# PERFORMANCE EVALUATION OF ATMOTUBE PRO SENSORS FOR AIR QUALITY MEASUREMENTS IN AN URBAN LOCATION

**Aishah I. Shittu[1,5], Kirsty J. Pringle[2], Stephen R. Arnold[1], Richard J. Pope[3,4], Ailish M. Graham[1], Carly Reddington[1], Richard Rigby[6], James B. McQuaid[1]**

[1]Institute of Climate and Atmospheric Sciences, University of Leeds, Leeds, UK
[2]Software Sustainability Institute, Edinburgh Parallel Computing Center, University of Edinburgh, Edinburgh, UK
[3]School of Earth and Environment, University of Leeds, Leeds, UK
[4]National Centre for Earth Observation, University of Leeds, Leeds, UK
[5]Science Laboratory Technology Department, Federal Polytechnic Ilaro, Nigeria
[6]Centre for Environmental Modelling and Computation, University of Leeds, UK

Correspondence: Aishah Shittu (eeais@leeds.ac.uk)

**Submitted to Atmospheric Measurement Techniques (AMT)**

## ABSTRACT

This study presents a performance evaluation of eight Atmotube Pro sensors using US Environmental Protection Agency (US-EPA) guidelines. The Atmotube Pro sensors were collocated side-by-side with a reference-grade FIDAS monitor in an outdoor setting for a 14-week period at Leeds city centre, UK. We assessed the linearity and bias for $PM_1$, $PM_{2.5}$ and $PM_{10}$. The result of the $PM_{2.5}$ assessment showed the Atmotube Pro sensors had particularly good precision with a coefficient of variation (CoV) of 28~~3~~%, 18~~5~~% and 15~~3~~% for minutes, hourly and daily $PM_{2.5}$ data averages, respectively. ~~The $PM_{2.5}$ data was cleaned prior to analysis to improve reproducibility between units. 6 out of 8 Atmotube Pro sensor units had particularly good precision.~~ The inter-sensor variability assessment showed two sensors with low bias and one sensor with a higher bias in comparison with the sensor average. Simple univariate analysis was sufficient to obtain good fitting quality to a FIDAS reference-grade monitor ($R^2 > 0.7$) at hourly averages although, poorer performance was observed using a higher time resolution of 15 minutes averaged $PM_{2.5}$ data ($R^2$; ~~0.43-0.54~~ 0.48-0.53). The average error bias, root mean square error (RMSE) and normalized root mean square error (NRMSE) were ~~4.19~~ 3.38 µgm$^{-3}$ and ~~2.17~~ 0.03% respectively. While there were negligible influences of temperature on Atmotube Pro measured $PM_{2.5}$ values, substantial positive biases (compared to a reference instrument) occurred at relative humidity (RH) values > 80%. The Atmotube Pro sensors correlated well with the purple air sensor ($R^2$=0.88~~6~~, RMSE=2.~~85~~ 9 µgm$^{-3}$). In general, the Atmotube Pro sensors performed well and passed the base testing metrics as stipulated by recommended guidelines for low-cost $PM_{2.5}$ sensors. Calibration using multiple linear regression model was enough to improve the performance of the $PM_{2.5}$ data of the Atmotube Pro sensors.

## 1.0 INTRODUCTION

Particulate matter (PM) with an aerodynamic diameter of less than 2.5 µm ($PM_{2.5}$) has been associated with several harmful effects on human health (Maynard et al., 2023; WHO, 2021; Williams et al., 2014). The acute effects of PM include an increase in hospital admissions, early development of asthma in children (Khreis et al., 2019; Mansourian et al., 2011). Long–term effects of outdoor $PM_{2.5}$ are associated with fatal cardiovascular and respiratory diseases and lung cancer with records of increased mortality rates in cities with a higher concentration of airborne PM (WHO, 2021). Another challenge is the exposure disparities amongst socioeconomic groups (Keswani et al., 2022). Understanding the health effects on a given population requires evaluation of their exposure to $PM_{2.5}$, which in turn relies on an understanding of the atmospheric concentration of $PM_{2.5}$. This is challenging as $PM_{2.5}$ concentrations can vary temporally and spatially on small scales (Liu et al., 2009). Low-cost air quality sensors represent recent technologies which are less expensive than typical air quality monitors and allow measurement of specific air pollutants such as PM and other gaseous pollutants. These low-cost sensors are portable allowing ongoing measurements of exposure of individuals as they move around their environments, they also offer an appealing way of obtaining additional atmospheric measurements to better characterise the distribution of $PM_{2.5}$ in a wide range of locations.

