of the PA-II $PM_{2.5}$ values. Therefore, the distance between the units did not impact the comparison.

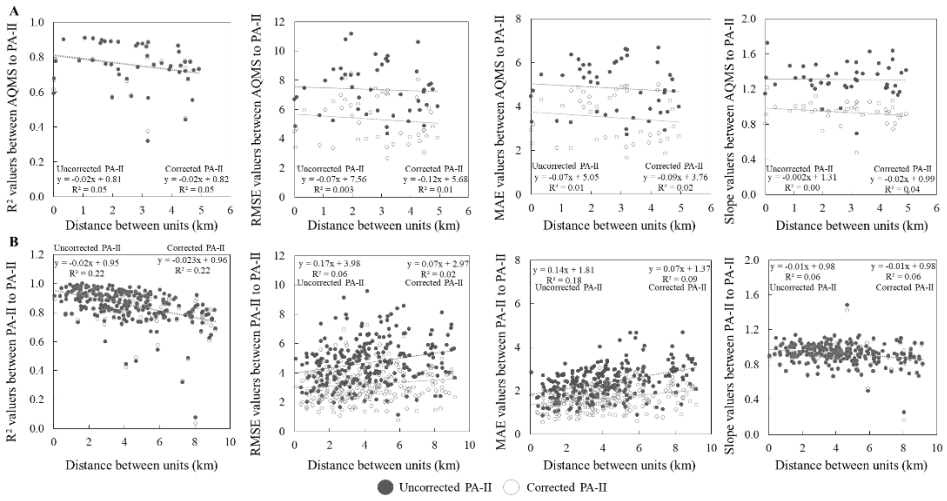

***Figure 7**: Comparison of distance (km) between PA-II to its nearest AQMS in all regions (A) and between each PA-II unit to all other PA-II units per region (B) to R^2, RMSE, MAE and slope values received from the $PM_{2.5}$ hourly measurements comparison.*

Section 3.4: I think that section could be improved by including some recommendations based on your findings from this study.

The original section was removed from the manuscript, instead we did implement it as part of section 3.5"Underlying Differences and Future Implications". As suggested by the reviewer we added several recommendations, this includes but not limited to how to use the measurements of the PA-II units, the necessity of Temperature and Relative Humidity measurements, steps for assuring the unit integrity and more.

This information was added to the manuscript:

3.5. Underlying Differences and Future Implications

While appropriate PA-II $PM_{2.5}$ value corrections can improve the comparison between the PA-IIs with reference units, there are other differences between PA-IIs and AQMS units that can influence

the comparison results, including the underlying technology and the manner in which units are placed. The $PM_{2.5}$ sensors in the AQMSs perform gravimetric measurements using the mass of the particle; by contrast, the PA-II units use a laser particle counter to count electric pulses generated as particles crossing through a laser beam. The method used by the PA-II might impact the count of particles during high humidity conditions or when a majority of the particles are volatile. Another difference is the physical location of the units; whereas AQMSs are meticulously positioned in an open area, the location of a PA-II is determined by its owner. Although PurpleAir recommends positioning the PA-II in an open area, ultimately, it is the owner's decision. In practice, most of the PA-II units are located in residential areas with low-rise housing. Furthermore, the height at which the sensor is located could affect the measurements. The height of the AQMS inlet is regulated and kept constant at each location; on the other hand, the owner of a PA-II unit can freely place it near the ground or higher up. The location of the PA-II units in residential areas can provide both an advantage and a disadvantage. For example, a single PA-II unit might be exposed to more localized PM sources such as a barbecue, lawn mower, or car, making it report different results compared with other units in its area. Therefore, an increase of PM by a single PA-II unit should be taken into account. When the PA-II is used as a network, as suggested by Ford et al. (2019), comparison of the PM values measured by all PA-II units will help identify such a localized source. Maintenance and calibration are other possible causes of differences between the two. The $PM_{2.5}$ sensors in the AQMSs have strict rules for the monthly evaluation of sensor performance, including through flow calibration or calibration based on minimum value threshold (which, in some cases, causes the recording of negative PM values). By contrast, PA-II units do not have any quality control other than that done by the company for each sensor before shipment to the customer (PurpleAir personal communication, 2019). Another point that should be taken into account is the lifetime of the PA-II units. The manufacturer of the PMS5003 sensor used in the PA-II units states that it has a lifetime expectancy of ~3 years (Yong, 2016). Bi et al. (2020) found that the PA-II unit's efficiency is affected even after only two years of being operational.

Based on the findings from this work, we believe that there are several needed steps that will allow the usage of the PA-II units in air quality and health related research. First, users should identify regions with multiple PA-II units, where at least one PA-II is co-located with an FRM or FEM

unit. Ideally the same location will also contain measurements of T and RH, or at least T and RH measurements will be nearby. Keep in mind that it is not recommended to use the PA-II internal sensors for T and RH values, as they are not representing the atmospheric measurements (Malinges et al., 2020; PurpleAir personal communication, 2019). However, we have found that in many regions there is no meteorological station that can serve as a reference for the correction process. It would be useful then, to devise a way in which the PA-II internal T and RH sensors can be used. To achieve this, an extensive study is necessary, to gain a better understanding of the issues related to the usage of the PA-II internal sensors and to formulate a calibration equation that then can be applied to the desired PA-II units.

Comparison of all PA-II units in each region will help to identify and remove outlier PA-II units from future analysis. Exposure to high PM concentration might affect the PA-II efficiency, as suggested by Sayahi et al. (2019), and therefore, its measurements will differ substantially from those of the AQMSs and other PA-II units. Ideally PurpleAir should monitor all active PA-II units and identify units that behave differently from surrounding PA-II units or identify PA-II units whose internal sensors (A and B) report different values, flag them on the online map, and communicate instructions to the unit owners on how to fix or replace the unit.

After PA-II units have been identified, users should conduct MLR between the co-located PA-II and AQMS units, including measurements of T and RH. For the MLR to be efficient it is important have a wide range of $PM_{2.5}$, T and RH measurements. This MLR will provide a coefficient that will be used to correct all the remaining $PM_{2.5}$ values of all PA-II units in that region. Evaluation of the PA-II $PM_{2.5}$ value corrections should be made for the duration of the study but also for specific events with spatial impact such as inversion, dust storms, biomass burning, and more. Such events should impact a larger area and therefore will allow detection of the PM changes in all PA-II units as a whole (network). Correction of PA-II $PM_{2.5}$ values should be performed per region, as they represent specific PM values as well as changes of T and RH values that the PA-II units were exposed to. This will help the public obtain information on the spatial and temporal distribution of PM concentrations in their area (Gupta et al., 2018; Morawska et al., 2018), which will enable them to monitor local air-quality conditions (Williams et al., 2018) and help make decisions related to events with high PM exposure.

In this study, we evaluated PA-II units that were up to 5 km away from an AQMS unit, as well as up to 10 km from each other. This raises the question of maximum effective distance. What is the maximum distance between an AQMS and PA-II units that will still allow for the MLR to successfully correct the measurement taken by PA-II units; a distance greater than this would carry the potential of introducing additional factors that might impact the comparisons. Another situation that requires further investigation is that of regions that include multiple PA-II units but do not have a co-located pair or completely lack a reference monitoring station. The question in mind, if and how we might use neighboring regions in which measurements were successfully corrected to compensate in the case of such problematic areas. For example, could we have used Vallejo and San Francisco, two regions that were included in this study to correct the measurements of the PA-II units in the region of Berkeley - Oakland that resides between the two?

Page 14, line 433: please call it instrument drift, as instrument efficiency is meaningless in this context.

This sentence was removed from the manuscript.

References

Bi, J. Wildani, A., Chang, H. H. and Liu, Y.: Incorporating Low-Cost Sensor Measurements into High-Resolution $PM_{2.5}$ Modeling at a Large Spatial Scale. Environ. Sci. Technol. 2020, 54, 2152−2162, DOI: 10.1021/acs.est.9b06046.

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

Di Antonio, A. Popoola, O. A. Ouyang, B. Saffell, J. and Jones, R. L.: Developing a relative humidity correction for low-cost sensors measuring ambient particulate matter. Sensors 18, 2790, https://doi.org/10.3390/s18092790 (2018).Ford, B., Pierce, J. R., Wendt, E., Long, M., Jathar, S., Mehaffy, J., Tryner, J., Quinn, C., van Zyl, L., L'Orange, C., Miller-Lionberg, D., and Volckens, J.: A low-cost monitor for measurement of fine particulate matter and aerosol optical depth – Part 2: Citizen-science pilot campaign in northern Colorado, Atmos. Meas. Tech., 12, 6385–6399, https://doi.org/10.5194/amt-12-6385-2019, 2019.

Gupta, P., Doraiswamy, P., Levy, R., Pikelnaya, O., Maibach, J., Feenstra, B., Polidori, A., Kiros, F., and Mills, K. C.: 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, doi.org/10.1029/2018GH000136, 2018.

Jayaratne, R., Liu, X., Thai, P., Dunbabin, M., and Morawska, L.: The influence of humidity on the performance of a low-cost air particle mass sensor and the effect of atmospheric fog, Atmos. Meas. Tech., 11, 4883-4890, doi.org/10.5194/amt-11-4883-2018, 2018.

Magi B. I., Cupini, C., Francis, J., Green, M. and Hauser C.: Evaluation of PM2.5 measured in an urban setting using a low-cost optical particle counter and a Federal Equivalent Method Beta Attenuation Monitor Aerosol Sci. Technol., 54,147-159, 2020.

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

Morawska, L., Thai, P. K., Liu, X., Asumadu-Sakyi, A., Ayoko, G., Bartonova, A., Bedini, A., Chai, F. Christensen, B., and Dunbabin. M.: Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: how far have they gone?, Environ Int., 116, 286-99, doi: 10.1016/j.envint.2018.04.018, 2018.

PurpleAir, PurpleAir Map, air quality Map: http://map.purpleair.org/, last access: 1 August 2019.

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

Williams, R., Nash, D., Hagler, G., Benedict, K., MacGregor, I., Seay, B., Lawrence, M., Dye, T., September 2018. Peer Review and Supporting Literature Review of Air Sensor Technology Performance Targets. EPA Technical Report Undergoing Final External Peer Review. EPA/600/R-18/324.

Yong, Z.: Digital universal particle concentration sensor, PMS5003 series data manual: http://www.aqmd.gov/docs/default-source/aq-spec/resources-page/plantower-pms5003-manual_v2-3.pdf, last access: 1 August 2018.

We would like to thank anonymous reviewers 2 for the helpful comments and suggestions. In line with the reviewer comments and suggestions, and in line with new publications that were published while this paper was under review, we modify and revised the manuscript. Below are all the comments (in bold) followed by the replies. The parts that are in italic are corrections that are included in the revised version of the paper.

Sincerely,
Karin Ardon-Dryer

Response to Reviewer 2

This is an informative manuscript that evaluates the performance of networks of the PurpleAir PA-II low-cost aerosol sensor in real-world use. These sensors are commonly purchased by private citizens and installed, sometimes haphazardly, in residential and commercial neighborhoods. They are quite low-cost (<$300/unit) and data from these sensors could be used to increase understanding of the spatial distribution of PM2.5 and supplement more comprehensive, but much more costly and less ubiquitous, air quality monitoring stations (AQMS). The real question is whether these sensors provide data of adequate quality to be useful. The paper is generally clear and well-written, and it makes a strong case that the sensors have value and can provide scientifically useful information, at least under the conditions evaluated. It is also nice to see a high school student involved in the study. That said, there are some changes that need to be made to improve the manuscript. In particular, the evaluation of the sensitivity of the sensors to relative humidity (RH) and temperature (T) needs to be reworked, and some of the information in the tables could be presented more effectively with graphics. Below are major concerns, followed by a couple of minor issues. I have not checked the references for completeness.

1) In section 3.3.1, the effect of RH and T on unit performance are evaluated by regressing these values against the PM2.5 values from the PA-II units. Unsurprisingly, there was no significant correlation against either of these parameters. Instead, what needs to be compared is RH and T against the *difference* between the PA-II units and the nearest AQMS values. Biases associated with T and RH are minimized in the AQMS sensors but

would show up in the PA-II sensors, which do not control sample RH or T (although T is higher inside the sensing elements; thus we would expect RH to be reduced significantly below ambient). Any large bias associated with RH or ambient T should show up in this comparison (except see minor comment (b) below).

We took into consideration the reviewer comments, therefore we made extensive changes in our manuscript. First, we added an evaluation of the PA-II sensitivity to RH and T, for co-located PA-II units with AQMS (Fig S4). We also added an entire paragraph that discusses the impact of RH and T on the PA-II. Also, based on the reviewer's comments and suggestions as well as new publications that were published while the original manuscript was under review, we added an entirely new analysis to the paper. We performed a multivariate linear regression (MLR) on the co-located units (PA-II and AQMS, that were at a distance up to 1.1 km) and used the coefficient from the MLR to correct that additional PA-II unit measurements taken in the same region. This correction of the PA-II PM2.5 values improve the comparison between the PA-II units and the AQMS as well as between the PA-II to other PA-II units, as showed by improving the slop and reduction of the root mean square error (RMSE) and mean absolute error (MAE) values.

The following information was added to the manuscript

The overestimating raises questions about the accuracy of the PA-II units. According to PurpleAir (PurpleAir, personal communication, 2019) the company does not calibrate the PA-II units; instead, before each PA-II unit is sent out to a customer, the company performs a comparison test with a dozen PA-II units to find and remove outliers from the shipment (PurpleAir, personal communication, 2019). Previous studies suggested that part of the problem with the PA-II unit results from the optical particle counter being impacted by changes of RH (Crilley et al., 2018; Malings et al., 2020; Magi et al., 2020). Water vapor can condense on aerosol particles, making them grow hygroscopically under high RH conditions (Lundgren and Cooper, 1969). The PA-II units do not have any heater or dryer at their inlets to remove water from the sample before measuring the particles; therefore, deliquescent or hygroscopic growth of particles, mainly under high RH conditions, can lead to higher reported PM concentrations (Di Antonio, 2018; Jayaratne et al., 2018; Bi et al., 2020), which ends as an overestimate of the PM compared to the reference units. Weather conditions can impact the values reported by low-cost sensors (Morawska et al.,

2018). Changes in T or RH have been found to affect the performance of the PA-II units, especially under atmospheric conditions, as they cannot be controlled (Bi et al., 2020). Therefore, MLR between a PA-II, and an AQMS, which also considers changes of T and RH, can help correct the reported $PM_{2.5}$ values of the co-located PA-II units. Similar corrections have been suggested and implemented in other locations with PA-II units (Bi et al., 2020; Magi et al., 2020) and other low-cost sensors (Malings et al., 2020). Most of these studies focus on co-located units or on units that were up to 1 km from the reference unit.

Calculations of the ratio between the measured $PM_{2.5}$ from the PA-II to the AQMS as a function of T and RH, known as a hunidogram, were performed (Fig S4). Some of the PA-II units seem to be impacted by T and RH more than others; these units also had relatively low R^2 values with the AQMS unit, as in the case of DE-PA-6 in Denver (Fig. S4A).

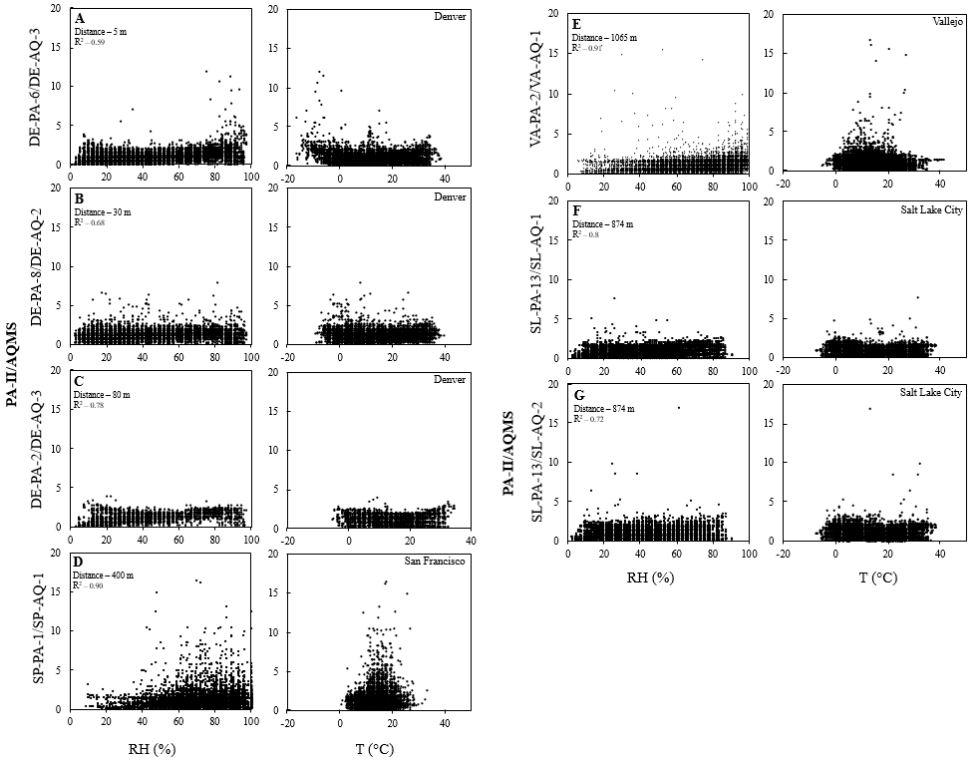

Figure S4: *Ratio between measured $PM_{2.5}$ from PA-II to the AQMS, as a function of temperature and relative humidity (hunidogram) for all collocated PA-II and AQMS pairs. Information on the distance and R^2 values between the two presented in each plot.*

2) There are a lot of values in tables in this manuscript, many of which really belong in the supplemental information. I would much prefer to see a new figure with scatterplots of each sensor against the AQMS values in the main text, and move Figs. S1 amd S3 there as well. The detail in the tables should be moved to the SI.

Per the reviewer's suggestions all tables were moved to the supplement (Now Tables S1-S3) and scatterplot of the PA-II compared to the AQMS were added to the manuscript (Fig. 4 and Fig.5). We compared the co-located units before and after we performed the MLR (Fig 4,) and observed the difference before and after we applied the coefficients from the MLR to the rest of the PA-II units (Fig 5.)

As suggested by the reviewer the figure with the map (originally Fig S1) was moved to the main manuscript, and it is now Fig. 2. Figure S3 was also moved to the main manuscript, it is now Fig. 7. We made changes in the figure, we evaluated the impact of the distance on the R^2, RMSE, MAE, and the slope values. This was performed both between the PA-II to the nearest AQMS as well as between the PA-II units.

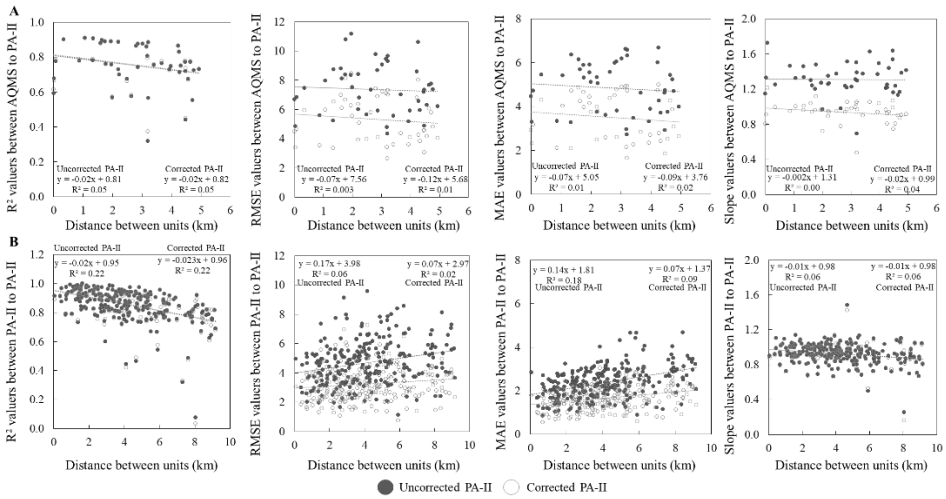

Figure 7*: Comparison of distance (km) between PA-II to its nearest AQMS in all regions (A) and between each PA-II unit to all other PA-II units per region (B) to R^2, RMSE, MAE and slope values received from the $PM_{2.5}$ hourly measurements comparison.*

3) The linear regressions should be performed with "2-sided" regressions because there is uncertainty in both the x and y values of the scatterplots. Standard linear regression assumes

uncertainty only in the y values. I also suggest you remove obvious outliers (for example, the July 4th fireworks smoke) before performing regressions; these outliers can severely torque the slopes and r2 values.

We apologize but we were unsure about the reviewer meaning for 2-sided regression and why he considers the AQMS measurements as uncertain. Our study, like others treats the AQMS as an absolute and does not question the validity or accuracy of its measurements.

As suggested by the reviewer we removed all the outlier's events before the statistical tests, we also performed an analysis that allows us to remove outlier PA-II units. A new section was added to the manuscripts about describing both.

The following information was added to the manuscript

2.5. Remove of outlier PA-II units and irregular hours

The first step was to identify outliers among the PA-II units, per region, meaning PA-II units that behave differently from the other PA-II units in their region. By comparing R^2 between the $PM_{2.5}$ values measured by each pair of PA-II units, using a linear regression, we identified the outlier units. A PA-II unit that did not have an $R^2 \geq 0.75$ with at least 75% of the other PA-II units in its region was considered an outlier unit, and therefore was removed from future analysis (Fig. S1 shows a comparison for each of the four regions). Only one unit from SF (SF-PA-9, see Fig. S1B) had very low R^2 when compared to all other PA-II units. Most PA-II units had high R^2 values (>0.9) with the other units. Irregular $PM_{2.5}$ hourly measurements were removed from all units (PA-II and AQMS). These irregular hourly measurements were identified as a large single hourly increase of $PM_{2.5}$ values (>70 $\mu g\ m^{-3}$) that was not measured by any other unit in the region. Such a large increase was caused most likely by a local source near a specific unit, such as a small-scale fire, lawn mower, barbecue, cigarette smoke, or fireworks (Zheng et al., 2018), and attributed to the location of many of the PA-II units in a residential area. Firework events were removed, as they were very localized events and were measured by a single unit. Overall, less than 0.03% of the hourly $PM_{2.5}$ measurements identified as irregular hours were removed from different PA-II and AQMS units.

4) There is lack of specificity in the abstract and throughout the text about "co-located" and "same location". I was quite confused when first reading the abstract, because it says that this manuscript reports analysis of PA-II units that are not "co-located" with AQMS sites, but then in the next sentence that "we selected eight different locations, where each location contains multiple PA-II units (minimum of seven per location, a total of 86 units) and at least one AQMS (total of 14)." This sounds to me like "colocated" because you have not specified the criteria used for selecting PA-II units. I suggest you use "nearby" or "regional" rather than "location" throughout the text to avoid confusion. And please define the distance criteria for which PA-II units were selected for comparison with AQMS instruments.

We apologize that our lack of clarity about the location aspect of the units. As the reviewer suggested we added more clarification to the manuscript. Co-located units are PA-II and AQMS units that are up to 1.1 km between each other, this is similar range to what was done by Bi et al. (2020) . We also changed the use of the word location to region as suggested by the reviewer. In addition, we provide the extract criteria for a distance that was used in our analysis to define each region.

This information was added to the abstract:
For this study, we selected four different regions, each containing multiple PA-II units (minimum of seven per region). In addition, each region needed to have at least one AQMS unit that was co-located with at least one PA-II unit, all units needed to be at a distance of up to 5 km from an AQMS unit and have up to 10 km between each other.

5) You may want to explore the seasonality of differences between the PA-II units and the AQMS values. For example, in winter in Utah, I would expect big gradients between airport sensors on the flat plains and residential sensors on the slopes. This may become evident in the analysis I suggest in comment (1) above.

Per the reviewer comments we analyzed the seasonality differences between the PA-II units and the AQMS values in all four regions, we attempted to identify the impact of T and RH as suggested by the reviewer. All regions had lower R^2, RMSE and MAE values in the spring compared to the

other seasons, however, this difference was not statistically significant for all cases. Next, we calculated the average RH and T for each season, and we compared it to the R^2, RMSE and MAE values. To our surprise there was no seasonal impact of RH or T on these values. We found that the lower R^2, RMSE and MAE values in the spring result from the overall lower PM2.5 values measured in that season for all four regions (as can be seen in Fig.3 in the manuscript). The PM concentrations had a stronger impact on the PA-II and AQMS comparisons than the T and RH had, therefore, we decided not to include this analysis in the manuscript.

As for the reviewer's example, we explored the spatial changes between the PA-II units, mainly in Salt Lack City, Utah as suggested by the reviewer. All the units that we used in the study were in residential area and not next to the airport. Overall, most sensors behaved in a similar way, as shown by the figure below. A similar range (bins of 5 μg m-3) of $PM_{2.5}$ concentration were measured by all the units. However, in the very few cases in which we observed some spatial differences (mainly in August 2018, as shown in the Figure below), we could determine the causes of these differences.

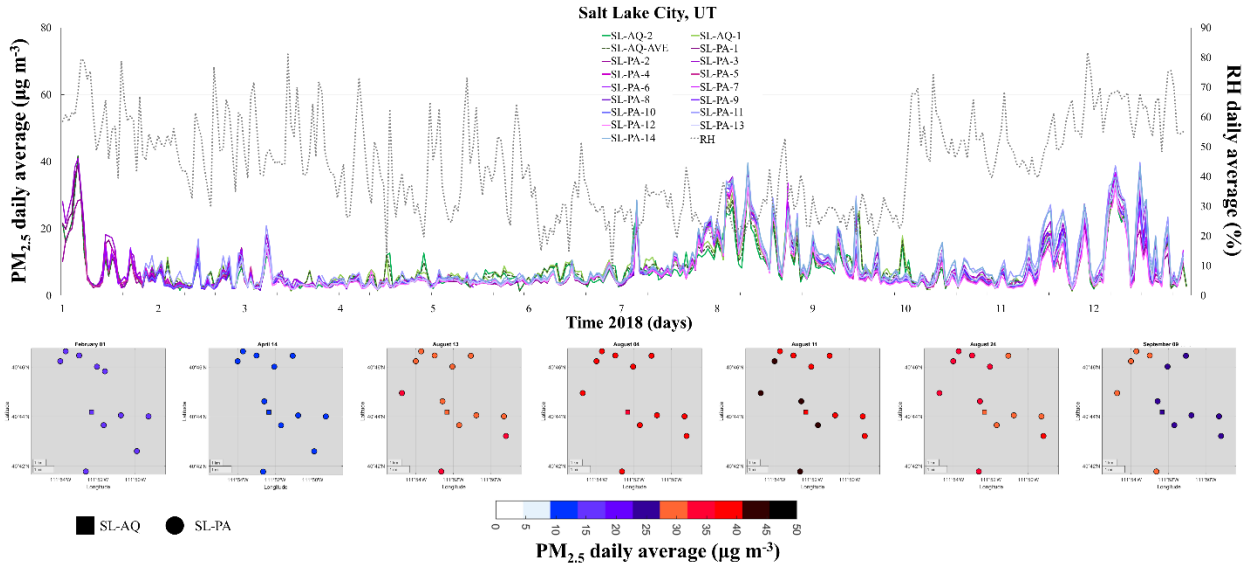

Fig 2. Time series of daily $PM_{2.5}$ measurements from the AQMS and PA-II units in Salt Lake City during 2018 (top). Measurements from AQMS are represented by the green lines and the PA-II units are indicated by purple lines, RH values represented by the gray dotted line. Maps of different days during 2018 with the spatial distribution of the daily $PM_{2.5}$ measurements (lower panel).

AQMS represented by the square and PA-II by round shape. Each color represents $PM_{2.5}$ values in bins of 5 μg m^{-3}.

Minor comments:

a) Lines 151-164. These are not needed; this information is already in the tables.

This entire paragraph was removed from the manuscript

b) In Sect. 3.2.2., these differences between the AQMS values and the PA-II data in Utah in winter may be associated with the volatility of ammonium nitrate, which dominates the aerosol composition there (Womack et al., https://doi.org/10.1029/2019GL082028). The PA-II instrument would be less likely to volatilize ammonium nitrate, while the NAAQS FRM does volatilize it (Grover et al., https://doi.org/10.1029/2004JD004995).

We would like to thank the reviewer for bringing up this point. The reviewer comments regarding the volatility of ammonium nitrate helped us to understand one of the causes for the increase of $PM_{2.5}$ in Salt Lake City during the winter months. We added this information to the manuscript

The following information was added to the manuscript

On average the PA-II values were higher by 2.1 ± 2.6 μg m^{-3} from those measured by the AQMS. The still higher $PM_{2.5}$ values could be due to the volatility of ammonium nitrate, which is a dominant aerosol composition at the region of Salt Lake City during the winter times (Moravek et al., 2019; Womack et al., 2019). It has been shown that sensors similar to the ones used in the PA-II units would be less likely to volatilize ammonium nitrate, unlike the one used in the AQMS units (Grover et al., 2005).

References

Bi, J. Wildani, A., Chang, H. H. and Liu, Y.: Incorporating Low-Cost Sensor Measurements into High-Resolution $PM_{2.5}$ Modeling at a Large Spatial Scale. Environ. Sci. Technol. 2020, 54, 2152−2162, DOI: 10.1021/acs.est.9b06046.

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

Di Antonio, A. Popoola, O. A. Ouyang, B. Saffell, J. and Jones, R. L.: Developing a relative humidity correction for low-cost sensors measuring ambient particulate matter. Sensors 18, 2790, https://doi.org/10.3390/s18092790 (2018).

Grover, B. D., Kleinman, M., Eatough, N. L., Eatough, D. J., Hopke, P. K., Long, R. W., Wilson, W. E., Meyer, M. B., and Ambs, J. L.: Measurement of total PM2.5 mass (nonvolatile plus semi-volatile) with the Filter

Dynamic Measurement System tapered element oscillating microbalance monitor, J. Geophys. Res. Atmos., 110(7), D07S03, doi:10.1029/2004JD004995, 2005

Jayaratne, R., Liu, X., Thai, P., Dunbabin, M., and Morawska, L.: The influence of humidity on the performance of a low-cost air particle mass sensor and the effect of atmospheric fog, Atmos. Meas. Tech., 11, 4883-4890, doi.org/10.5194/amt-11-4883-2018, 2018.

Lundgren, D.A.; Cooper, D.W. Effect of Humidity on Light-Scattering Methods of Measuring Particle Concentration. J. Air Pollut. Control Assoc., 19, 243–247, 1969.

Magi B. I., Cupini, C., Francis, J., Green, M. and Hauser C.: Evaluation of PM2.5 measured in an urban setting using a low-cost optical particle counter and a Federal Equivalent Method Beta Attenuation Monitor Aerosol Sci. Technol., 54,147-159, 2020.

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

Moravek, A., Murphy, J. G., Hrdina, A., Lin, J. C., Pennell, C., Franchin, A., Middlebrook, A. M., Fibiger, D. L., Womack, C. C., McDuffie, E. E., Martin, R., Moore, K., Baasandorj, M., and Brown, S. S.: Wintertime spatial distribution of ammonia and its emission sources in the Great Salt Lake region, Atmos. Chem. Phys., 19, 15691–15709, https://doi.org/10.5194/acp-19-15691-2019, 2019.

Morawska, L., Thai, P. K., Liu, X., Asumadu-Sakyi, A., Ayoko, G., Bartonova, A., Bedini, A., Chai, F. Christensen, B., and Dunbabin. M.: Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: how far have they gone?, Environ Int., 116, 286-99, doi: 10.1016/j.envint.2018.04.018, 2018.

PurpleAir, PurpleAir Map, air quality Map: http://map.purpleair.org/, last access: 1 August 2019.

Womack, C. C., McDuffie, E. E., Edwards, P. M., Bares, R., Gouw, J. A. A., Docherty, K. S., Dubé, W. P., Fibiger, D. L., Franchin, A., Gilman, J. B., Goldberger, L., Lee, B. H., Lin, J. C., Long, R., Middlebrook, A. M., Millet, D. B., Moravek, A., Murphy, J. G., Quinn, P. K., Riedel, T. P., Roberts, J. M., Thornton, J. A., Valin, L. C., Veres, P. R., Whitehill, A. R., Wild, R. J., Warneke, C., Yuan, B., Baasandorj, M., and Brown, S. S.: An Odd Oxygen Framework for Wintertime Ammonium Nitrate Aerosol Pollution in Urban Areas: NOx and VOC Control as Mitigation Strategies, Geophys. Res. Lett., 46, 4971–4979, https://doi.org/10.1029/2019GL082028, 2019.

Measurements of $PM_{2.5}$ with PurpleAir under atmospheric conditions

Karin Ardon-Dryer[1], Yuval Dryer[1], Jake N. Williams[1] and Nastaran Moghimi[2]
[1]Department of Geosciences, Atmospheric Science Group, Texas Tech University, TX
[2]Thomas S. Wootton High School, North Potomac, MD

Correspondence to: Karin Ardon-Dryer (karin.ardon-dryer@ttu.edu)

Abstract. The PurpleAir PA-II unit is a low-cost sensor for monitoring changes in the concentrations of Particulate Matter (PM) of various sizes. There are currently more than ~~9000~~ 10,000 PA-II units in use worldwide; some of the units~~m~~ are located in areas where no other reference air monitoring system is present. Previous studies have examined the performance of these PA-II units (or the sensors within them) in comparison to a ~~co-located~~co-located reference air monitoring system. However, because PA-II units are installed by PurpleAir customers, most of the PA-II units are not ~~co-located~~co-located with a reference air monitoring system and, in many cases, are not near one. This study ~~aimed~~ aims to examine how each PA-II unit~~s~~ performs under atmospheric conditions when exposed to a variety of pollutants and $PM_{2.5}$ concentrations (PM with an aerodynamic diameter smaller than 2.5 μm), when not co-located with a reference sensor. ~~We were interested in knowing how accurate these PA-II units are when measuring $PM_{2.5}$ concentrations with their sensitivity to concentration changes~~ We examine how PA-II units perform in comparison to other PA-II units ~~the~~ and Environmental Protection Agency (EPA) Air Quality Monitoring Stations (AQMSs) that are not ~~co-located~~co-located with them. For this study, we selected ~~eight~~ four different regions~~locations~~, ~~where~~ each ~~location contains~~ containing multiple PA-II units (minimum of seven per ~~location~~region~~, a total of 86 units) and at least one AQMS~~). ~~(total of 14)~~ In addition, each region needed to have at least one AQMS unit that was co-located with at least one PA-II unit, all units needed to be at a distance of up to 5 km from an AQMS unit and up to 10 km between each other. Correction of $PM_{2.5}$ values of the co-located PA-II units was implemented by multivariate linear regression (MLR), taking into account changes of temperature and relative humidity. The fit coefficients, received from the MLR, was then used to correct the $PM_{2.5}$ values in all the remaining PA-II units in the region. Hourly $PM_{2.5}$ measurements from each PA-II unit were compared to those from the AQMSs and other PA-II units in its region. The correction of the $PM_{2.5}$ values improved the R-squared (R^2), root mean square error (RMSE), and mean absolute error (MAE) and slope values between all units. ~~$PM_{2.5}$ measurements from each PA-II unit were compared to those from the AQMS and other PA-II units in its area. The comparisons were made based on hourly and daily $PM_{2.5}$ measurements.~~ In most cases, the AQMSs and the PA-II units were found to be in good agreement (75% of the comparisons had a $R^2 > 0.8$); they measured similar values and followed similar trends, that is, when the $PM_{2.5}$ values measured by the AQMSs increased or decreased, so did those of the PA-II units. In some high-pollution events, the corrected PA-II ~~measured~~ had slightly higher $PM_{2.5}$ values compared to those measured by the AQMS. Distance between the units did not impact the comparison between units. ~~We found PA-II $PM_{2.5}$ measurements to remain unaffected by changes in temperature or Relative Humidity (RH).~~ Overall, the PA-II unit, after corrections of $PM_{2.5}$ values, seems to be a promising tool for identifying relative changes in $PM_{2.5}$ concentration with the potential to complement sparsely distributed

monitoring stations and to aid in assessing and minimizing the public exposure to PM~~, particularly in areas lacking the presence of an AQMS~~.