Several low-cost sensors ($200 - $2500) have become commercially available (Williams et al., 2014). These sensors are portable in size, lightweight and provide high-resolution data in near real-time (Morawska et al., 2018; Rai et al., 2017). The advent of these low-cost sensors has the potential to change the paradigm of air pollution monitoring as it allows for the possibility of more frequent measurements, which could improve our knowledge, especially in areas where monitoring is sparse and lacks expensive equipment operated by the government or research agencies (Chatzidiakou et al., 2019; Morawska et al., 2018). In addition, these sensors can be used easily without much training, enabling widespread access to air quality data, and making it possible for individuals and communities to monitor air quality both indoors and outdoors by themselves. Recent research has demonstrated that low-cost sensors may be used to identify and apportion various pollution sources in urban environments (Bousiotis et al., 2023; Hagan et al., 2019; Pope et al., 2022; Westervelt et al., 2024; Yang et al., 2022) ~~The use of a network of low-cost sensors is increasing in low-and-middle-income countries (LMIC) countries where a reference-grade monitor for continuous measurement of air pollutants is sparse or lacking~~. Assessing the performance of low-cost sensors and their behaviour relative to reference instruments is crucial, given the growing popularity and use of these sensors for citizen science projects, community engagement initiatives, personal exposure monitoring (Borghi et al., 2017), and building community sensor networks to supplement official reference-grade monitoring networks.

Previous studies have found that some low-cost sensors exhibit significant variation in performance, influenced by several factors such as environmental conditions and choice of reference instrument used. (Kang et al., 2022; Karaoghlanian et al., 2022). Environmental factors such as humidity and temperature have been reported to impact their accuracy (Hagan and Kroll, 2020; Pawar and Sinha, 2020). The results of these evaluations can help determine the suitability of low-cost sensors for measuring pollution in different settings and applications and guide the development of better sensor technologies in the future. Numerous studies have found that some low-cost sensors performed well for measuring ultrafine particles while others were less accurate and had higher measurement variability as reported by (Alfano et al., 2020; Kang et al., 2022). Overall, these studies highlight that careful evaluation of low-cost sensors for particulate pollution measurement is required.

Studies have examined the performance of different brands of low-cost sensors in comparison with a reference grade monitor (Bulot et al., 2019; Feenstra et al., 2019; Jovašević-Stojanović et al., 2015; Sousan et al., 2017) and several calibration methods using linear regression, multiple linear regression, gaussian process regression, ridge regression and random forests have been used to improve the raw $PM_{2.5}$ data (Badura et al., 2019; Barkjohn et al., 2021, 2022; Karaoghlanian et al., 2022; Malings et al., 2019, 2020; Raheja et al., 2023). These calibration methods allow the sensors to be better suited for implementation as a supplement for reference monitors in smaller communities or cities. However, the use of different methodologies developed by various research groups may impact the accuracy and reliability of the data obtained from low-cost sensors (Alfano et al., 2020). Performance evaluation of low-cost sensors for particulate pollution measurements thus far has focused on assessing the accuracy and reliability of low-cost sensors used for measuring particulate matter pollution in field studies, and only a few papers have investigated in detail inter-sensor variability of identical sensor types. Inconsistencies among devices from the same manufacturer might emerge, leading to varying readings under similar conditions. Sensor performance can be highly variable between different devices and end users need to be provided with inter-sensor precision, accuracy, long-term drift and calibration transferability to decide on the right measurement tool for their specific application (Diez et al., 2024). ~~.~~

There are two ways of evaluating the performance of low-cost PM sensors; colocation with a reference instrument and laboratory-based evaluation. The US-EPA (Environmental Protection Agency) refer to this as base testing and enhanced testing respectively (Duvall et al., 2021). This paper focuses on the well detailed metrics for the base testing methods for the performance assessment of Atmotube Pro sensors and the benefits of data cleaning prior to the assessment of the $PM_{2.5}$ data. ~~There are no detailed performance assessment studies available for this sensor model.~~ AQ-SPEC program is a testing centre for low-cost air monitoring sensors to establish performance standards by which low-cost sensors are evaluated both in the field under ambient conditions and laboratory testing under controlled environmental conditions for sensors measuring criteria pollutants (Feenstra et al., 2019; Polidori et al., 2017). Three Atmotube Pro units were previously used in a field evaluation by the well-known South Coast Air Quality Management District (AQMD) which set up the AQ-SPEC (Air Quality Sensor Performance Evaluation

Centre) using the GRIMM and Met-One BAM reference instruments; $R^2 > 0.7$ (AQMD, 2020). The report focused on limited evaluation statistics. Following the AQ-SPEC report in 2020, a few other studies have made use of these Atmotube Pro sensors for occupational and household $PM_{2.5}$ exposure monitoring, and community citizen science (Masri et al., 2022, 2023; Voultsidis et al., 2023; Wang et al., 2020) thus there is a need for a detailed performance assessment on these sensors to ensure confidence in the data being collected.