1. Introduction

Atmospheric particulate matter (PM) with an aerodynamic diameter smaller than 2.5 μm ($PM_{2.5}$) is one of the leading contributors to the global burden of disease (GBD, Cohen et al., 2017; Forouzanfar et al., 2015; Lim et al., 2012). These particles are small enough to penetrate deep into the human lungs (Ling and van Eeden, 2009), where they have a negative impact on human health (Shiraiwa et al., 2017). Exposure to high $PM_{2.5}$ concentrations was found to be correlated with the daily number of hospitalizations and mortality cases (Schwartz et al., 1996; Klemm and Mason, 2000; Di et al., 2017). In the US, 3 ~~%~~ 5 % of annual deaths are attributed to $PM_{2.5}$ (Cohen et al., 2017). Determining the pollution-level $PM_{2.5}$ exposure can be challenging as a limited number of in-situ instruments are available for monitoring ground-level $PM_{2.5}$ concentrations (Ford et al., 2019).

In the United States, the Environmental Protection Agency (EPA) monitors ambient $PM_{2.5}$ concentrations by using ~~air~~ Air ~~quality~~ Quality ~~monitoring~~ Monitoring ~~stations~~ Stations (AQMSs). These stations use equipment that implements either a federal reference method or federal equivalent method (FRM and FEM, respectively; Clements et al., 2017). The FRM is a gravimetric measurement method in which particles are collected on a filter and the difference in filter weight before and after exposure is used to determine the 24-h PM concentration (Watson et al., 2017). The FEM measures PM using optical, beta ray attenuation and trapped element oscillation to provide hourly PM concentrations. A single FEM $PM_{2.5}$ sensor in each AQMS costs thousands of dollars. Further, the operation of these AQMSs requires trained personnel and significant infrastructure; they are subject to strict maintenance and calibration routines to ensure high-quality data and comparability between different locations (Castell et al., 2017). AQMSs generally have sparse geographic coverage and are located at fixed sites, mainly in large population centers; they are not present in smaller cities and underdeveloped regions. The high temporal and spatial resolution of $PM_{2.5}$ concentrations may vary significantly within a region, therefore, $PM_{2.5}$ concentration values provided by a single AQMS site may not accurately represent the $PM_{2.5}$ concentrations present near people who are concerned about their possible health effects (Wang et al., 2015). These limitations create a growing need for air quality sensor networks that ~~will~~ produce both temporal and spatial high-resolution pollution maps that can be used to identify peak events across large areas (Morawska et al., 2018).

Recent advancements in technology and a rise in public awareness have led to an increase in the popularity of low-cost air-quality sensors that are relatively cheap and easy ~~-~~ to ~~-~~ use (Commodore et al., 2017; Woodall al., 2017). Such sensors enable communities and individuals alike to obtain granular information on the spatial and temporal distribution of PM concentrations

in their area (Gupta et al., 2018; Morawska et al., 2018), thereby enabling them to monitor local ~~air~~ air-quality conditions (Williams et al., 2018). Many types of low-cost air-quality sensors are available, and they vary in performance (Williams et al., 2018); however, despite the proposed benefits of these sensors, their accuracy and precision remain unknown (Kuula et al., 2017). Data quality remains a major concern that hinders the widespread adoption of low-cost sensor technology. To assure data quality, it is important to test these sensors and compare them to FRM/FEM measurements under both laboratory and field conditions, particularly under atmospheric conditions with various air pollution levels in which the sensors are expected to operate (Kelly et al., 2017; Morawska et al., 2018). Testing these sensors at multiple locations will allow for exposure to different atmospheric conditions and pollutant types (AQ-SPEC, 2018).

Among the limitations of low-cost sensors are environmental factors that affect the sensor~~'~~s' abilities. Some low-cost sensors have exhibited sensitivity to temperature (T) and relative humidity (RH) (Clements et al., 2017). ~~When working in the~~In laboratory, these environmental conditions can be controlled; however, it is impossible to achieve such stability in the field under atmospheric conditions. Therefore, additional measurements under a variety of ambient conditions are needed (Kelly et al., 2017). In addition, some sensors have exhibited a drift in sensitivity over time (reduction of efficiency). The rate of drift over time is a crucial parameter in sensor characterization as it determines the interval of calibration as well as the overall useable lifetime of the sensor (Clements et al., 2017; Hagan et al., 2018).

The PA-II unit is a low-cost sensor sold by the company PurpleAir ~~company~~. It is meant for outdoor usage and is the subject of ~~our~~ this study. Each PA-II unit contains two Plantower particulate matter sensors (PMS5003 sensors) that provide real-time measurements of $PM_{1.0}$, $PM_{2.5}$, and PM_{10}. The usage of PA-II has grown rapidly in the last ~~two~~ few years, ~~with the result that~~to date more than ~~9000~~ 10,000 such sensors are in use across five continents, with the majority being operated in the US and Europe. PurpleAir provides live information on their website in the form of a color-coded air quality index (AQI) together with actual PM concentrations (PurpleAir, 2019). Several studies have already evaluated the PA-II unit or the sensors (PMS5003) ~~it~~ the unit contains; however, in all such studies, the PA-II unit (or the PMS5003 sensor) was ~~co-located~~co-located with a reference unit. The AQ Sensor Performance Evaluation Center (AQ-SPEC) evaluated the performance of a PA-II unit using FEM sensors as reference under laboratory and field conditions in the Los Angeles area. Their evaluation showed a very good comparison between the two for both $PM_{2.5}$ and PM_{10} (AQ-SPEC, 2018). An additional comparison between three different PA-II sensors and a single FEM was performed for eight weeks between December 2016 and January 2017 at the South Coast Air Quality Management District Rubidoux Air Monitoring Station. Good correlation ($R^2 > 0.9$) was found between the three PA-II units and the FEM unit. However, although the PA-II unit follows diurnal and day-to-day fluctuations very well, it consistently overestimated the $PM_{2.5}$ concentrations measured by the FEM (Gupta et al., 2018). Sayahi et al. (2019) conducted a long-term comparison (320 days) between two PMS5003 sensors and both FRM and FEM units that were all ~~co-located~~co-located ~~at~~ in Salt Lake City, Utah. One of their PMS5003 sensors overestimated the $PM_{2.5}$ concentration whereas the other measured similar values to those measured by the FEM. According to Gupta et al. (2018), the performance

of PA-II compared against FEM units in a high-pollution environment ($PM_{2.5}$ > 100 µg m^{-3}) is unknown and requires further evaluation. ~~In addition, the sensitivity of the PA-II sensors to changes in RH, temperature, and other environmental parameters remains a topic of further investigation (Gupta et al. 2018). Answers to these questions~~

Multivariate linear regression (MLR) models with T and RH have been widely used to calibrate the PA-II sensors against co-located reference monitors, which help improve the accuracy of the PA-II units (Bi et al., 2020; Magi et al., 2020). Magi et al. (2020) performed a comparison of multiple co-located PA-II units with an FEM unit using a MLR that used measurements of $PM_{2.5}$ (using FEM unit), RH, and T as predictors to model the correct PA-II $PM_{2.5}$ values up to 50 µg m^{-3}. They concluded that the PA-II is suitable for air quality, health, and urban aerosol research. Bi et al. (2020) matched a PA-II unit to its nearest AQMS unit within a 500 m radius; they found that co-located pairs were robust within a range 100 to 1,000 m. Most of these studies so far focused on co-located units or units that were up to 1 km from the reference unit. But in reality, most PA-II units are not near any reference unit; many are positioned more than 1 km away. Several questions can be raised based on that: Can MLR of co-located units be used to improve the accuracy of the measurements taken by PA-II units that are further away from the AQMS unit? Can MLR of multiple regions be used to compensate for the lack of a co-located pair of a neighboring region? Such usage of PA-II units ~~-~~at various distances ~~are~~ is crucial if we are to assess the possibility of using measurement data from multiple PA-II units to properly represent the air quality of an area, thus allowing the residents to protect themselves when high pollution events occur.

This study ~~aimed~~ aims to examine how PA-II units perform under atmospheric conditions when exposed to a variety of pollutants and $PM_{2.5}$ concentrations. For the scope of this study, we chose to focus only on regions that contain at least one pair of co-located PA-II and AQMS units. Corrections of $PM_{2.5}$ values for co-located PA-II and AQMS units, based on MLR, were performed and applied to all the other PA-II units in that region. Comparison of ~~PA-II units to~~ $PM_{2.5}$ measurements taken by all units in each region, ~~an~~ AQMSs ~~that was no~~and PA-II units~~t co-located~~ (when $PM_{2.5}$ values were measured or corrected) are presented ~~with them are presented~~. The presented comparisons were done for both the entire study period and for specific events that we wanted to examine in greater detail. ~~Further, a comparison of PA-II units to other nearby PA-II units and their efficiency as a network of low-cost sensors are discussed.~~

2. Method

2.1. PurpleAir PA-II Unit Structure and Data

The PurpleAir PA-II unit ~~has size of~~is 85 × 125 mm in size. It contains two PMS5003 sensors (see two blue rectangles in Fig. 1A), a BME280 environmental sensor, and an ESP8266 microcontroller. The BME280 sensor is used to monitor the units' inner pressure, temperature, and humidity; the sensor measurements are not to be used for monitoring ambient conditions (PurpleAir, personal communication, 2019). The ESP8266 microcontroller is used to communicate with both the two PMS5003 sensors and with the PurpleAir server over Wi-Fi, thereby allowing the PM concentration to be presented live on the PurpleAir map (https://www.purpleair.com/map). The PMS5003 sensors provide real-time measurements of $PM_{1.0}$, $PM_{2.5}$, and PM_{10} concentrations; the sensors are based on the light scattering principle, and a photodiode detector converts the scattered light to a voltage pulse. A fan draws the particles into the sensor and past the laser path (Fig. 1B) at a flow rate of 0.1 L/min. The particle count is calculated by counting the pulses from the scattering signal and converting the number of pulses to a mass concentration for six diameters between 0.3 and 10 μm using an algorithm for outdoor PM (CF_ATM - average particle density). Each PMS5003 sensor has an effective measurement range for $PM_{2.5}$ concentration of 0~~–~~500 μg m^{-3} with a resolution of 1 μg m^{-3}, and the maximum standard $PM_{2.5}$ concentration is above 1000 μg m^{-3} According to the manufacturer, each PMS5003 sensor will work effectively in a ~~temperature~~ T range of -10 °C to 60 °C and RH range of 0~~ %~~ 99 % (Yong, 2016).

The microcontroller in the PA-II unit reads the $PM_{1.0}$, $PM_{2.5}$, and PM_{10} concentrations from the PMS5003 sensors every second; it averages the concentration values across 20 s and displays the results using UTC time (PurpleAir, personal communication, 2019). The use of a dual PMS5003 sensor setup serves as an internal check for the PA-II unit's integrity. The similarity/difference in the PM concentrations obtained from the two PMS5003 sensors (named ~~as~~ A and B) allows users to evaluate the efficiency and validity of their PA-II unit. The two PMS5003 sensors, A and B, should agree with each other at all ~~the~~ times; failure to report the same value indicates that something is wrong with one of the sensors. PurpleAir does not calibrate the units~~ir devices~~; instead, before each PA-II unit is sent out to a customer, the company performs a comparison test with a dozen other PA-II units to find and remove outliers from the shipment (PurpleAir, personal communication, 2019).

All the data regarding the PA-II units and their measurements was downloaded from the PurpleAir website. Information about all the PA-II units was downloaded in a JSON formatted file. Each PA-II unit has a name (given by the owner), a unique ID number (designated by the company for each sensor), the unit location (latitude and longitude), and ~~a~~ the date on which the unit was installed. We initially selected all the PA-II units that were active between January 1, 2017, and December 31, 2018 (UTC time). For each selected PA-II unit, we downloaded an Excel file containing the measurement data in 20-s intervals for both PMS5003 sensors (A and B). Because our focus was on $PM_{2.5}$ measurements, we calculated the $PM_{2.5}$ hourly average and standard deviation (SD) based on the original measurement values and the daily average and standard deviation based on hourly averages that we had calculated previously. Our final dataset included only days that had a minimum of ~~20~~ 13 h of measurements per day (~~80~~ >50 % of the day). Only times ~~which~~ that had a good agreement (R^2 > 0.9) of hourly $PM_{2.5}$ measurements between the two PMS5003 sensors (A and B) were used.

2.2. $PM_{2.5}$ Measurements from AQMS

Hourly measurements of $PM_{2.5}$ (FRM/FEM Mass code - 88101 file) from all AQMSs collected by the EPA from January 1, 2017, to December 31, 2018, were selected from the EPA website (https://aqs.epa.gov/api). The location of each AQMS was provided in the same file. Each AQMS is identified by the combination of state code, county code, site number, and Parameter Occurrence Code (POC) number. The POC is used to represent cases in which more than one unit performs $PM_{2.5}$ measurements at the same site. All timestamps were converted to UTC to match the PA-II measurement timestamps. The $PM_{2.5}$ daily average and standard deviation were calculated based on the hourly $PM_{2.5}$ measurements; only days with a minimum of ~~20~~ 13 h of measurements per day (~~80~~ >50 % of the day) were considered.

2.3. Identification of Locations for Analysis - Areas with Multiple PA-II units and at least one AQMS

By using the JSON file for the PA-II units and the 88101 file for the AQMS, we calculated the distances between all the units ~~was calculated~~ to identify ~~locations~~ regions with multiple PA-II units (a minimum of five units) and at least one AQMS. At least one AQMS unit needed to be at a distance up to 1.1 km from at least one PA-II unit (defined as a *co-located pair*, a similar range used by Bi et al., 2020). All the units in these regions~~locations~~ needed to be active during the designated time period of January 1, 2017, to December 31, 2018. In each region PA-II units needed to be less than 5 km from at least one AQMS unit and up to 10 km from each other. ~~Eight~~ Four different ~~locations~~ regions containing a total of ~~14~~ seven different AQMSs (all FEM type) and ~~86~~ 46 different PA-II units were identified: ~~Pittsburgh, PA;~~ Denver, CO; ~~Berkeley-Oakland, CA;~~ San Francisco, CA; Vallejo, ~~CA; Ogden-South Ogden, UT; Lindon-Orem, UT;~~ and Salt Lake City, UT. Fig~~.~~ ure ~~S1~~ 2 shows a map with all the PA-II units and AQMSs at each ~~location~~region. Table S1 provides information on each of the four regions with the names of the units, their locations, first and last times of measurement, and the number of hours measured by each unit. For simplifications, each region was defined by two letters to represent its name (DE for Denver, SF for San Francisco, VA for Vallejo, and SL for Salt Lake City). Also, each unit type received a two letter code (AQ for AQMS and PA for PA-II). Each unit received a number instead of an ID, as shown in Table S1. More than 50% of the units were at a distance of 4 km from each other. The highest distance between two PA-II units (9.2 km) was in SL. Table S2 lists the distance between each unit per region. The number of concurrent hourly measurements of PA-II units and AQMS units in each comparison varies per region. Overall, the number of concurrent hourly measurements ranged from 95 to 16,658 h with an average of 6,412 ± 2,924 h. Table S2 lists the number of concurrent $PM_{2.5}$ hourly measurements between all units in each of the regions.

To evaluate the similarities and differences between the PA-II units and the AQMSs and other PA-II units, a set of calculations and comparisons was performed using Matlab and Excel. R-squared (R^2), root mean square error (RMSE) and mean absolute error (MAE) values as well as the best fit information, including the slope, were used for the comparison.

~~Table 1 provides information on each of the eight locations with the names of the units, their location, first and last time of measurement, and the minimum and maximum $PM_{2.5}$ hourly values.~~

~~In Pittsburgh, two AQMSs (42-3-8-3 and 42-3-1376-1) and eleven PA-II units (ID=3723, 3981, 9016, 9026, 9038, 9096, 9878, 9880, 9892, 9896, and 9906) were used. In Denver, three AQMS (8-31-26-3, 8-31-27-3, and 8-31-28-3) and eight PA-II units (ID=2249, 2267, 2269, 2719, 2900, 3924, 4022, and 7956) were used. In in Berkeley-Oakland, three AQMSs (6-1-11-3, 6-1-12-3, and 6-1-13-3) and ten PA-II units (ID=2574, 3082, 3854, 4335, 4506, 4795, 4825, 5414, 6410, and 10114) were used. San Francisco, Vallejo, Ogden-South Ogden, and Lindon-Orem all had a single AQMS (6-75-5-3, 6-95-4-4, 49-57-2-5, and 49-49-4001-5, respectively) but multiple PA-II units. San Francisco had nine PA-II units (ID=1226, 2031, 2910, 3348, 3996, 4372, 4770, 5776, and 6344); Vallejo had 15 units (the maximum; ID=1142, 1870, 1874, 1878, 1882, 2480, 2906, 3686, 3758, 3769, 3782, 3784, 3960, 4928, and 5127); Ogden-South Ogden had seven PA-II units (the minimum; ID=465, 1104, 5178, 5454, 6604, 7858, and 7860); and Lindon-Orem had 12 PA-II units (ID=5135, 5143, 5145, 5728, 5732, 5736, 5750, 5754, 5760, 6304, 6948, and 6986). Salt Lake City had two AQMSs at the same location (49-35-3006-4 and 49-35-3006-5, different POCs) and 14 PA-II units (ID=884, 3388, 5014, 5460, 5742, 5802, 5990, 6078, 6356, 6360, 6434, 6608, 6622, and 10050).~~

2.~~5~~4. Meteorological Information

Meteorological measurements including ~~temperature~~T, RH, and wind speed/direction were used from the EPA website (https://www.epa.gov/outdoor-air-quality-data). Only a ~~some~~few AQMSs had these meteorological measurements: ~~42-3-1376-1 and 42-3-8-3 from Pittsburgh,~~ DE-AQ-1, and DE-AQ-3 in ~~8-31-26-3 and 8-31-28-3 from~~ Denver, ~~49-57-2-5 from Ogden-South Ogden, 49-49-4001-5 from Lindon-Orem,~~ and SL-AQ-1~~49-35-3006-4~~ from Salt Lake City. Additional meteorological measurements such as ~~temperature~~T, RH, wind speed and gust, wind direction, and visibility of different meteorological stations were obtained from the Iowa Environmental Mesonet website (https://mesonet.agron.iastate.edu/request/download.phtml). For meteorological information about the selected ~~locations~~regions, the following meteorological stations were used: ~~AGC Pittsburgh/ Allegheny station in Pittsburgh, the~~ Denver International Airport (DEN) station in Denver, ~~the Ogden Hinckley Muni (OGD) station in Utah, the Provo Muni (PVU) station in Ogden-South Ogden,~~ the Salt Lake City International ~~airport~~ Airport (SLC) station ~~in Lindon-Orem, the California Oakland (OAK) station in Berkeley-Oakland and~~ San Francisco International Airport (SFO) in San Francisco, and the Napa County (APC) station ~~in~~ for Vallejo.

2.5. Remove of outlier PA-II units and irregular hours

The first step was to identify outliers among the PA-II units, per region, meaning PA-II units that behave differently from the other PA-II units in their region. By comparing R^2 between the $PM_{2.5}$ values measured by each pair of PA-II units, using a linear regression, we identified the outlier units. A PA-II unit that did not have an $R^2 \geq 0.75$ with at least 75% of the other PA-II units in its region was considered an outlier unit, and therefore was removed from future analysis (Fig. S1 shows a comparison for each of the four regions). Only one unit from SF (SF-PA-9, see Fig. S1B) had very low R^2 when compared to all other PA-II units. Most PA-II units had high R^2 values (>0.9) with the other units. Irregular $PM_{2.5}$ hourly measurements were removed from all units (PA-II and AQMS). These irregular hourly measurements were identified as a large single hourly increase of $PM_{2.5}$ values (>70 $\mu g\ m^{-3}$) that was not measured by any other unit in the region. Such a large increase was caused most likely by a local source near a specific unit, such as a small-scale fire, lawn mower, barbecue, cigarette smoke, or fireworks (Zheng et al., 2018), and attributed to the location of many of the PA-II units in a residential area. Firework events were removed, as they were very localized events and were measured by a single unit. Overall, less than 0.03% of the hourly $PM_{2.5}$ measurements identified as irregular hours were removed from different PA-II and AQMS units.

~~**2.4. Comparison between PA-II and AQMS**~~

~~To evaluate the similarities and differences between the AQMS and the PA-II units, a set of calculations and comparisons was performed. First, graphs showing the distribution of $PM_{2.5}$ values were plotted. Second, a regression~~ ~~We performed different comparisons for both the entire study period and for specific events that we wanted to examine in greater detail.~~ ~~between the AQMS and each PA-II unit was made based on hourly and daily $PM_{2.5}$ measurements. From the regression, R-squared (R^2) and root mean square error (RMSE) values as well as the best fit information, including the slope and intercept, were obtained.~~ ~~We performed different comparisons for both the entire study period and for specific events that we wanted to examine in greater detail.~~

~~**2.6. AQI Calculations**~~

~~The AQI is used for the reporting air quality levels. It allows the public to know how clean the air is and indicates the health effects a person may experience within a few hours or days of breathing unhealthy air. The AQI has six categories, each of which corresponds to a different level of health concern (EPA, 2014): Good (0–50, green), Moderate (51–100, yellow), Unhealthy for Sensitive Groups (101–150, orange), Unhealthy (151–200, red), Very Unhealthy (201–300, purple), and Hazardous (301–500, maroon) (see Table S1). In our study, we calculated the AQI for $PM_{2.5}$ daily average as follows:~~

~~$$AQI = \frac{(\text{measured } PM_{2.5} - PM_{min})(AQI_{max} - AQI_{min})}{(PM_{max} - PM_{min})} + AQI_{min} \quad (1)$$~~

~~where the measured $PM_{2.5}$ is the daily average $PM_{2.5}$ value, PM_{max} and PM_{min} are respectively the maximum and minimum concentration of the AQI color category for the measured $PM_{2.5}$, AQI_{max} is the maximum AQI value for a color category that corresponds to the measured $PM_{2.5}$, and AQI_{min} is the minimum AQI value for a color category that corresponds to the measured $PM_{2.5}$. Table S1 lists the different values and categories of PM_{max}, PM_{min}, AQI_{max}, and AQI_{min}.~~

3. Results and Discussion

3.1. Hourly and Daily ~~$PM_{2.5}$ Comparisons~~measurements from ~~of~~ AQMS and PA-II units.

This study examined measurements ~~for~~ from a two-year period from January 1, 2017, to December 31, 2018, resulting in ample overlapping measurement times between the different PA-II units and ~~the~~ different AQMSs. ~~The number of concurrent hourly measurements in each comparison varies per location. Overall, the number of concurrent hourly measurements ranged from 1017 to 13975 h with an average of 6652 ± 2822 h per comparison. Other than the Lindon-Orem area where the local AQMS was active only from November 2017, measurements from January 2017 were available in all the other areas.~~ Most of the PA-II units became active only at the end of 2017. ~~The distance between the different AQMSs and PA-II units ranged from 0.01 km to 13 km with an average of 4.2 ± 2.4 km. Table 2 lists the exact distance and number of $PM_{2.5}$ hourly measurements used in comparisons of each AQMS and PA-II unit. Based on the overlap times, we identified and examined the distribution of daily $PM_{2.5}$ values measured by the PA-II units and AQMS for each location and also performed additional comparisons between the units in these locations.~~ The frequency of hourly $PM_{2.5}$ measurements from PA-II units and AQMSs as well as measurements of RH and T, during the study period, in each region were observed to understand the conditions each region had (shown in Fig. S2). Some regions had high frequency of hourly measurements at low RH (30 - 40 %), while others had high RH (>90 %). Most of the measurements were performed under T of 5 - 20 °C. All regions had a high frequency of $PM_{2.5}$ between 10 - 20 µg m^{-3} for both PA-IIs and AQMSs.

~~3.1.1 Distribution of Daily $PM_{2.5}$ Values~~

~~Fig. 2 shows the~~ Time series ~~distribution~~ of daily $PM_{2.5}$ values for each unit at each of the ~~eight~~ four ~~locations~~regions are presented in Fig. 2. Overall, the daily $PM_{2.5}$ values obtained from both the AQMSs and the PA-II units seem to follow similar trends. When the AQMS values increase~~/~~ or decrease, the PA-II values also increase~~/~~ or decrease. The PA-II unit measurements of daily $PM_{2.5}$ values start at 0 µg m^{-3}, and the AQMS can measure negative values owing to its calibration process. In some cases ~~(locations and times)~~, the AQMS measured higher $PM_{2.5}$ daily values compared to the PA-II units, mainly at days with low $PM_{2.5}$ values, as seen in April - June 2018 in Vallejo (Fig. 3C) and Salt Lake City (Fig. 3D). These differences were observed mainly in days with low RH values, and low $PM_{2.5}$ daily values (Fig. S3).~~, as seen during April–July 2018 in Berkeley–~~

~~Oakland (Fig. 2C), Lindon-Orem (Fig. 2G), and Salt Lake City (Fig. 2H)~~. However, overall, regardless of the $PM_{2.5}$ concentration, the PA-II units usually measured higher values compared to those measured by the AQMSs (see July and August 2018 in ~~Pittsburgh~~Denver, Fig. ~~2A~~3A, and November 2018 in San Francisco and Vallejo, Fig. 3B and Fig. 3C). This overestimating of PM values by the PA-II units (or PMS sensors) compared to FRM and FEM units has also been observed in previous studies ~~previously~~ (Kelly et al., 2017; AQ-SPEC, 2018; Gupta et al., 2018; Sayahi et al., 2019) when the two units were ~~co-located~~co-located.

The overestimating raises questions about the accuracy of the PA-II units. According to PurpleAir (PurpleAir, personal communication, 2019) the company does not calibrate the PA-II units; instead, before each PA-II unit is sent out to a customer, the company performs a comparison test with a dozen PA-II units to find and remove outliers from the shipment (PurpleAir, personal communication, 2019). Previous studies suggested that part of the problem with the PA-II unit results from the optical particle counter being impacted by changes of RH (Crilley et al., 2018; Malings et al., 2020; Magi et al., 2020). Water vapor can condense on aerosol particles, making them grow hygroscopically under high RH conditions (Lundgren and Cooper, 1969). The PA-II units do not have any heater or dryer at their inlets to remove water from the sample before measuring the particles; therefore, deliquescent or hygroscopic growth of particles, mainly under high RH conditions, can lead to higher reported PM concentrations (Di Antonio, 2018; Jayaratne et al., 2018; Bi et al., 2020), which ends as an overestimate of the PM compared to the reference units. Weather conditions can impact the values reported by low-cost sensors (Morawska et al., 2018). Changes in T or RH have been found to affect the performance of the PA-II units, especially under atmospheric conditions, as they cannot be controlled (Bi et al., 2020). Therefore, MLR between a PA-II, and an AQMS, which also considers changes of T and RH, can help correct the reported $PM_{2.5}$ values of the co-located PA-II units. Similar corrections have been suggested and implemented in other locations with PA-II units (Bi et al., 2020; Magi et al., 2020) and other low-cost sensors (Malings et al., 2020). Most of these studies focus on co-located units or on units that were up to 1 km from the reference unit.

3.2. Correction of PA-II $PM_{2.5}$ hourly values using a Multivariate Liner Regression

Seven PA-II units were co-located with at least one AQMS unit. In Denver, three PA-II units were co-located with AQMS units. The closest PA-II unit was DE-PA-6, which was only 5.8 m from DE-AQ-3. Unit DE-PA-8 was 30 m from DE-AQ-2, while DE-PA-2 was 79 m from DE-AQ-3. In San Francisco only one PA-II unit was co-located with the AQMS unit. SF-PA-1 was 400 m from SF-AQ-1. In Salt Lake City the two co-located AQMS units (SL-AQ-1 and SL-AQ-2) were 874 m from SL-PA-13. Vallejo unit VA-PA-2 was 1.06 km from VA-AQ-1.

Calculations of the ratio between the measured $PM_{2.5}$ from the PA-II to the AQMS as a function of T and RH, known as a hunidogram, were performed (Fig. S4). Some of the PA-II units seem to be impacted by T and RH more than others; these

units also had relatively low R^2 values with the AQMS unit, as in the case of DE-PA-6 in Denver (Fig. S4A). Only co-located pairs with $R^2 > 0.65$ were used, reducing the co-located pairs to six. The fact that not all units seem to be impacted in a similar way by the changes of T and RH can explain parts of the debate that exists in the literature. For example, Sayahi et al. (2019) found very low correlation values between measurements from the PMS5003 sensor (used in PA-II) to T and RH under atmospheric conditions. Holstius et al. (2014) found a negligible effect of T or RH on measurements performed using low-cost sensors under ambient conditions. However, several studies that used old PMS units such as PMS1003 which was used in PA-I, or PMS3003 which was never used in any PA units found that these sensors were affected by RH (Kelly et al., 2017; Jayaratne et al., 2018; Zheng et al., 2018). AQ-SPEC (2018) tested the PA-II unit in a laboratory setting under different RH conditions and found that most RH combinations had a minimal effect on the PA-II's precision. On the other hand, Magi et al. (2020) found an impact of T and RH conditions on the PA-II $PM_{2.5}$ measurements in atmospheric conditions. Therefore, consideration of T and RH was used in the MLR.

An MLR following Magi et al. (2020) was performed on each co-located PA-II and AQMS pair, including meteorological measurements (T and RH). Based on the MLR, the multivariable linear dependence of PA-II $PM_{2.5}$ on AQMS, RH and T created the predictors of PA-II as:

$$PA\text{-}II(PM_{2.5}) = A_1 + A_2 AQMS(PM_{2.5}) + A_3 T + A_4 RH \quad (1)$$

where A_1, A_2, A_3, and A_4 fit coefficients received from the MLR, PA-II ($PM_{2.5}$) and AQMS($PM_{2.5}$) are in units of μg m^{-3}, T is in Celsius, and RH is in percentage. Based on these parameters and fit coefficients, a calculation of the corrected PA-II $PM_{2.5}$ hourly values for each PA-II was performed using the following:

$$PA\text{-}II(PM_{2.5}), corrected = \frac{PA-II(PM_{2.5}), uncorected - A_1 - A_3 T - A_4 RH}{A_2} \quad (2)$$

Details of the coefficients received in the MLR as well as the regression output including R^2, RMSE, MAE, and slope for each correction of $PM_{2.5}$ values in the PA-II units, for each region, can be found Table 1. Figure 4 presents a comparison of the $PM_{2.5}$ values from the uncorrected PA-II unit to the AQMS as well as the PA-II $PM_{2.5}$ values hourly after correction, per region.

3.2.1. Correction of $PM_{2.5}$ values in co-located PA-II unit per region

San Francisco had only one co-located PA-II unit (SF-PA-1) with the single AQMS unit (SF-AQ-1). There were 9,910 h of $PM_{2.5}$ measurements overlapping from both units. The hourly $PM_{2.5}$ measurements of SF-PA-1 before correction ranged from 0.1 up to 263.8 μg m^{-3}, while SF-AQ-1 ranged from -10 up to 241 μg m^{-3}. Meteorological measurements from SFO meteorological station, located 16 km from the two units, were used. The meteorological measurements ranged from 0.6 to 33 °C for T and 9.2 up to100 % for RH. The MLR improved the comparison between SF-AQ-1 and SF-PA-1, as shown in Fig 4A. While there was no change in the R^2 values (0.91), the RMSE and MAE values improved. RMSE decreased from 7.3 μg

m^{-3} before the $PM_{2.5}$ value corrections to 5.8 µg m^{-3} after the $PM_{2.5}$ value corrections, while MAE changed from 5.4 to 4.3 µg m^{-3}. The slope changed with the MLR from 1.3 to 1.0.

Vallejo had one co-located PA-II unit (VA-PA-2) that had 11,506 h of overlapping $PM_{2.5}$ measurements with VA-AQ-1. The uncorrected $PM_{2.5}$ measurements from the PA-II unit ranged from 0 up to 468.5 µg m^{-3}, while AQMS $PM_{2.5}$ measurements ranged from -10 up to 435 µg m^{-3}. The meteorological station that was used for this PA-II $PM_{2.5}$ values correction (APC) was 11 km away from the AQMS. The meteorological measurements during this comparison ranged from -5 to 41 °C and 5.8 up to 100 % for T and RH, respectively. The MLR improved the comparison between the PA-II and the AQMS (Fig. 4B). There was no change in the R^2 values, which was 0.91, yet RMSE and MAE values decreased. RMSE decreased from 8.0 µg m^{-3} to 6.0 µg m^{-3}, while MAE decreased from 5.4 µg m^{-3}, before the $PM_{2.5}$ values corrections, to 4.0 µg m^{-3} , after the $PM_{2.5}$ value corrections. The slope also improved from 1.3 to 1.0.