In summary, low-cost sensors hold great potential to provide widespread useful air quality information for researchers and community members. ~~However, at present there are only limited ways to assess the accuracy of these low-cost sensors due to the absence or low spatial density of reference-grade monitors, especially in LMICs. By demonstrating a good framework for testing the precision, accuracy, and the likelihood of using good sensors in a network of sensors, the results will provide the users with some constraint on the in-situ $PM_{2.5}$ levels measured.~~ The aim of the study is to assess the inter-sensor variability and accuracy of Atmotube Pro sensors to provide an insight on the reliability and robustness of these sensors $PM_{2.5}$ measurements. By demonstrating a good framework for testing the precision, accuracy, and the reliability of sensors within a sensor network, the results will provide the users a clear understanding of the limitations as well as the confidence in the in-situ $PM_{2.5}$ levels measurements obtained for Atmotube Pro sensors. In addition, we investigated the performance of the sensors at higher time resolution (15 minutes) to test the feasibility of their application in capturing short-time events that may be missed at lower resolution.

## 2.0 MATERIALS AND METHODS

### 2.1 Sampling site and data collection

We conducted a sensor colocation exercise aimed at evaluating the performance of Atmotube Pro (manufacturer) sensors compared to a reference monitor placed alongside them. The colocation exercise took place in an ambient environment at Corn Exchange, Leeds city centre (next to a bus stop) where 8 Atmotube Pro sensors and 1 Purple air sensor were collocated side-by-side with a Fine Dust Analysis System (FIDAS 200S) reference-grade air quality monitor in an urban location at the Leeds city centre (53°47'51'N, 1°33'8'W) at a height of about 3.5 metres. The duration of the colocation exercise was done during Autumn from September 26th, 2023, to January 1st, 2024. The city centre is representative of an ideal urban cenntre, which included frequent stops from public buses (vehicular emissions). ~~The colocation exercise took place at the Leeds city centre where 8 Atmotube Pro sensors and 1 purple air sensor were collocated side-by-side with a Fine Dust Analysis System (FIDAS 200S) reference-grade air quality monitor (Leeds City Council) situated at the Corn Exchange at the city centre, Leeds, UK (53°47'51'N, 1°33'8'W). The duration of the colocation exercise was from September 26, 2023, to January 2, 2024.~~

Atmotube Pro is a small and lightweight sensor (0.104 kg) classified as a low-cost device ($250) and is commercially available. Atmotube Pro device have sensirion SPS30 sensors which uses~~ a~~ laser scattering principle to radiate and detect suspended particles in an air chamber. A micro fan draws in air through an inlet, and the air passes through the laser beam where the scattered light reflected by~~off~~ the particles is captured by a photodiode. A signal is transmitted to the micro control unit based on M~~IE~~ie theory where a proprietary algorithm processes the data and supplies outputs for the concentration of the particulate (µgm$^{-3}$). Atmotube Pro sensors report the estimated mass concentration of particles with an aerodynamic diameter of <1µm ($PM_1$), < 2.5µm ($PM_{2.5}$) and <10µm ($PM_{10}$). In addition to the sensirion sensors for PM measurements, the Atmotube device also contains BOSCH BME280 sensors for measuring temperature and relative humidity values. The sensors also log data every second and store it in memory every minute (Atmotube, 2023). One of the limitations of the Atmotube Pro device is the data retrieval memory with limited history size of 10 days after which data not downloaded would be overwritten.

The Purple Air sensors contain two Plantower PMS5003 sensors, which record two-minute averaged data. The Purple Air sensor uses a similar principle to the Atmotube Pro sensors described above, based on scattering of laser light. The Plantower sensors also estimate mass of particles with aerodynamic diameters <1 µm, <2.5 µm and <10 µm which are reported as cf_1 and cf_atm which both have channels

A and B in the Purple Air dataset. The cf_atm data is displayed on the Purple Air map (BarkJohn, 2021) and this sensor input is the dataset used in this study. The Purple Air sensor was deployed at the colocation site since June 2022.

The reference monitor used for the study was a FIDAS 200S consisting of a sampling head that also enables representative sampling in strong wind. The control unit is integrated in an IP 65 weather protected housing which can be set up as a standalone outdoor instrument. It uses optical light scattering according to M~~IE~~ie theory using bright and durable white LED light as a light source. It measures $PM_1$, $PM_{2.5}$, $PM_4$, $PM_{10}$, TSP, temperature, and relative humidity parameters. The measuring range in mass is 0-10,000 µgm$^{-3}$ and particle size range is 0.18-18 µm. The monitor records 15-minutes averages. The FIDAS 200S are certified and developed for compliance monitoring of PM in accordance with EU and UK legislation. The uncertainty between FIDAS devices is 0.44 µgm$^{-3}$ (FIDAS, 2024).

The performance of the low-cost sensors will be assessed using US-EPA guidelines (Duvall et al., 2021) for base and enhanced testing metrics. For sensor accuracy, coefficient of determination ($R^2$), root mean square error (RMSE), mean normalised bias (MNB), normalized root mean square error (NRMSE), slope and intercept will be determined. The $R^2$ value is a metric that provides information about the proportion of the variance in the dependent variable (Atmotube Pro sensor) that can be explained by the independent variable (reference monitor). The RMSE helps to understand the error associated with sensor $PM_{2.5}$ concentration in comparison with the reference concentration. For sensor precision, the standard deviation (SD) and coefficient of variation (CoV) will be determined. It is recommended that low-cost sensor used for performance evaluation test should have 75% data completeness during the colocation study period (Duvall et al., 2021; Zimmerman, 2022). Other performance metrics include detection range, detection limit and response time.