Denver had two different PA-II units that were co-located with two different AQMS units. Unit DE-PA-8 had 2,134 h of overlapping $PM_{2.5}$ measurements with DE-AQ-2, while DE-PA-2 had 6,800 h of overlapping $PM_{2.5}$ measurements with DE-AQ-3. The range of the $PM_{2.5}$ values were similar for both PA-II units. DE-PA-2 ranged from 0.1 up to 76.9 µg m^{-3}, while DE-PA-8 ranged from 0.1 up to 78.5. The AQMS units also had relatively similar $PM_{2.5}$ measurements, which ranged from 0.3 up to 57.9 µg m^{-3} for DE-AQ-2 and from 0.9 up to 46.6 µg m^{-3} for DE-AQ-3. Both AQMS units had meteorological measurements as part of the AQMS units that were used for the MLR. Temperature measurements for DE-PA-2 ranged from -9.4 up to 39.4 °C, while DE-PA-8 T measurements ranged from -5.6 up to 34.4 °C. RH measurements were very similar as well, ranging from 3 % for DE-PA-2 and 4% for DE-PA-3 and up to 98 % for both. Although there were similar ranges of $PM_{2.5}$, T and RH measurements were taken, and there were differences between the PA-II's comparison to their co-located AQMS units. DE-PA-2 had better correlation values before and after the MLR (R^2 of 0.75 and 0.78, respectively) compared to DE-PA-8 (R^2 of 0.68 and 0.69, respectively, Fig 4D). RMSE and MAE values for both cases improved by more than 1.1 µg m^{-3} for the RMSE and 0.8 µg m^{-3} for the MAE. The slope values, which were 1.7 and 1.3 before the MLR, reduced to 1.0 in both cases. While the DE-PA-2 with DE-AQ-3 had higher R^2 values, it also had higher RMSE and MAE values compared to the DE-PA-8 and DE-AQ-1 pair. The two co-located pairs were combined and compared before and after the MLR (Fig. 4E). R^2, RMSE, and MAE values were in the same range in the two separate comparisons. R^2 improved from 0.71 to 0.72 before and after the MLR respectively. RMSE changed from 5.4 to 3.8 µg m^{-3}, and MAE changed from 3.7 to 2.6 µg m^{-3}. The value of the slope also improved from 1.4 to 1.0 after the correction of PA-II $PM_{2.5}$ values. The combined MLR had a higher number of observations and R^2, RMSE, and MAE values that were in the range of each of the separate comparisons.

The last region with co-located units was Salt Lake City. In this region one PA-II unit (SL-AP-13) was co-located with two AQMS units that were in the same location (SL-AQ-1 and SL-AQ-2). Measurements of T and RH were used from the SL-AQ-1 meteorological station. SL-AQ-1 had 6,216 hours of overlapping $PM_{2.5}$ hourly measurements with SL-AP-13, while SL-

AQ-2 had slightly more overlapping measurements (6,409 h). The meteorological parameters during these comparisons ranged from -7.2 to 38.3 °C for T and 2 up to 91% for RH. The uncorrected $PM_{2.5}$ measurements from the PA-II unit ranged from 0 up to 128.5 µg m^{-3}, while the $PM_{2.5}$ measurements ranged from 0.1 up to 87.5 µg m^{-3} for SL-AQ-1 and 0.1 up to 89.1 µg m^{-3} for SL-AQ-2. We first evaluated the MLR for each of the AQMS units separately (Fig. 4F-G). Different R^2, RMSE, and MAE values were obtained. While both showed an improvement of the RMSE, MAE, and slope value after the $PM_{2.5}$ values corrections, the MLR with SL-AQ-1 had better results with a higher R^2 (0.88 compared to 0.78) and, lower RMSE (2.7 µg m^{-3} compared to 3.7 µg m^{-3}) and MAE (1.8 µg m^{-3} compared to 2.4 µg m^{-3}) values. Combining the hourly $PM_{2.5}$ values from the two AQMSs together, since both were in the same location, was performed by averaging the AQMS $PM_{2.5}$ values (Fig. 4H). The MLR results showed an increase of R^2 and a decrease of RMSE and MAE values. Averaging of the AQMS units provided lower RMSE and MAE values and higher R^2 values compared to one of the separate options. This MLR had low RMSE and MAE values (3.0 µg m^{-3} and 2.0 µg m^{-3}, respectively) and high R^2 (0.84), making this MLR better than the one used by the pair SL-PA-13 with SL-AQ-2.

3.2.2. Corrections of PA-II $PM_{2.5}$ values of other PA-II units per region based on MLR

Based on the different coefficient values received in each MLR (Table 1), we implemented Eq. 2 on each of the uncorrected PA-II units' $PM_{2.5}$ hourly values, using the same meteorological parameter used for the MLR corrections. San Francisco and Vallejo each only had one set of comparisons and coefficients, while there were several options for Denver and Salt Lake City. The new $PM_{2.5}$ hourly values (corrected) from each PA-II unit were compared to the nearest AQMS unit and to all the other PA-IIs in the region using a linear regression. Corrected $PM_{2.5}$ hourly values of PA-II unit measurements improved the comparison between the other PA-IIs and AQMS units, as shown by the general reduction of RMSE, MAE, and the slope (Fig. 5 shows comparison to AQMSs. See Table S3 for full comparison of all comparisons).

The comparison of each unit in San Francisco and Vallejo before and after implementing the PA-II $PM_{2.5}$ value corrections did not change the R^2 values between the PA-IIs and AQMS units (Table S3A, and Table S3B). The average R^2 value between the PA-IIs and AQMS units for Vallejo was 0.79 ± 0.13, while San Francisco's average R^2 value was 0.83 ± 0.11. No changes were observed among the comparison of the PA-IIs themselves (0.93 ± 0.06 for San Francisco and 0.89 ± 0.07 for Vallejo). However, reductions in RMSE and, MAE values were observed in both regions (see Fig. 5A for San Francisco and Fig. 5A Vallejo).

The average RMSE values for San Francisco for the PA-IIs with the AQMSs changed from 8.23 ± 0.66 µg m^{-3} to 6.53 ± 0.54 µg m^{-3}. Similar improvements were observed when the PA-II units were compared to the other PA-II units; average RMSE changed from 4.23 ± 1.05 µg m^{-3} to 3.39 ± 0.83 µg m^{-3}. Similar reductions were also observed for the MAE; average MAE values for the PA-IIs with the AQMSs changed from 5.84 ± 0.31 µg m^{-3} to 4.61 ± 0.26 µg m^{-3}. Even when the PA-IIs were

compared to the other PA-II units, a reduction in MAE was observed; the average MAE changed from 2.16 ± 0.35 μg m^{-3} to 1.71 ± 0.28 μg m^{-3}. A reduction was also observed in the average slope value (Table S3A).

Vallejo also had a reduction in RMSE and MAE values after the MLR. The average RMSE values for the PA-II with the AQMS changed from 8.95 ± 1.35 μg m^{-3} to 6.73 ± 1.05 μg m^{-3}. Similar improvements were observed when the PA-II units were compared to the other PA-II units. RMSE changed from 5.14 ± 1.48 μg m^{-3} before the $PM_{2.5}$ value corrections to 3.89 ± 1.12 μg m^{-3} after the PA-II $PM_{2.5}$ value corrections. Similar reductions were also observed for the MAE; average MAE values for the PA-II with the AQMS changed from 5.81 ± 0.61 μg m^{-3} to 4.33 ± 0.46 μg m^{-3}. Even when the PA-IIs were compared to the other PA-II units, a reduction in MAE was observed; the average MAE changed from 2.5 ± 0.56 μg m^{-3} to 1.89 ± 0.42 μg m^{-3}. A reduction was also observed in the average slope value (Table S3B).

In Denver three different correction options were evaluated based on two separate pairs of PA-IIs with AQMSs as well as a combination of both together. Since there were several AQMS units, each PA-II unit was compared to its nearest AQMS unit (see Table S2A for distances, the distances ranged from 0 to 4 km). The coefficient values were different between each MLR option (Table 1). The R^2, RMSE, MAE, and slope were very similar (Fig. 5C-E). No change in R^2 was observed in each MLR type. A reduction in RMSE was observed after the PA-II $PM_{2.5}$ values were corrected. The average RMSE value, between each PA-II to the nearest AQMS unit, before the correction, was 5.7 ± 0.8 μg m^{-3}. All three correction types had lower average RMSE values. Correction of PA-II $PM_{2.5}$ values based on DE-PA-2 with DE-AQ-3 had an average RMSE value of 3.6 ± 0.3 μg m^{-3}, lower than the average RMSE from the corrections that were based on DE-PA-8 with DE-AQ-1 (4.3 ± 0.5 μg m^{-3}). The correction that combined both units had an average RMSE in the range of the two correction options (4.0 ± 0.5 μg m^{-3}). Similar reduction trends were observed for the MAE values. The average MAE values between each PA-II to the nearest AQMS unit, before the PA-II $PM_{2.5}$ value corrections, was 3.8 ± 0.6 μg m^{-3}. Correction of PA-II $PM_{2.5}$ values based on DE-PA-2 with DE-AQ-3 had an average MAE value of 2.4 ± 0.2 μg m^{-3}, lower than the average MAE based on DE-PA-8 with DE-AQ-1 (2.8 ± 0.4 μg m^{-3}). Yet the combined option had MAE in the range of the other two (average of 2.6 ± 0.3 μg m^{-3}). Reductions of RMSE and MAE were observed when the PA-II units were compared to the other PA-II units (Table S3C). The average RMSE and MAE values between all PA-II units before the correction were 3.9 ± 1.0, and 2.4 ± 0.6 μg m^{-3} (respectively); after the corrections of PA-II $PM_{2.5}$ both RMSE and MAE values decreased. Correction of PA-II $PM_{2.5}$ values based on DE-PA-2 with DE-AQ-3 had average RMSE and MAE values of 2.4 ± 0.6 μg m^{-3} and 1.5 ± 0.4 μg m^{-3} (respectively). These values were lower than those that were based on the corrections type of DE-PA-8 with DE-AQ-1 (average RMSE was 3.0 ± 0.7 μg m^{-3} while the average MAE was 1.8 ± 0.5 μg m^{-3}). The PA-II $PM_{2.5}$ value corrections that combined both units had average RMSE and MAE values in the range of the two correction options.

Three different options of PA-II $PM_{2.5}$ value corrections were performed in Salt Lake City, one of SL-AP-13 with each of the two AQMS units (SL-AQ-1 and SL-AQ-2), and another after the $PM_{2.5}$ values of the AQMS units were averaged. The

corrections of the PA-II ($PM_{2.5}$ values) units in Salt Lake City varied depending on the type of corrections and coefficient used. While all options of PA-II $PM_{2.5}$ values corrections improved the comparison between the PA-II to the AQMS units and between the PA-II themselves (Fig. 5F-H, Table S3D), there was no significant change in the R^2 values when the PA-II units were compared to the AQMSs. Overall corrections of PA-II $PM_{2.5}$ values based on SL-AQ-2 had lower R^2 and higher RMSE, and MAE values compared to the other two correction options. The average RMSE, between the PA-II to the AQMS units, based on SL-AQ-1 MLR was 6.0 ± 1.2 μg m^{-3}. After the PA-II $PM_{2.5}$ value corrections the average RMSE reduced to 3.8 ± 0.7 μg m^{-3}. A reduction was also observed for the average MAE value, which changed from 3.7 ± 0.8 μg m^{-3} to 2.3 ± 0.5 μg m^{-3} after implementing the SL-AQ-1 MLR. Reductions of RMSE and MAE values were also observed when the PA-II $PM_{2.5}$ value corrections were based on averaging the $PM_{2.5}$ values from both AQMS units. The average RMSE, between the PA-II to the AQMS units, was 6.1 ± 0.9 μg m^{-3} before the correction of PA-II $PM_{2.5}$ values and 3.9 ± 0.6 μg m^{-3} after the correction of PA-II $PM_{2.5}$ values. A reduction was also observed for the average MAE value, which changed from 3.8 ± 0.7 μg m^{-3} to 2.5 ± 0.4 μg m^{-3} after implementing the average AQMS option. A reduction of RMSE and MAE was also observed when the PA-II units were compared to other PA-II units. Before the PA-II $PM_{2.5}$ value corrections, the average RMSE and MAE values were 4.4. ± 1.7 μg m^{-3} and 2.24 ± 0.96 μg m^{-3}, respectively. After the corrections of PA-II $PM_{2.5}$ values with SL-AQ-1, the average RMSE and MAE values were reduced to 2.65. ± 0.98 μg m^{-3} and 1.44 ± 0.62 μg m^{-3}, respectively. Similar values were found for the PA-II $PM_{2.5}$ value corrections that were based on average AQMSs. The average RMSE and MAE values after the PA-II $PM_{2.5}$ value corrections were 2.64. ± 0.97 μg m^{-3} and 1.43 ± 0.62 μg m^{-3}, respectively.

Overall, almost all the PA-II units had high correlation values when compared with the other PA-IIs or AQMSs in their region. Two PA-II units, SL-PA-6 and SL-PA-8 had low R^2 values with the AQMS, they also had a relatively low correlation with the other PA-II units. It is feazable, that if stricter rules for identifying outlier PA-II units were in use, these two units would have been considered as such and subsequently removed from the data set.

Although improvements of RMSE, MAE, and slope values were observed for the entire research time period, the comparison only provides a general overview on the units' behaviors, but cannot provide information on the variability of the $PM_{2.5}$ values under different conditions mainly under high pollution events. Therefore, in order to evaluate whether the PA-II $PM_{2.5}$ value corrections improved the performance of PA-II units, a comparison of $PM_{2.5}$ values at different locations that experienced similar meteorological conditions or pollution types needed to be performed.

~~**3.1.2 Linear Regression Tests**~~

~~To evaluate the overall trends of the PA-II units compared to the AQMS, we performed a series of regression tests for each site. As in previous works (Gupta et al., 2018; Sayahi et al., 2019) and as commonly used (Clements et al., 2017), these comparisons were performed using linear regression. Each AQMS was compared to all the PA-II units in its area based on hourly $PM_{2.5}$ measurements. Table 2 lists R^2, RMSE values, and the slope and intercept of the linear fit. In general, the linear~~

regression results were mixed. The total R^2 values for the hourly $PM_{2.5}$ measurements ranged from 0.1 to 0.91 with an average of 0.63 ± 0.17, which is relatively high. The RMSE values ranged from 3.89 to 13.13 μg m^{-3} with an average of 7.73 ± 2.05 μg m^{-3}. The slope ranged from 0.03 to 3.12, but was mostly around 1, with an average of 1.15 ± 0.35.

In some locations such as Denver (Table 2B) and Vallejo (Table 2F), high correlation values were found between the local AQMS and the PA-II units. Denver had three AQMSs; each comparison had a high R^2 value in the range of 0.53 to 0.91 (average of 0.72 ± 0.1 for all three AQMSs), average RMSE of 5.65 ± 0.89 μg m^{-3}, and average slope of 1.4 ± 0.18. Vallejo had one AQMS with fifteen PA-II units; the R^2 values ranged from 0.55 to 0.91 with an average of 0.79 ± 0.13. The RMSE values in Vallejo were higher than those in Denver, with an average of 8.95 ± 1.28 μg m^{-3} but with lower average slope of 1.27 ± 0.11. These high correlation values and relatively low RMSE indicate that although the PA-II units and the AQMS are not co-located, they still tend to behave in a similar way. At the other locations, except for Ogden-South Ogden, more than 75 % of the comparisons had high correlation values (>0.5) and only a few with low R^2 value. Several PA-II units had low R^2 values when compared to an AQMS, as in the case of unit 5414 in Berkeley-Oakland and unit 6344 in San Francisco. These two units also had low correlation values compared to the other PA-II units in their region (data not shown). We noticed that unit 6344 was exposed to very high $PM_{2.5}$ concentrations (up to 250 μg m^{-3} for a duration of 3 h) on May 13, 2018. We suspect that this exposure might have affected the instrument efficiency, as was suggested by Sayahi et al. (2019), and therefore, its measurements differ substantially from those of the AQMS. Another exception was Ogden-South Ogden, as all of the comparisons had very low R^2 values (ranging from 0.11 to 0.36 with an average of 0.28 ± 0.1) and high RMSE values (ranging from 8.27 to 10.6 μg m^{-3}). However, when the PA-II units were compared to each other (and not to the AQMS), they showed high correlation values ranging from 0.83 to 0.98 with an average of 0.92 ± 0.05 (Fig. S2). These low correlation values and high RMSE values for the PA-II and AQMS comparisons were most likely caused by specific events and the location of each of the units, as explained below.

A comparison based only on hourly $PM_{2.5}$ values lower than 40 μg m^{-3}, as performed by Sayahi et al. (2019), did not improve the hourly correlation values, as shown in Table S2. Around 88 % of the comparisons had lower correlation values compared to the case when all $PM_{2.5}$ concentrations were used; the R^2 values ranged from 0.04 to 0.9 with an average of 0.57 ± 0.16. Some locations such as Pittsburgh (Table S2A) showed no change in their correlation values for $PM_{2.5}$ <40 μg m^{-3} comparisons whereas others such as Ogden-South Ogden (Table S2F) and Lindon-Orem (Table S2G) showed improved correlation values. Unlike the correlation values, the RMSE values in the comparison of $PM_{2.5}$ < 40 μg m^{-3} improved in 93 % of the cases, resulting in lower RMSE values compared to those found when all $PM_{2.5}$ values were used. The RMSE values ranged from 2.89 to 12.96 μg m^{-3} with an average of 6.83 ± 1.54 μg m^{-3}.