This paper focuses on the reproducibility of the 8 Atmotube Pro sensor units (identical model) and developing an appropriate data cleaning method for the obtained $PM_{2.5}$ data. CoV was calculated using one-minute, fifteen-minutes and one-hour averages. Low CoV values indicate high reproducibility in the measurements across the Atmotube Pro sensors units. US-EPA recommends CoV of <30% between sensors of identical models.

### 2.2 Atmotube Pro Quality Assurance ~~and Data Cleaning~~

The Atmotube Pro came assembled and needs to be charged frequently. The sensor requires a charging time of about 2.5 hours. The battery requires daily charging when set to “always on” mode thus we left the sensor plugged in throughout the entire duration of the study alongside the Purple Air and reference monitor. The Atmotube Pro also stores historical data in an onboard flash memory when not connected to a smartphone. The historical data can be transferred to a smartphone during data synchronization each time the sensor is connected to a smartphone. The Atmotube Pro sensors are designed for mobile monitoring, and to protect the sensors from rain at the colocation site, a makeshift cover was used to enclose all the sensors used in the study.

One-minute $PM_1$, $PM_{2.5}$, $PM_{10}$ relative humidity and temperature data were retrieved from 8 Atmotube Pro sensors. Data completeness as shown in equation 1 for Atmotube Pro sensors is the percentage ratio of minute-wise data available for each sensor and the total number of minutes expected for the study period (Polidori, et al., 2017). This ranged from 73-84% for $PM_1$, $PM_{2.5}$, $PM_{10}$. ~~-~~Preliminary analysis ~~focused~~ investigated ~~on~~ the performance of $PM_1$, $PM_{2.5}$ and $PM_{10}$ size distributions using Atmotube Pro sensors and reference FIDAS monitor data. To understand reproducibility between Atmotube Pro sensor units, ~~.~~ o~~O~~ne minute, fifteen minutes and one-hour averaged $PM_{2.5}$ outputs were used for calculating a coefficient of correlation (r) between Atmotube Pro sensors ~~for the raw $PM_{2.5}$ data (Fig. 1~~)~~. $PM_{2.5}$ data filtering was achieved by eliminating all data at each time stamp where 4 or more sensors had missing data. Where there were sufficient sensor measurements, the sample $PM_{2.5}$ mean and SD values were derived to calculate the CoV (as in Equation 1). Where there were large CoV values (e.g. > 100%), it indicates large variability between the sensors (i.e. one or more sensors had anomalously large $PM_{2.5}$ values). If the CoV values tend towards zero, it indicates low variability between the sensors and a more homogeneous sample. There is no standard for this method of data cleaning, however by utilising the~~

~~CoV method, using minute-wise data, it allowed for the removal of major anomalies in the sensor data, while retaining a good degree of data coverage~~.

$$\cancel{CoV}\text{Data recovery} = \frac{N\ valid\ data\cancel{SD_{mi}}}{N\ study\ period\cancel{\mu_{mi}}} * 100 \qquad (1)$$

~~Where; $SD_{mi}$ = Standard deviation of $PM_{2.5}$ concentration for Atmotube Pro sensors for each minute (μgm$^{-3}$)~~

~~$\mu_{mi}$ = Mean of $PM_{2.5}$ concentration for 8 Atmotube Pro sensors for each minute (μgm$^{-3}$)~~

where ; Nvalid= number of valid sensor data points during test period

Ntest= total number of data points for the study period

~~The CoV ranged from 0 to 244% and a threshold <50% was used to filter the $PM_{2.5}$ data. Lower CoV values indicate higher reproducibility between sensor units. Applying this filter resulted in a PM2.5 data loss of 3.8% temporally.~~

## 3.0 RESULTS AND DISCUSSION

During the colocation period, there were some data gaps, mostly due to failure to download data within the 10-day data buffering period of the internal sensor storage. The Atmotube Pro device erases old records to create room for new ones after storing a maximum of 14,400 data points, or 10 days. The sensors were connected to external power continually throughout the study period. The data downloader tool allows fetching data from the sensor unit for a period of up to 7 days via a simple user interface (Atmotube, 2023) and this is a limitation for long-term data collection due to limited space. This indicates that these sensors need frequent data download to avoid data loss.
The performance of Atmotube Pro sensors for $PM_1$, $PM_{2.5}$ and $PM_{10}$ in comparison with the reference FIDAS monitor are shown in Figure 1. The average of all 8 sensors was computed at hourly time resolution. $PM_1$ had a very low error bias of 1.7 μgm$^{-3}$ and a strong $R^2$ of 0.94. $PM_{2.5}$ had a larger error bias of 3.2 μgm$^{-3}$ and a decrease in the $R^2$ value to 0.86 in comparison to $PM_1$. The poorest performance was recorded for $PM_{10}$ with a larger error bias of 6.2 μgm$^{-3}$ and a further decline in $R^2$ of 0.49. Similar result were recorded in the study by Molina et al., 2023 using Plantower, Sensirion and Piera low-cost sensors. The rest of the paper will focus on particle size <2.5 μm ($PM_{2.5}$) as $PM_{2.5}$ is the key standard by WHO and other regulatory agencies for health-related research.

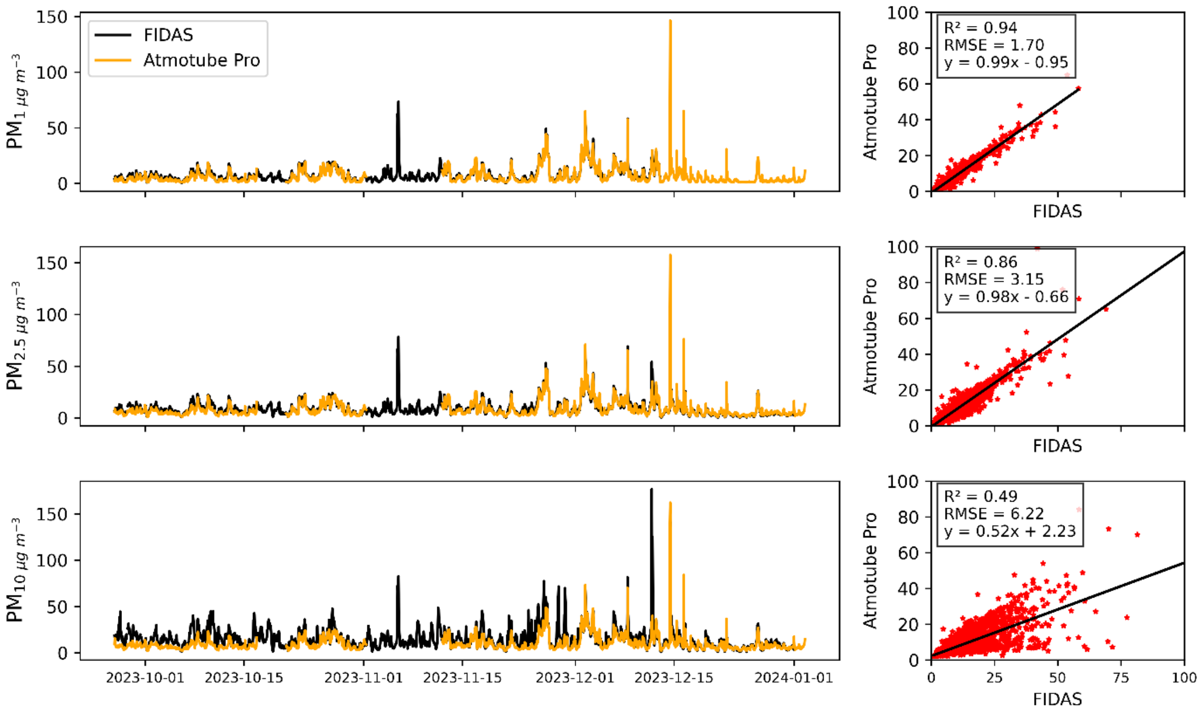



Figure 1: Comparison of Atmotube Pro and FIDAS reference monitor data for $PM_1$, $PM_{2.5}$ and $PM_{10}$ (hourly averaged data).

### ~~3.1 Data Cleaning~~

~~-~~We have approximately three months of observational data with $PM_{2.5}$ in a range of 1 - ~~500~~ 120 µg $m^{-3}$ a(using hourly averaged data) as shown in Fig. ~~1~~2(a). During this period, there are spikes in the order of ~~~200~~50 µg $m^{-3}$ and above which is probably due to episodic events such as buses driving past, tobacco smoke and the annual Guy Fawkes bonfire night~~.~~. The sensors exhibit comparable temporal variability between the sensors, however, in absolute terms Sensor 6 (S6) has higher concentrations in some cases. We correlated minute-wise data of each sensor against the other and the coefficient of correlation ranged generally between 0.8 to nearly 1.0. Sensor 5 (S5) had slightly lower values of about 0.8 - 0.9, while S6 clearly was the poorest or the outlier because the r value was between 0.5 - 0.6. This suggests some anomalous data recorded as shown in Fig 2, highlighting a fault in sensor S6., ~~thus requiring a data filtering method as shown in Eq. (1). Fig. 1 demonstrates the usefulness in removing the erroneous spikes, particularly in S6, where other sensors did not exhibit these, strongly suggesting that the high values S6 had recorded were erroneous. After the data cleaning, we can clearly see in Fig 1 (c) that removing the erroneous data marked an improvement in the agreement between the sensors. S6 is clearly the outlier but the r value was improved by 0.2 with r >0.74. S5 is much more closely aligned with the other sensors of r > 0.89, indicating an increase in the r value by 0.05. There is an overall benefit of applying the filtering and the filtered data was used for the scientific analysis.~~