Comparisons based on the $PM_{2.5}$ daily values improved the results (Table S3). The numbers of concurrent $PM_{2.5}$ daily measurements ranged from 18 to 574 days, with an average of 270 ± 119 days per comparison. The correlation values ranged

~~from 0.17 to 0.97 with an average of 0.78 ± 0.15. Further, the RMSE values had a wide range of 2.1–12.8 µg m^{-3} with an average of 4.98 ± 1.77 µg m^{-3}. Overall, 95 % of the comparisons had a higher R^2 and 98 % of the comparisons had lower RMSE values compared to the hourly comparison. Even Ogden–South Ogden, which did not show an improvement in previous comparisons, exhibited better results (Table S3F). The average correlation values in Ogden-South Ogden improved from 0.28 ± 0.1 in the hourly comparison to 0.53 ± 0.12 in the daily comparison. The RMSE values also improved; they decreased from an average of 9.51 ± 0.83 µg m^{-3} in the hourly comparisons to 6.95 ± 0.46 µg m^{-3} in the daily comparisons.~~

3.~~2~~3. Comparison of PA-II units in High Pollution Events

Observations of the PA-II units under high pollution conditions were performed based on daily measurements of $PM_{2.5}$ values from different regions that experienced different pollution types. It is known that ~~D~~different meteorological conditions such as wind direction or speed as well as pollution type (traffic, industrial, wildfire, ~~fireworks,~~ etc.) or source (local vs. regional) may affect the comparison between the AQMSs and the PA-II units. We aimed to determine how the PA-II units behaved (before and after PA-II $PM_{2.5}$ value corrections) in a high-pollution event when the daily $PM_{2.5}$ concertation exceeded~~s~~ the EPA daily regulation of 35 µg m^{-3} Therefore, we ~~decided to~~ investigated specific events with high $PM_{2.5}$ concentrations in different time frames under different atmospheric conditions in each region included in this study.

~~3.2.1. Fireworks in Ogden- South Ogden~~

~~**In Ogden South** Ogden, major differences were observed in the $PM_{2.5}$ values measured during July 2018 (Fig. 3) by the PA-II units and the single AQMS. During this month, we noticed that the AQMS measured very high hourly $PM_{2.5}$ values (with peaks over 400 µg m^{-3}), whereas none of the PA-II units exceeded 20 µg m^{-3}. The regression test results for this month also showed low R^2 values with an average of 0.03 ± 0.01. The location of the units (Fig. S1F), pollution type during this event, and meteorological conditions at the time revealed the cause of these differences. The increase in $PM_{2.5}$ was due to 4th of July fireworks (correlated to July 5, UTC time) that caused an increase in AQMS hourly $PM_{2.5}$ values > 100 µg m^{-3} for a duration of 5 h. The AQMS was located downwind from the main fireworks event (Friendship Park, south of the AQMS) whereas all the PA-II units were far from any fireworks in a residential area on the slopes of Mt. Ogden. Local regulations did not allow the use of fireworks in a residential area (east of road 203; Ogden City Fire Department, 2019) where most of the PA-II units are located. Wind direction information obtained from the local metrological station (see Methods) revealed that the wind was blowing from the fireworks location toward the AQMS but was not reaching the PA-II units. Therefore, the PA-II units could not detect this increase. A similar result was seen in the previous year in July 2017 when only one PA-II unit was active (see Fig. 2F). We also noticed that on July 9, one of the PA-II units (ID 6604) measured high $PM_{2.5}$ values (up to 135 µg m^{-3}) whereas all the other units measured much lower $PM_{2.5}$ values. This high concentration was measured during only one hour~~

~~(23:00 UTC time); therefore, we suspected that this increase was caused by a local source near this specific unit, such as a small-scale fire, lawn mower, or barbeque.~~

~~In both cases, the presence of the PA-II sensors significantly benefited the areas' residents by allowing them to make informed decisions. In the case of the fireworks, if the residents were to base their actions solely on the AQMS data, they would assume that the air quality is unhealthy when actually it is not. If the wind direction was to change and blow from the fireworks toward the residential area, the AQMS data would not prepare the residents at all. In the second case, the localized pollution was identified by the PA-II unit; the AQMS did not measure any changes owing to its location. Overall, the probability of any event being identified by a single AQMS is significantly lower than that of it being identified using multiple PA-II sensors.~~

~~The remaining days included both low-pollution days (July 1–5 and after July 9) and elevated-pollution days (July 7–8). During these days, the PA-II sensors and the AQMS exhibited similar trends, identified the same changes in $PM_{2.5}$ concentrations, and measured similar values. A repeat of the regression tests for only these days (without the fireworks and local event data) resulted in a significant improvement in correlation values; specifically, the average R^2 value increased to 0.69 ± 0.03.~~

~~**3.2.2. Inversion in Utah**~~

~~In Utah, all three locations—Ogden–South Ogden, Lindon–Orem, and Salt Lake City–followed similar daily $PM_{2.5}$ trends during December 4–13, 2018 (Fig. 4). The entire area was affected by an inversion for several days (December 3–13) that increased the daily $PM_{2.5}$ values up to 67.2 ± 4.17 μg m^{-3} and reduced the visibility to almost zero (see photos in Williams, 2019). Overall, at each of these three locations, the values measured by the PA-II units increased at the same time and followed a similar trend to the AQMS measurements. However, whereas all the PA-II units measured similar $PM_{2.5}$ values, the AQMS measured lower $PM_{2.5}$ concentrations. $PM_{2.5}$ values only decreased after precipitation occurred on December 13. The linear regression for each area shows good correlation. In Ogden–South Ogden, Salt Lake City, and Lindon–Orem, the average R^2 was 0.93 ± 0.01, 0.98 ± 0.01 for both AQMSs, and 0.96 ± 0.01, respectively. Overall, at each of these three locations, the PA-II units measured similar values, but these seemed to be overestimated when compared to the AQMS measurements.~~

3.2.~~3~~1. Wildfire in California

~~The three locations~~Two regions in California~~–~~ Vallejo~~, Berkeley–Oakland, ~~and San Francisco are relatively close to each other, and both were affected by a large wildfire that occurred in November 2018. According to the California Statewide Wildfire Recovery Resources (2019), the wildfire started on November 8 at Butte County (north of Vallejo) owing to a combination of strong winds and very dry conditions. A southwesterly wind transferred the wildfire smoke from Butte County toward Vallejo~~, Berkeley–Oakland, ~~and San Francisco. Very high daily $PM_{2.5}$ values (>200 μg m^{-3}) were measured from

November 9 to 21 (Fig. ~~5~~6A and Fig. 6B). During this period, the area had stable meteorological conditions, with low wind speed~~,~~ that reduced visibility down to 1.6 km (1 mile). The high daily $PM_{2.5}$ values decreased only after precipitation started on November 21. Overall, at each of the ~~three locations~~regions, the values measured by the PA-II units increased at the same time and followed a similar trend to the AQMS measurements. A comparison of the $PM_{2.5}$ values measured by the PA-IIs before and after the correction (Fig. 6A and Fig. 6B) showed that the measured (uncorrected) PA-II $PM_{2.5}$ values were higher compared to the AQMS values. In San Francisco, during the wildfire days, the PA-II measured on average 30.3 ± 13.2 μg m^{-3} higher $PM_{2.5}$ daily values than the AQMS. Similar values were also found for Vallejo (30.2 ± 13.0 μg m^{-3}). However, after the correction of PA-II $PM_{2.5}$ values, the daily $PM_{2.5}$ values were lower than before, and they were in a similar range to those measured by the AQMS units. The corrected PA-II $PM_{2.5}$ daily values, during the wildfire, were still slightly higher than those measured by the AQMS during the same time. In San Francisco the corrected $PM_{2.5}$ daily values were on average 6.7 ± 12.0 μg m^{-3} higher than those measured by the AQMS. Lower values were found for Vallejo (1.7 ± 11.8 μg m^{-3}). However, a closer look on the daily values in each of the two regions found that after the correction of PA-II $PM_{2.5}$ values for specific daily values that exceeded 100 μg m^{-3}, the daily values of the corrected PA-IIs were lower than the daily $PM_{2.5}$ values measured by the AQMS. During one such day the corrected PA-II $PM_{2.5}$ daily values were lower by 29 μg m^{-3} compared to the daily $PM_{2.5}$ value measured by the AQMS. The underestimation may be a result of not having enough hourly measurements with such high $PM_{2.5}$ values. Both regions had very few amounts of hours with $PM_{2.5} > 100$ μg m^{-3}, only 0.15% of the hourly measurements in San Francisco had $PM_{2.5}$ hourly measurements from the AQMS that exceed 100 μg m^{-3}. Vallejo had a lower value of 0.1%. Therefore, there were not enough data points to train the MLR model, which resulted in lower PA-II values.

~~Regression test results of each area also show very similar results to each other. In Vallejo, the average R^2 was 0.97 ± 0.01, and in Berkeley-Oakland, where there are three AQMSs, two of them had an average R^2 of 0.95 ± 0.04 and the third had average R^2 of 0.94 ± 0.03~~. ~~In both Vallejo (nine PA-II units) and Berkeley-Oakland (six PA-II units), the average daily $PM_{2.5}$ values of the PA-II units were higher than those measured by the AQMS (Fig. 5A-B). There was no active AQMS at San Francisco during these days, and therefore, only the PA-II units are shown in Fig. 5C. Out of the eight PA-II units located in Berkeley-Oakland (Fig. 5B), two PA-II units (5414 and 10114) measured lower daily $PM_{2.5}$ values compared to the other PA-II units and even compared to the local AQMS.~~

3.2.2. Inversion in Utah

In Utah, ~~all three locations- Ogden-South Ogden, Lindon-Orem, and-~~Salt Lake City~~-followed similar~~ had higher daily $PM_{2.5}$ values ~~trends~~ during December 4-13, 2018 (Fig. ~~4~~6C). The entire area was affected by an inversion for several days (December 3–13) that increased the daily $PM_{2.5}$ values ~~up to 67.2 ± 4.17 μg m^{-3}~~ and reduced the visibility to almost zero (see photos in Williams, 2019). Overall, ~~at each of these three locations,~~ the values measured by the PA-II units increased at the same time

and followed a similar trend to the AQMS measurements. However, whereas all the PA-II units measured similar $PM_{2.5}$ values, uncorrected $PM_{2.5}$ daily values from the PA-II during these days were~~the~~ much higher than those measured by the -AQMS (on average 9.2 ± 7.4 μg m^{-3} more each day). ~~measured lower $PM_{2.5}$ concentrations.~~ $PM_{2.5}$ values only decreased after precipitation occurred on December 13. ~~The linear regression for each area shows good correlation. In Ogden-South Ogden, Salt Lake City, and Lindon-Orem, the average R^2 was 0.93 ± 0.01, 0.98 ± 0.01 for both AQMSs, and 0.96 ± 0.01, respectively. Overall, at each of these three locations, the PA-II units measured similar values, but these seemed to be overestimated when compared to the AQMS measurements.~~ There were three corrections of PA-II $PM_{2.5}$ values for the PA-II units in Salt Lake City. After each of the PA-II $PM_{2.5}$ value corrections, the corrected $PM_{2.5}$ daily values seemed to be similar to those measured by the AQMS units (Fig 6C). The average daily concentration during these days was slightly higher. Both PA-II $PM_{2.5}$ value corrections of PA-II values that were based on SL-PA-13 with SL-AQ-2 or SL-AQ-1, had higher concentrations compared to the AQMS; on average the PA-II daily average, during these days, was higher by 1.7 ± 2.4 μg m^{-3} for SL-PA-13 with SL-AQ-1 and by 1.7 ± 2.5 μg m^{-3} for SL-PA-13 with SL-AQ-2. The PA-II $PM_{2.5}$ value corrections that used averaged AQMS values had slightly higher $PM_{2.5}$ daily values. On average the PA-II values were higher by 2.1 ± 2.6 μg m^{-3} from those measured by the AQMS. The still higher $PM_{2.5}$ values could be due to the volatility of ammonium nitrate, which is a dominant aerosol composition at the region of Salt Lake City during the winter times (Moravek et al., 2019; Womack et al., 2019). It has been shown that sensors similar to the ones used in the PA-II units would be less likely to volatilize ammonium nitrate, unlike the one used in the AQMS units (Grover et al., 2005).

3.2.3. Haze in Denver

On September 4, 2017, a thick hazy smoke from western wildfires settled into eastern Colorado (Spears, 2020). Haze was reported by the DEN meteorological station from 6:00 until the end of the day. Low wind speeds (average of 3.7 ± 0.9 m min^{-1} until 20:00) were recorded and visibility was reduced to 4 km (2.5 miles). Visibility started to increase only around 22:00. $PM_{2.5}$ daily measurements from AQMS units in Denver during this day increased up to 37.1 μg m^{-3}. Before the PA-II $PM_{2.5}$ value corrections, $PM_{2.5}$ daily measurements from the four PA-II units that were active during this period were almost double (the average concentration of the four units was 69.3 ± 2.1 μg m^{-3}). The $PM_{2.5}$ daily measurements from the four PA-II units were higher than those taken by the AQMS units for almost the entire duration of September 1-11, but lower than the remaining days (Fig. 6D). On average the PA-IIs measured 8.3 μg m^{-3} more $PM_{2.5}$ daily concentration than the AQMS units during the entire period. There were three PA-II $PM_{2.5}$ values corrections options in this area, two were based on two different co-located PA-II and AQMS units, and another combined these two pairs. All PA-II $PM_{2.5}$ value corrections showed a reduction in the daily $PM_{2.5}$ values compared to the measured (uncorrected) case, yet the corrected daily $PM_{2.5}$ values were higher for almost this entire period. On average the corrected PA-II $PM_{2.5}$ values were 1.2 μg m^{-3} higher than the AQMS values (during this entire period), when the PA-II $PM_{2.5}$ values were corrected based on DE-PA-2 with DE-AQ-3. Higher values were calculated

in the other two PA-II $PM_{2.5}$ value correction types (3.5 µg m^{-3} and 2.8 µg m^{-3} based on DE-PA-8 with DE-AQ-1), and the one that combined both co-located pairs (respectively).

~~Using AQI maps is another good way to see the spatial and temporal changes in $PM_{2.5}$ measurements; it is also important as the public's behavior is based on the interpretation of the AQI values. We calculated the AQI values for both the PA-II units and the AQMS of all three areas; these calculations were based on the daily $PM_{2.5}$ values (see Methods). We drew maps of all three areas for each day (Fig 6) that show the locations of the AQMS and PA-II units; the locations on the maps are color-coded based on the AQI value at that location on that day. Examining these maps shows us how, as the wildfire and smoke progressed, the air quality worsened. On November 6, before the wildfire started, the AQI for the entire area was moderate. As the fire progressed, the air quality changed from unhealthy on November 11 to very unhealthy on November 16; the air quality became good again only on November 22. Overall, the AQMS and PA-II units in these areas reported similar values and followed similar trends; AQI values differed between the AQMS and PA-II units on a few days are a result of the differences in the $PM_{2.5}$ values used in the calculation. Having multiple PA-II units in each area allows us to track air quality changes with higher resolution, as multiple sensors provide more data than a single AQMS. In the case of the San Francisco area where no AQMS was active, the PA-II units are the only source of data for providing the residents with crucial information about the air quality in their region.~~

~~3.3. Factors That May Impact PA-II Performance~~

~~Meteorological conditions such as wind direction and speed, pollutant type, and pollution source are some of the factors that might affect the performance of the PA-II units. It is therefore important to also evaluate and consider additional factors such as other meteorological conditions and underlying technology used when comparing the behavior and measurements of the PA-II units and the AQMS.~~

~~3.3.1. Temperature and RH~~

~~The sensitivity of the PA-II unit to changes in temperature and RH remains unknown (Gupta et al., 2018). We can assume that changes in temperature or RH may affect the performance of the PA-II unit especially under atmospheric conditions as they cannot be controlled. Jayaratne et al. (2018) tested an older version of the PMS unit (PMS1003) and reported such an effect. Most low-cost sensors have no heater or dryer at their inlet to remove water from the sample before measuring the particles; therefore, deliquescent or hygroscopic growth of particles, mainly under high RH conditions (>75 %), can lead to higher reported PM concentrations (Jayaratne et al., 2018). According to Rai et al. (2017), most low-cost sensors show some sensitivity to RH conditions but not to temperature. It is therefore important to evaluate whether the PA-II unit will be affected by changes in temperature or RH. To do so, we used temperature and RH measurements from the nearest available~~

~~meteorological stations (see Methods for station information) and, in some cases, additional measurements from the AQMS (e.g., in Pittsburgh, Denver, Ogden–South Ogden, Lindon–Orem, and Salt Lake City).~~

~~The hourly temperature measurements from the meteorological stations were compared with the hourly $PM_{2.5}$ measurements from each PA-II unit (86 units in total) using linear regression. The regression resulted in very low R^2 values that ranged from 1×10^{-9} to 0.07 with an average of 0.02 ± 0.02. Similar results were found when the AQMS temperature measurements were used (52 units in total, Table S4); the R^2 values ranged from 6×10^{-5} to 0.13 with an average of 0.04 ± 0.03. For the RH, two different comparisons were made: a comparison using all RH values and a comparison for only those cases in which the RH value was higher than 75 %. When using RH data from the meteorological stations and for the entire RH range, very low R^2 values were found. The correlations values ranged from 7.5×10^{-7} to 0.1 with an average of 0.02 ± 0.03. Comparison results obtained using RH measurements from the AQMS were similar (Table S4); the R^2 values ranged from 1.01×10^{-5} to 0.17 with an average of 0.05 ± 0.04. Even when only RH > 75 % was tested, the R^2 values ranged from 1.6×10^{-7} to 0.1 with an average of 0.01 ± 0.01 for RH measurements from the meteorological station. Similar values were also found for RH measured by the AQMS; R^2 values ranged from 5.5×10^{-6} to 0.18 with an average of 0.02 ± 0.04. Similar results have been reported previously as well. For example, Sayahi et al. (2019) found very low correlation values between measurements from the PMS5003 sensor and the temperature/RH under atmospheric conditions. Holstius et al. (2014) found a negligible effect of temperature or RH on measurements performed using low-cost sensors under ambient conditions. However, several studies that used old PMS units, such as PMS1003 that was used in PA-I or PMS3003 that was never used in any PA units, found that these sensors were affected by RH (Kelly et al., 2017; Jayaratne et al., 2018; Zheng et al., 2018). AQ-SPEC (2018) tested the PA-II unit in a laboratory setting under different temperature and RH conditions and found that most temperature and RH combinations had a minimal effect on the PA-II's precision. Our findings for PA-II units in the field under atmospheric conditions are in agreement with those of the AQ-SPEC (2018).~~

3.4. Impact of Distance on Comparisons Between the Units

Previous studies obtained good results when comparing the PA-II unit or PMS5003 sensor and the FRM and FEM units when the two units were co-located. The AQ-SPEC (2018) recently released a report comparing PA-II units to two FEM instruments under laboratory and field conditions. They found good correlations for hourly and daily values of both $PM_{2.5}$ and PM_{10} under field conditions with higher correlation values for $PM_{2.5}$ compared to those for PM_{10}. Gupta et al. (2018) compared three PA-II units in California to a single FEM unit and obtained good correlation values ($R^2 > 0.9$). Sayahi et al. (2019) co-located reference air monitors (tapered element oscillating microbalance, TEOM), and FRM unit, next to a PMS5003 (used in the PA-II unit) in Salt Lake City. The PMS5003 $PM_{2.5}$ measurements correlated well with the hourly TEOM measurements ($R^2 > 0.87$) and with the daily FRM measurements ($R^2 > 0.88$). In our study, we did not position the PA-II units. Further, in most cases, the AQMS and the PA-II units were not co-located; therefore, they might have been exposed to different particle types

and concentrations. Some might claim that not having the PA-II and FRM units co-located, as was done in previous studies, might diminish the accuracy of the comparison between these units. Although lower correlation values were in fact observed in our study, as we were using PA-II units in their natural locations, this was expected. Further, as we saw that the correlation values are not much lower than those in the co-located cases described in previous studies, they are still statistically significant. Because the AQMS and the PA-II units were not co-located, we wanted to verify whether the distance between all the units affected the R^2, RMSE, MAE and slope values. We compared the R^2, RMSE, MAE and slope values received from the comparisons of hourly $PM_{2.5}$ measurements with the corresponding distances between the units (Fig. 7). There was no correlation between the two. Not when the PA-II units were compared to the nearest AQMS units (Fig. 7A), or between the PA-II units (Fig.7B), before or after the corrections of the PA-II $PM_{2.5}$ values. Therefore, the distance between the units did not impact the comparison.