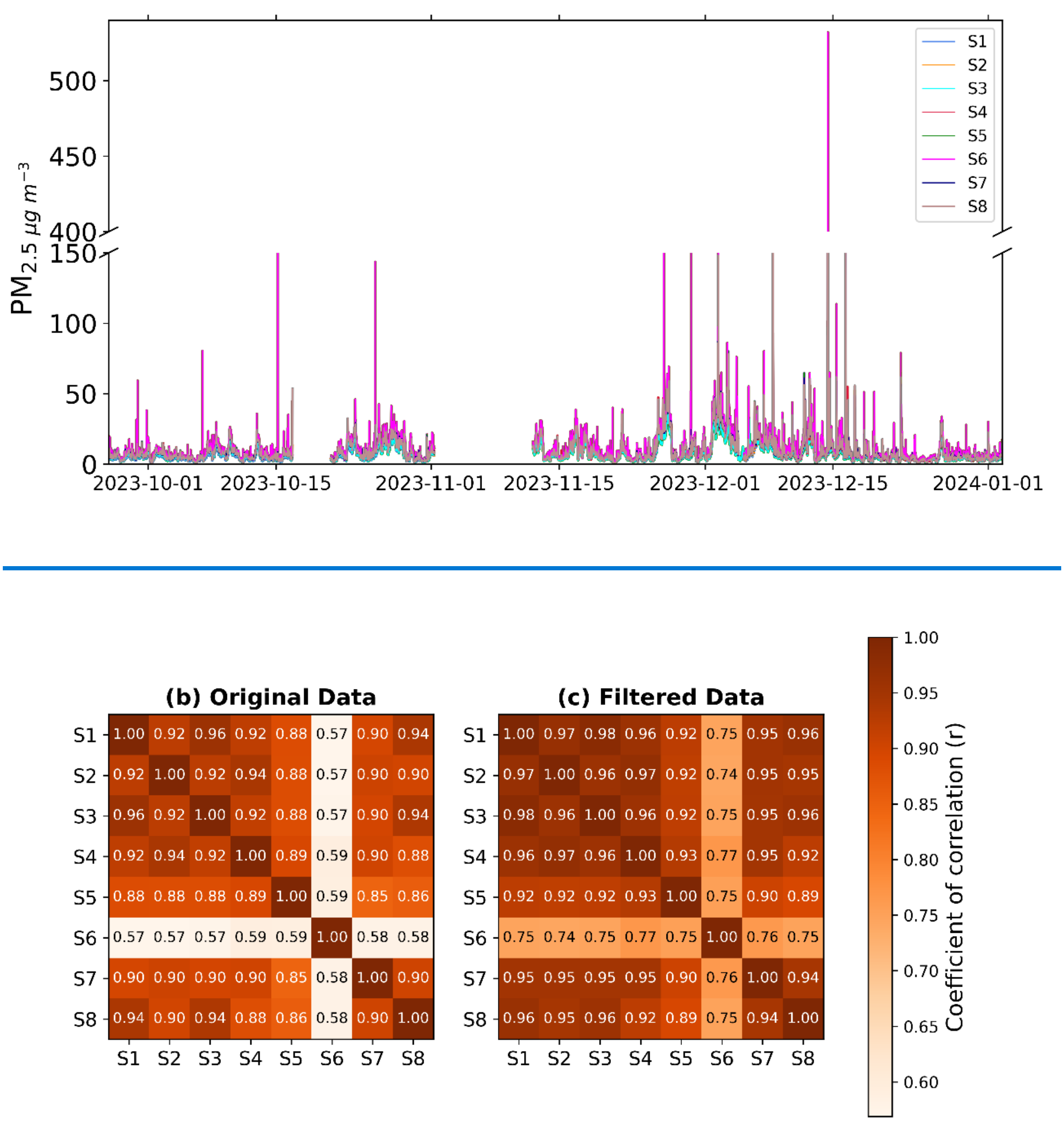



Figure 1: $PM_{2.5}$ data of 8 Atmotube Pro sensors (a) showing the time series of each sensor, (b) coefficient of correlation (r) for raw $PM_{2.5}$ data indicating inferior performance of S6 in comparison to others and (c) improved r values for all sensors after data cleaning.