~~**3.3.2. Technology, Maintenance, and Placement**~~

3.5. Underlying Differences and Future Implications

While appropriate PA-II $PM_{2.5}$ value corrections can improve the comparison between the PA-IIs with reference units, t~~T~~here are ~~many~~ other differences between PA-IIs and AQMS units that can influence the comparison results, including the underlying technology and the manner in which units are placed. The $PM_{2.5}$ sensors in the AQMSs perform gravimetric measurements using the mass of the particle; by contrast, the PA-II units use~~s~~ a laser particle counter to count electric pulses generated as particles crossing through a laser beam. The method used by the PA-II might impact the count of particles during high humidity conditions or when a majority of the particles are volatile. Another difference is the physical location of the units; whereas AQMSs are meticulously positioned in an open area, the location of a PA-II ~~sensor~~ is determined by its owner. Although PurpleAir recommends positioning the PA-II ~~sensor~~ in an open area, ultimately, it is the owner's decision. In practice, most of the PA-II units are located in residential areas with low-rise housing. Furthermore, the height at which the sensor is located could affect the measurements. ~~Whereas t~~The height of the AQMS inlet is regulated and kept constant at each location~~;~~, on the other hand, the owner of a PA-II unit can freely place it near the ground or higher up. The location of the PA-II units in residential areas can provide both an advantage and a disadvantage. For example, ~~as in the case of Ogden South Ogden,~~ a single PA-II unit might be exposed to more localized PM sources such as a ~~barbeque~~barbecue, lawn mower, or car, making it report different results compared with other units in its area. Therefore, an increase of PM by a single PA-II unit should be taken into account. When the PA-II is used as a network, as suggested by Ford et al. (2019), comparison of the PM values measured by all PA-II units will help identify such a localized source. Maintenance and calibration are other possible causes of differences between the two. The $PM_{2.5}$ sensors in the AQMSs have strict rules for the monthly evaluation of sensor performance, including through flow calibration or calibration based on minimum value threshold (which, in some cases, causes the recording of negative PM values). By contrast, PA-II units do not have any quality control other than that done by

the company for each sensor before shipment to the customer (PurpleAir personal communication, 2019). Another point that should be taken into account is the lifetime of the PA-II units. The manufacturer of the PMS5003 sensor used in the PA-II units states that it has a lifetime expectancy of ~3 years (Yong, 2016). Bi et al. (2020) found that the PA-II unit's efficiency is affected even after only two years of being operational.

Based on the findings from this work, we believe that there are several needed steps that will allow the usage of the PA-II units in air quality and health related research. First, users should identify regions with multiple PA-II units, where at least one PA-II is co-located with an FRM or FEM unit. Ideally the same location will also contain measurements of T and RH, or at least T and RH measurements will be nearby. Keep in mind that it is not recommended to use the PA-II internal sensors for T and RH values, as they are not representing the atmospheric measurements (Malinges et al., 2020; PurpleAir personal communication, 2019). However, we have found that in many regions there is no meteorological station that can serve as a reference for the correction process. It would be useful then, to devise a way in which the PA-II internal T and RH sensors can be used. To achieve this, an extensive study is necessary, to gain a better understanding of the issues related to the usage of the PA-II internal sensors and to formulate a calibration equation that then can be applied to the desired PA-II units.

Comparison of all PA-II units in each region will help to identify and remove outlier PA-II units from future analysis. Exposure to high PM concentration might affect the PA-II efficiency, as suggested by Sayahi et al. (2019), and therefore, its measurements will differ substantially from those of the AQMSs and other PA-II units. Ideally PurpleAir should monitor all active PA-II units and identify units that behave differently from surrounding PA-II units or identify PA-II units whose internal sensors (A and B) report different values, flag them on the online map, and communicate instructions to the unit owners on how to fix or replace the unit.

After PA-II units have been identified, users should conduct MLR between the co-located PA-II and AQMS units, including measurements of T and RH. For the MLR to be efficient it is important have a wide range of $PM_{2.5}$, T and RH measurements. This MLR will provide a coefficient that will be used to correct all the remaining $PM_{2.5}$ values of all PA-II units in that region. Evaluation of the PA-II $PM_{2.5}$ value corrections should be made for the duration of the study but also for specific events with spatial impact such as inversion, dust storms, biomass burning, and more. Such events should impact a larger area and therefore will allow detection of the PM changes in all PA-II units as a whole (network). Correction of PA-II $PM_{2.5}$ values should be performed per region, as they represent specific PM values as well as changes of T and RH values that the PA-II units were exposed to. This will help the public obtain information on the spatial and temporal distribution of PM concentrations in their area (Gupta et al., 2018; Morawska et al., 2018), which will enable them to monitor local air-quality conditions (Williams et al., 2018) and help make decisions related to events with high PM exposure.

In this study, we evaluated PA-II units that were up to 5 km away from an AQMS unit, as well as up to 10 km from each other. This raises the question of maximum effective distance. What is the maximum distance between an AQMS and PA-II

units that will still allow for the MLR to successfully correct the measurement taken by PA-II units; a distance greater than this would carry the potential of introducing additional factors that might impact the comparisons. Another situation that requires further investigation is that of regions that include multiple PA-II units but do not have a co-located pair or completely lack a reference monitoring station. The question in mind, if and how we might use neighboring regions in which measurements were successfully corrected to compensate in the case of such problematic areas. For example, could we have used Vallejo and San Francisco, two regions that were included in this study to correct the measurements of the PA-II units in the region of Berkeley - Oakland that resides between the two?

~~**3.3.3. Distance and Number of Comparisons Between the Units**~~

~~Other factors that could affect the comparisons with the AQMS are the distances between the units or the number of observations. Previous studies obtained good results when comparing between the PA-II unit or PMS5003 sensor and the FRM and FEM units when the two units were co-located. The AQ-SPEC (2018) recently released a report comparing PA-II units to two FEM instruments under laboratory and field conditions. They found good correlations for hourly and daily values of both $PM_{2.5}$ and PM_{10} under field conditions with higher correlation values for $PM_{2.5}$ compared to those for PM_{10}. Gupta et al. (2018) compared three PA-II units in California to a single FEM unit and obtained good correlation values ($R^2 > 0.9$). Sayahi et al. (2019) co-located reference air monitors (tapered element oscillating microbalance, TEOM), and FRM unit, next to a PMS5003 (used in the PA-II unit) in Salt Lake City. The PMS5003 $PM_{2.5}$ measurements correlated well with the hourly TEOM measurements ($R^2 > 0.87$) and with the daily FRM measurements ($R^2 > 0.88$). In our study, we did not position the PA-II units. Further, in most cases, the AQMS and the PA-II units were not located at the same place; therefore, they might have been exposed to different particle types and concentrations. Some might claim that not having the PA-II and FRM units co-located, as was done in previous studies, might diminish the accuracy of the comparison between these units. Although lower correlation values were in fact observed in our study, as we were using PA-II units in their natural locations, this was expected. Further, as we saw that the correlation values are not much lower than those in the co-located cases described in previous studies, they are still statistically significant. Because the AQMS and the PA-II units were not co-located, we wanted to verify whether the distance between the AQMS and the PA-II units affected the R^2 values. We compared the R^2 values that we previously calculated for the hourly $PM_{2.5}$ measurements with the corresponding distances between the PA-II units and AQMS (Fig. S3A). There was no correlation between the two, and similar results were found when the RMSE values were tested (Fig. S3B). The number of observations used for the comparison was also tested; comparing the same R^2 from the measurements with the number of observations revealed no effect of the number of observations on R^2 or RMSE values (Fig. S3C-D).~~

~~**3.4. Next Steps with PA-II units**~~

~~Ford et al. (2019) suggested the use of PA-II units as a network installed by residents in an in North Colorado. This seems like a good solution for locations that are lacking FRM or FEM units as multiple sensors can provide more data. However, it is~~

~~important to consider the limitations of the PA-II unit. The PA-II unit needs to be monitored for changes in unit behavior. We recommend PurpleAir to monitor the measurements of the PA-II units, identify units that behave differently from other surrounding units or units whose internal sensors (A and B) report different values, flag them on the online map, and communicate instructions to the unit owners on how to clean the unit. The manufacturer of the PMS5003 sensor that is used in the PA-II units noted that it has a lifetime of ~3 years (Yong, 2016). None of the current units have been active for that long; therefore, the efficiency of PA-II units over such a long period remains unknown and should be evaluated. It is possible that, after this duration, they will lose their efficiency (a behavior known as drift) and will become outliers.~~

4. Conclusions

PA-II units are becoming a common low-cost tool to monitor changes in the concentrations of PMs of various sizes. Previous studies have examined the performance of these PA-II units (or the sensor in them) by comparing them with a ~~co-located~~co-located EPA AQMS. However, a majority of PA-II units are not ~~co-located~~co-located in practice, and some of them are placed in areas where there is no reference air monitoring system. This study aimed to examine the behavior of PA-II units under atmospheric conditions when exposed to a variety of pollutants and different $PM_{2.5}$ concentrations. For this purpose, we used PA-II units that have already been active for some time, irrespective of where they might be located. ~~Eight~~ Four ~~locations~~ regions with multiple PA-II units and at least a single AQMS were identified. Each region had at least one co-located pair of a PA-II with an AQMS. Corrections of PA-II $PM_{2.5}$ values using MLR based on the AQMSs' $PM_{2.5}$, RH, and T values of these co-located units improved the comparison of the PA-II (co-located and not co-located) with the AQMS unit (higher R^2 and, lower RMSE and MAE as well as better slope values). ~~Each PA-II unit was compared to the AQMS and to other PA-II units in its surrounding area based on hourly or daily $PM_{2.5}$ measurements.~~ Overall, the PA-II units behaved in ~~a~~ similarly way to the other PA-II units ~~at~~ in their ~~locations~~regions. Without corrections of PA-II $PM_{2.5}$ values, the majority ~~We found that even though some~~ PA-II units measured much higher values than the AQMSs. ~~overestimated or underestimated at times, the AQMS~~After corrections of PA-II $PM_{2.5}$ values, ~~and~~ PA-II units were ~~mostly~~ in agreement and measured overall similar $PM_{2.5}$ concentrations. ~~PA-II was also found to not be affected by temperature or RH.~~ We think that the PA-II unit is a promising tool for measuring $PM_{2.5}$ concentrations and identifying relative concentration changes as long as the PA-II $PM_{2.5}$ values can be corrected. ~~Further, through the use of AQI, the current air quality can be successfully conveyed to the public.~~ ~~The PA-II unit has the potential to complement sparsely distributed monitoring stations, particularly in areas lacking a nearby AQMS.~~

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

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

Acknowledgment

The authors are thankful to the PurpleAir team for their help and explanations about the PA-II units. Further, they are thankful to Mr. Mangus Nick from the National Air Data Group at US EPA for his help with the EPA data. Finally, they are thankful to Dr. Amber McCord, College of Media & Communication at Texas Tech University, for her help with the graphic abstract. Use of the sensor manufacturer's name does not imply endorsement.

References

AQ-SPEC, the Air Quality Sensor Performance Evaluation Center- PurpleAir PA-II evaluation summary: http://www.aqmd.gov/docs/default-source/aq-spec/summary/purpleair-pa-ii---summary-report.pdf?sfvrsn=4, last access: 1 August 2019.

Bi, J. Wildani, A., Chang, H. H. and Liu, Y.: Incorporating Low-Cost Sensor Measurements into High-Resolution $PM_{2.5}$ Modeling at a Large Spatial Scale. Environ. Sci. Technol. 2020, 54, 2152−2162, DOI: 10.1021/acs.est.9b06046.

California wildfires statewide recovery resources. November 2018 Fires: http://wildfirerecovery.org/general-info/, last access: June 21, 2019.

Castell, N., Dauge, F. R., Schneider, P., Vogt, M., Lerner, U., Fishbain, B., Broday, D., and Bartonova, A.: Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?, Environ. Int., 99, 293-302, 2017.

Clements, A., Griswold, W., R. S, A., Johnston, J. E., Herting, M. M., Thorson, J., Collier-Oxandale, A., and Hannigan, M.: Low cost air quality monitoring tools: from research to practice (a workshop summary), Sensors., 17(11), 2478, doi:10.3390/s17112478, 2017.

Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope,C. A., Shin, H., Straif, K., Shaddick, G., Thomas, M., Dingenen, R. van, Donkelaar, A. van, Vos, T., Murray, C. J. L. and Forouzanfar, M. H.: Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015, The Lancet, 389(10082), 1907-1918, doi:10.1016/S0140-6736(17)30505-6, 2017.

Commodore, A., Wilson, S., Muhammad, O., Svendsen, E., and Pearce, J.: Community-based participatory research for the study of air pollution: a review of motivations, approaches, and outcomes, Environ. Monit. Assess., 189(8), 378, doi:10.1007/s10661-017-6063-7, 2017.

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

Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F., Schwartz, J.: Air pollution and mortality in the Medicare population, N. Engl. J. Med., 376, 2513-2522, doi:10.1056/NEJMoa1702747pmid:28657878, 2017.

EPA, Air Quality Index Guide to Air Quality and Your Health, EPA-456/F-14-002, 2014, https://www3.epa.gov/airnow/aqi_brochure_02_14.pdf, last access: 1 August 2018.

Di Antonio, A. Popoola, O. A. Ouyang, B. Saffell, J. and Jones, R. L.: Developing a relative humidity correction for low-cost sensors measuring ambient particulate matter. Sensors 18, 2790, https://doi.org/10.3390/s18092790 (2018).

Ford, B., Pierce, J. R., Wendt, E., Long, M., Jathar, S., Mehaffy, J., Tryner, J., Quinn, C., van Zyl, L., L'Orange, C., Miller-Lionberg, D., and Volckens, J.: A low-cost monitor for measurement of fine particulate matter and aerosol optical depth – Part 2: Citizen-science pilot campaign in northern Colorado, Atmos. Meas. Tech., 12, 6385–6399, https://doi.org/10.5194/amt-12-6385-2019, 2019. ~~Ford, B., Pierce, J. R., Wendt, E., Long, M., Jathar, S., Mehaffy, J., Tryner, J., Quinn, C., van Zyl, L., L'Orange, C., Miller-Lionberg, D., and Volckens, J.: A low-cost monitor for measurement of fine particulate matter and aerosol optical depth - Part 2: Citizen science pilot campaign in northern Colorado, Atmos. Meas. Tech. Discuss., doi.org/10.5194/amt-2019-109, 2019.~~

Forouzanfar, M. H., Alexander, L., Anderson, H. R., Bachman, V. F., Biryukov, S., Brauer, M., Burnett, R., Casey, D., Coates, M. M., Cohen, A., Delwiche, K., Estep, K., Frostad, J. J., KC, A., Kyu, H. H., Moradi-Lakeh, M., Ng, M., Slepak, E. L., Thomas, B. A., Wagner, J., Aasvang, G. M., Abbafati, C., Ozgoren, A. A., Abd-Allah, F., Abera, S. F., Aboyans, V., Abraham, B., Abraham, J. P., Abubakar, I., Abu-Rmeileh, N. M. E., Aburto, T. C., Achoki, T., Adelekan, A., Adofo, K., Adou, A. K., Adsuar, J. C., Afshin, A., Agardh, E. E., Al Khabouri, M. J., Al Lami, F. H., Alam, S. S., Alasfoor, D., Albittar, M. I., Alegretti, M. A., Aleman, A. V., Alemu, Z. A., Alfonso-Cristancho, R., Alhabib, S., Ali, R., Ali, M. K., Alla, F., Allebeck, P., Allen, P. J., Alsharif, U., Alvarez, E., Alvis-Guzman, N., Amankwaa, A. A., Amare, A. T., Ameh, E. A., Ameli, O., Amini, H., Ammar, W., Anderson, B. O., Antonio, C. A. T., Anwari, P., Cunningham, S. A., Arnlöv, J., Arsenijevic, V. S. A., Artaman, A., Asghar, R. J., Assadi, R., Atkins, L. S., Atkinson, C., Avila, M. A., Awuah, B., Badawi, A., Bahit, M. C., Bakfalouni, T., Balakrishnan, K., Balalla, S., Balu, R. K., Banerjee, A., Barber, R. M., Barker-Collo, S. L., Barquera, S., Barregard, L., Barrero, L. H., Barrientos-Gutierrez, T., Basto-Abreu, A. C., Basu, A., Basu, S., Basulaiman, M. O., Ruvalcaba, C. B., Beardsley, J., Bedi, N., Bekele, T., Bell, M. L., Benjet, C., Bennett, D. A., et al.: Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013, The Lancet, 386(10010), 2287-2323, doi:10.1016/S0140-6736(15)00128-2, 2015.