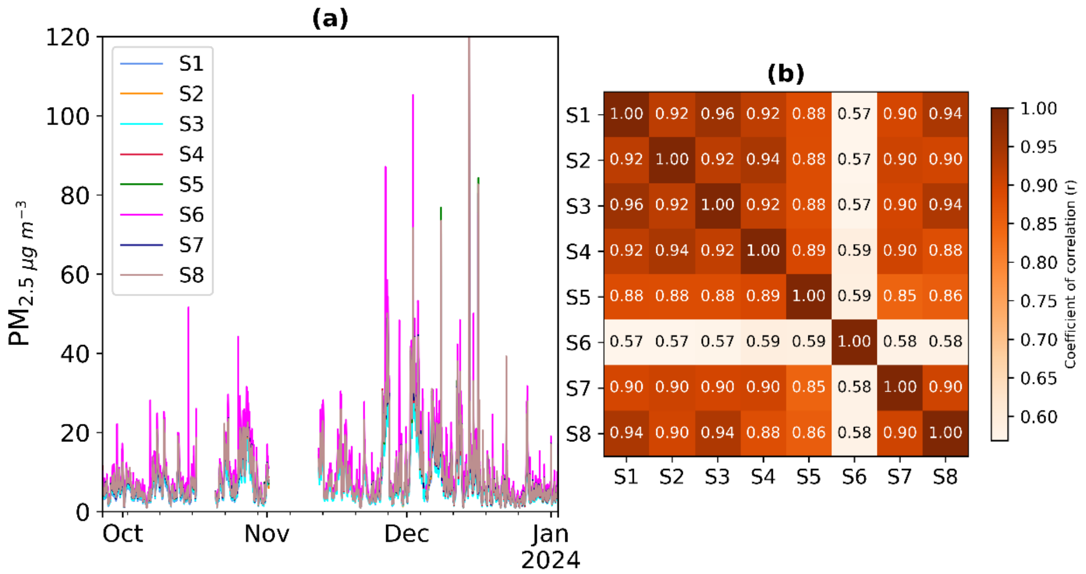
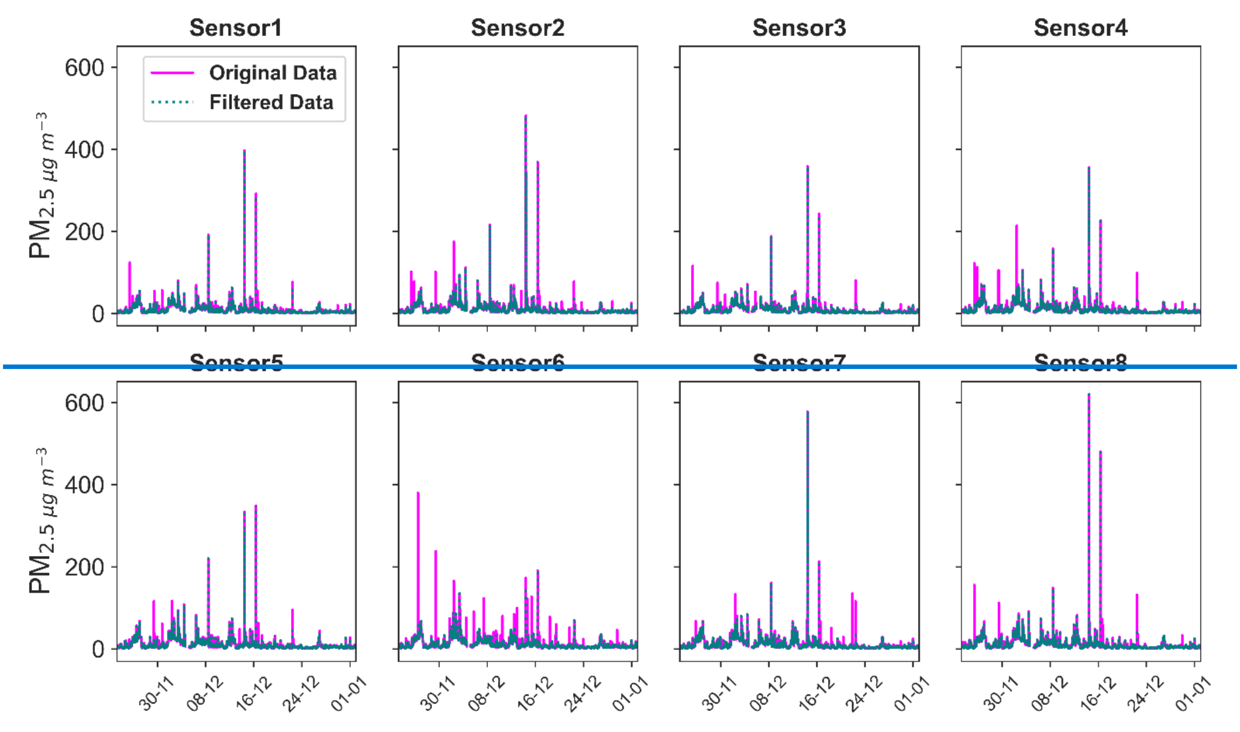



Figure 2: $PM_{2.5}$ data of 8 Atmotube Pro sensors (a) showing the time series of each sensor (hourly average), (b) coefficient of correlation (r) (minute-wise data)