Grover, B. D., Kleinman, M., Eatough, N. L., Eatough, D. J., Hopke, P. K., Long, R. W., Wilson, W. E., Meyer, M. B., and Ambs, J. L.: Measurement of total PM2.5 mass (nonvolatile plus semi-volatile) with the Filter Dynamic Measurement System tapered element oscillating microbalance monitor, J. Geophys. Res. Atmos., 110(7), D07S03, doi:10.1029/2004JD004995, 2005.

Gupta, P., Doraiswamy, P., Levy, R., Pikelnaya, O., Maibach, J., Feenstra, B., Polidori, A., Kiros, F., and Mills, K. C.: 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, doi.org/10.1029/2018GH000136, 2018.

Hagan, D. H., Isaacman-VanWertz, G., Franklin, J. P., Wallace, L. M. M., Kocar, B. D., Heald, C. L., and Kroll, J. H.: Calibration and assessment of electrochemical air quality sensors by co-location with regulatory-grade instruments, Atmos. Meas. Tech., 11, 315-328, doi.org/10.5194/amt-11-315-2018, 2018.

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

Jayaratne, R., Liu, X., Thai, P., Dunbabin, M., and Morawska, L.: The influence of humidity on the performance of a low-cost air particle mass sensor and the effect of atmospheric fog, Atmos. Meas. Tech., 11, 4883-4890, doi.org/10.5194/amt-11-4883-2018, 2018.

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

Klemm, R. J. and Mason, Jr R. M.: Aerosol Research and Inhalation Epidemiological Study (ARIES): air quality and daily mortality statistical modelling - interim results, J Air Waste. Manag. Assoc., 50,1433-1439. 2000;

Kuula, J., Mäkelä, T., Hillamo, R., and Timonen, H.: Response characterization of an inexpensive aerosol sensor, Sensors., 17, 2915, doi:10.3390/s17122915, 2017.

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., AlMazroa, M. A., Amann, M., Anderson, H. R., Andrews, K. G., Aryee, M., Atkinson, C., Bacchus, L. J., Bahalim, A. N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M. L., Blore, J. D., Blyth, F., Bonner, C., Borges, G., Bourne, R., Boussinesq, M., Brauer, M., Brooks, P., Bruce, N. G., Brunekreef, B., Bryan-Hancock, C., Bucello, C., Buchbinder, R., Bull, F., Burnett, R. T., Byers, T. E., Calabria, B., Carapetis, J., Carnahan, E., Chafe, Z., Charlson, F., Chen, H., Chen, J. S., Cheng, A. T.-A., Child, J. C., Cohen, A., Colson, K. E., Cowie, B. C., Darby, S., Darling, S., Davis, A., Degenhardt, L., Dentener, F., Des Jarlais, D. C., Devries, K., Dherani, M., Ding, E. L., Dorsey, E. R., Driscoll, T., Edmond, K., Ali, S. E., Engell, R. E., Erwin, P. J., Fahimi, S., Falder, G., Farzadfar, F., Ferrari, A., Finucane, M. M., Flaxman, S., Fowkes, F. G. R., Freedman, G., Freeman, M. K., Gakidou, E., Ghosh, S., Giovannucci, E., Gmel, G., Graham, K., Grainger, R., Grant, B., Gunnell, D., Gutierrez, H. R., Hall, W., Hoek, H. W., Hogan, A., Hosgood III, H. D., Hoy, D., Hu, H., Hubbell, B. J., Hutchings, S. J., Ibeanusi, S. E., Jacklyn, G. L., Jasrasaria, R., Jonas, J. B., Kan, H., Kanis, J. A., Kassebaum, N., Kawakami, N., Khang, Y.-H., Khatibzadeh, S., Khoo, J.-P., et al.: A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010, The Lancet, 380(9859), 2224–2260, doi:10.1016/S0140-6736(12)61766-8, 2012.

Ling, S. H., and van Eeden, S. F.: Particulate matter air pollution exposure: Role in the development and exacerbation of chronic obstructive pulmonary disease, Int. J. Chron. Obstruct. Pulmon. Dis., 4, 233-243, 2009.
Lundgren, D.A.; Cooper, D.W. Effect of Humidity on Light-Scattering Methods of Measuring Particle Concentration. J. Air Pollut. Control Assoc., 19, 243–247, 1969.
Magi B. I., Cupini, C., Francis, J., Green, M. and Hauser C.: Evaluation of PM2.5 measured in an urban setting using a low-cost optical particle counter and a Federal Equivalent Method Beta Attenuation Monitor Aerosol Sci. Technol., 54,147-159, 2020.
Malings, C., Tanzer, R., Hauryliuk, A., Saha, P.K., Robinson, A.L., Preso, A.A. and Subramanian, R.: Fine particle mass monitoring with low-cost sensors: Corrections and long-term performance evaluation. Aerosol Sci. Technol., 54, 160-174, 2020.
Moravek, A., Murphy, J. G., Hrdina, A., Lin, J. C., Pennell, C., Franchin, A., Middlebrook, A. M., Fibiger, D. L., Womack, C. C., McDuffie, E. E., Martin, R., Moore, K., Baasandorj, M., and Brown, S. S.: Wintertime spatial distribution of ammonia and its emission sources in the Great Salt Lake region, Atmos. Chem. Phys., 19, 15691–15709, https://doi.org/10.5194/acp-19-15691-2019, 2019.
Morawska, L., Thai, P. K., Liu, X., Asumadu-Sakyi, A., Ayoko, G., Bartonova, A., Bedini, A., Chai, F. Christensen, B., and Dunbabin. M.: Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: how far have they gone?, Environ Int., 116, 286-99, doi: 10.1016/j.envint.2018.04.018, 2018.
~~Ogden City Fire Department, fireworks restriction July 2, 2019: https://www.facebook.com/626740550676142/photos/a.729348300415366/2787170824633093/?type=3&theater, last access: 1 August 2019.~~
PurpleAir, PurpleAir Map, air quality Map: http://map.purpleair.org/, last access: 1 August 2019.
Rai, A. C., Kumar, P., Pilla, F., Skouloudis, A. N., Di Sabatino, S., Ratti, C., Yasar, A., and Rickerby, D.: End-user perspective of low-cost sensors for outdoor air pollution monitoring, Sci. Total Environ., 607-608, 691-705, 2017.
Sayahi, T., Butterfield, A., and Kelly K.E.: Long-term field evaluation of the Plantower PMS low-cost particulate matter sensors, Environ. Pollut., 245, 932-940, 2019.
Schwartz, J., Dockery, D. W., Neas, L.M.: Is daily mortality associated specifically with fine particles?, J Air Waste. Manag. Assoc., 46, 927-939, 1996.
Shiraiwa, M., Ueda, K., Pozzer, A., Lammel, G., Kampf, C. J., Fushimi, A., Enami, S., Arangio, A. M., Frohlich-Nowoisky, J., Fujitani, Y., Furuyama, A., Lakey, P. S. J., Lelieveld, J., Lucas, K., Morino, Y., Poschl, U., Takaharna, S., Takami, A., Tong, H. J., Weber, B., Yoshino, A., and Sato, K.: Aerosol Health Effects from Molecular to Global Scales, Environ. Sci. Technol., 51, 13545-13567, 2017.
Spears, C. Western Fires Cause Denver's Mountain View To Go Missing, September 04th, 2017: https://denver.cbslocal.com/2017/09/04/wildfire-smoke-flows-into-colorado/, last access: 1 July 2020.
Wang, Y., Li, J., Jing, H., Zhang, Q., Jiang, J., and Biswas, P.: Laboratory Evaluation and Calibration of Three Low-Cost Particle Sensors for Particulate Matter Measurement, Aerosol Sci. Technol., 49, 1063-1077, 2015.
Watson, J. G., Tropp, R. J., Kohl, S. D., Wang, X. L., Chow, J. C.: Filter processing and gravimetric analysis for suspended particulate matter samples, Aerosol Sci. Eng., 1,93-105, 2017.
Williams, C., Before and after: How pollution trapped by the inversion changes Salt Lake's scenery, posted - Dec 10th, 2018: https://www.ksl.com/article/46445294/before-and-after-how-pollution-trapped-by-the-inversion-changes-salt-lakes-scenery, last access: 1 August 2019.
Williams, R., Nash, D., Hagler, G., Benedict, K., MacGregor, I., Seay, B., Lawrence, M., Dye, T., September 2018. Peer Review and Supporting Literature Review of Air Sensor Technology Performance Targets. EPA Technical Report Undergoing Final External Peer Review. EPA/600/R-18/324.
Womack, C. C., McDuffie, E. E., Edwards, P. M., Bares, R., Gouw, J. A. A., Docherty, K. S., Dubé, W. P., Fibiger, D. L., Franchin, A., Gilman, J. B., Goldberger, L., Lee, B. H., Lin, J. C., Long, R., Middlebrook, A. M., Millet, D. B., Moravek, A., Murphy, J. G., Quinn, P. K., Riedel, T. P., Roberts, J. M., Thornton, J. A., Valin, L. C., Veres, P. R., Whitehill, A. R., Wild, R. J., Warneke, C., Yuan, B., Baasandorj, M., and Brown, S. S.: An Odd Oxygen Framework for Wintertime Ammonium Nitrate Aerosol Pollution in Urban Areas: NOx and VOC Control as Mitigation Strategies, Geophys. Res. Lett., 46, 4971–4979, https://doi.org/10.1029/2019GL082028, 2019.

Woodall, G. M., Hoover, M. D., Williams, R., Benedict, K., Harper, M., Soo, J., Jarabek, A. M., Stewart, M. J., Brown, J. S., Hulla, J. E., Caudill, M., Clements, A. L., Kaufman, A., Parker, A. J., Keating, M., Balshaw, D., Garrahan, K., Burton, L., Batka, S., Limaye, V. S., Hakkinen, P. J., and Thompson, B.: Interpreting Mobile and Handheld Air Sensor Readings in Relation to Air Quality Standards and Health Effect Reference Values: Tackling the Challenges, Atmosphere., 8, 182, doi:10.3390/atmos8100182, 2017.

Yong, Z.: Digital universal particle concentration sensor, PMS5003 series data manual: http://www.aqmd.gov/docs/default-source/aq-spec/resources-page/plantower-pms5003-manual_v2-3.pdf, last access: 1 August 2018.

Zheng, T., Bergin, M. H., Johnson, K. K., Tripathi, S. N., Shirodkar, S., Landis, M. S., Sutaria, R., and Carlson, D. E.: Field evaluation of low-cost particulate matter sensors in high- and low-concentration environments, Atmos. Meas. Tech., 11, 4823-4846, doi.org/10.5194/amt-11-4823-2018, 2018.

Table legends

slopeTable 1: Detail of the coefficients received in each MLR as well as the liner regression output including R^2, RMSE, MAE and slope for each correction PA-II unit

Figure legends

Figure 1. (A) Picture from the bottom of the PA-II unit containing two PMS5003 sensors (in blue). (B) Schematic of a single PMS5003 sensor. A fan draws the particles through the inflow (rounded holes) at the lower level of the sensor. The particles travel to the upper part of the sensor where they come out through the air flow holes and then pass through the laser path, causing the beam to scatter. Finally, the particles exit from the fan.

Figure 2: Maps of location with AQMS and PA-II units, each map (A-D) represent a different region. (A) Denver; (B) San Francisco; (C) Vallejo, and Salt Lake City (D). Maps created using Google map. AMQS unit is represented by the green points and the PA-II units, by the purple points.

Figure 23. ~~Distribution~~ Time series of daily $PM_{2.5}$ measurements from the AQMS and PA-II units in each of the ~~eight~~ four areas: (A) ~~Pittsburgh; (B)~~ Denver; ~~(C) Berkeley-Oakland; (~~DB) San Francisco; (~~E~~C) Vallejo~~; (F) Ogden-South Ogden; (G) Lindon-Orem~~, and (~~H~~D) Salt Lake City. Measurements from AQMS are represented by the green lines and the PA-II units are indicated by purple lines. ~~The numbers are the units' ID numbers.~~

Figure 4. Comparison of hourly uncorrected PA-II ($PM_{2.5}$) values compared to co-located AQMS ($PM_{2.5}$) values (black) and corrected PA-II ($PM_{2.5}$) values compared to AQMS ($PM_{2.5}$) values (gray). Dash lines represent a 1:1 line. Statistics of each case included the R2, RMSE and MAE for the uncorrected and corrected (MLR) data. N represent the number of data points used in the MLR. (A) for San Francisco, (B) Vallejo, (C-E) different MLR from Denver, DE-PA-2 with DE-AQ-3 (C), DE-PA-8 with DE-AQ-1 (D) and new MLR based on DE-PA-2 with DE-AQ-3, and DE-PA-8 with DE-AQ-1 (E). (F-I) different MLR from Salt Lake City, (F) for SL-PA-13 with SL-AQ-1, (G) SL-PA-13 with SL-AQ-2, and (H) SL-PA-13 with an average of both AQMS units.

Figure 5. Comparison of hourly uncorrected PA-II ($PM_{2.5}$) values compared to nearest AQMS ($PM_{2.5}$) values (black) and corrected PA-II ($PM_{2.5}$) values compared to AQMS ($PM_{2.5}$) values (gray). Dash lines represent a 1:1 line. Statistics of each case included the R2, RMSE and MAE for the uncorrected and corrected (MLR) data. N represent the number of data points used in the MLR. Each figure represents corrections based on different MLR type. (A) SF-PA-1 with SF-AQ-1 in San Francisco, (B) Vallejo was based on VA-PA-1 with DE-AQ-1, (C-E) different corrections in Denver, (C) based on DE-PA-2 with DE-AQ-3, (D) based on DE-PA-8 with DE-AQ-1, and E based on combining of DE-PA-2 with DE-AQ-3, and DE-PA-8 with DE-AQ-1 (E). (F-I) different corrections in Salt Lake City, (F) based on SL-PA-13 with SL-AQ-1, (G) based on SL-PA-13 with SL-AQ-2, and (H) based on SL-PA-13 with an average of both AQMS units.

Figure 6. Daily measurements of $PM_{2.5}$ at (A) San Francisco and (B) Vallejo during the November 2018 wildfire (UTC time). Daily $PM_{2.5}$ measurements at Salt Lake City during December 1-14, 2018 during inversion (C). Daily $PM_{2.5}$ measurements at Denver during haze, September 1-15, 2017 (D). Each location show measurements before the PA-II PM2.5 measurements were corrected (left) and those after each of the correction options. AMQS unit is represented by the different green lines and the PA-II units, by the different purple lines. Bars represent the standard deviation values per day.

~~**Figure 3**. Hourly $PM_{2.5}$ measurements at Ogden-South Ogden in UT during July 1-11, 2018 (UTC time). Measurements from the AMQS unit are represented in green and those from the PA-II units, in different shades of purple. Each number represent the ID of the unit. Error bars represent the standard deviation values for each hour on each of the PA-II units. Note that local PA-II unit 465 was not active during this time.~~

~~**Figure 4**. Hourly measurements of $PM_{2.5}$ at (A) Ogden-South Ogden, (B) Lindon-Orem, and (C) Salt Lake City during December 1-14 2018 (UTC time). An increase in average daily $PM_{2.5}$ values was observed from December 4-13. The AMQS unit is represented by the different green lines and the PA-II units, by the different purple lines. Each number represents the ID of the unit. Bars represent the standard deviation values per day. Several PA-II units were not operating during these times.~~

~~**Figure 5**. Hourly measurements of $PM_{2.5}$ at (A) Vallejo, (B) Berkeley-Oakland (B), and (C) San Francisco during the November 2018 wildfire (UTC time). An increase in average daily $PM_{2.5}$ values was observed during November 9–20. The AMQS unit is represented by the different green lines and the PA-II units, by the different purple lines. Each number represent the ID of the unit. Bars represent the standard deviation values per day.~~

~~**Figure 6**. Spatial and temporal changes of AQI in California at Berkeley-Oakland, San Francisco, and Vallejo during November 8-22, 2018. Squares represent AQMS and circles, PA-II units. The colors of units represent the different AQI values.~~

Figure 7: Comparison of distance (km) between PA-II to its nearest AQMS in all regions (A) and between each PA-II unit to all other PA-II units per region (B) to R^2, RMSE, MAE and slope values received from the $PM_{2.5}$ hourly measurements comparison.