Most air quality networks implement regular quality assurance and control measures, although outliers can still happen because of sensor malfunctions or differences in monitoring configuration such as inlet orientation. There are possibilities of using some of these low-cost sensors where there are no reference monitors present, but it is imperative to check if a network of low-cost sensors have malfunctioning sensors. Outliers pose challenges for statistical analysis. S6 was the sensor with the most erroneous $PM_{2.5}$ data, which showed anomalous, data spikes (low and high) relative to the other sensors. ~~The importance of data cleaning is illustrated in Fig. 2 and the time series focused on data from November 2023 to January 2024, highlighting the comparison between raw and filtered~~ ~~$PM_{2.5}$~~ ~~data. Figure 2 shows how data cleaning has improved the time series for the filtered data among all sensors in comparison to the original data where there were differences in the concentrations reported by S6 in comparison with others. S6 had~~ ~~$PM_{2.5}$~~ ~~values as high as 400 µg~~ ~~$m^{-3}$~~ ~~while other sensors had values less than 200 µg~~ ~~$m^{-3}$~~ ~~on the 29th of November. There were other instances such as 15th December where S6 misses large concentrations while other sensors recorded high values as high as 600 µg~~ ~~$m^{-3}$~~~~.~~




~~Figure 2: Comparison of raw and filtered data highlighting the effectiveness of anomalous spike removal, most evident in Sensor 6 (S6)~~

~~The effectiveness of the data cleaning was evaluated using a coefficient of correlation between the sensor before and after data cleaning. The coefficient of correlation for the raw $PM_{2.5}$ data showed r value that ranged from 0.57-0.96 as shown in Fig. 1(b) while that of the filtered method (CoV using threshold of <50%) improved the coefficient of correlation (r) to 0.74-0.98~~. One of the important benefits of the performance evaluation assessment for multiple sensors is to identify less robust individual sensors in a sample of sensors.

### ~~3.2~~3.1 Inter-sensor variability

Using methods stated in the US-EPA guidelines for low-cost performance metrics, results showed that the SD metrics in this study just exceeded the US-EPA recommended limit of <5 $\mu gm^{-3}$ while the CoV was below recommended limit of <30% as shown in Table 1. For this section, the CoV for determining inter-sensor variability is calculated as described by (Duvall et al., 2021; Zimmerman, 2022). This indicates reasonable variability in sensors of identical models and the high SD values can be attributed to high concentration short events such as Christmas market barbeques, smoking next to the sensors. Although the CoV values are within the recommended limits the values are higher than values seen in the report made by the South Coast AQMD (AQMD, 2020) where the relative inter-sensor variability (CoV) was 6.7% and the standard deviation was also reported to be 0.57 $\mu gm^{-3}$ for $PM_{2.5}$ values using 5 minutes averages of 3 Atmotube Pro sensors. There is a difference in the environment, duration of the study and the $PM_{2.5}$ concentrations. For the AQ-SPEC, the collocation was for done for a 2-week period and the 5-minutes averages had a maximum of 50 $\mu gm^{-3}$. This suggests the Atmotube Pro inter-sensor variability is less at lower $PM_{2.5}$ concentrations. Previous studies have reported CoV <10% for Plantower sensors (Badura et al., 2019; Zimmerman, 2022) while other models of low-cost sensors have also reported a higher CoV > 25% for Dylos (Carvlin et al., 2017), Plantower and Syhitech (indoor colocation) had CoV > 30% (Zamora et al., 2020).

~~Previous studies have reported CoV <10% for Purple Air (Zimmerman, 2022) and Plantower (Badura et al., 2019). Others models of low-cost sensors have reported a higher CoV >25% for Dylos (Carvlin et al., 2017), Plantower and Syhitech (indoor colocation) had CoV > 30% (Zamora et al., 2020).~~

Table 1: Coefficient of variation and standard deviation between 8 Atmotube Pro sensors

| $PM_{2.5}$ data | | CoV (%) | SD ($\mu gm^{-3}$) |
|---|---|---|---|
| Raw data | 1 minute | 27.8 | 12.2 |
| | 1 hour | 17.7 | 8.8 |
| | 1 day | 15.0 | 6.0 |
| ~~Filtered data~~ | ~~1 minute~~ | ~~22.7~~ | ~~11.3~~ |
| | ~~1 hour~~ | ~~14.8~~ | ~~9.1~~ |
| | ~~1 day~~ | ~~13.3~~ | ~~5.8~~ |

~~For the filtered data, t~~The CoV reduced further when Atmotube Pro sensor “S6” was removed from the analysis; CoV=~~18~~20%, 11% and 10% for minute, hourly and daily averages, respectively. This compares well with the range of CoV values from field evaluation results of different low-cost sensors of 0.9 to 31.0% with an average of 12.8% for 24-hour averages as described by (Duvall et al., 2021) using resources from AQ-SPEC sensor evaluation, US-EPA sensor evaluations and peer reviewed literature. Atmotube Pro sensors sit well within this range for both lower and higher resolution ~~raw and filtered~~ $PM_{2.5}$ ~~24-hour averaged~~ data of ~28~~15.0~~%, 18% and ~~13.3~~15% for minute, hourly and daily averaged data respectively.

Our results indicate one anomalous sensor can drive an increased inter-sensor variability in the measurements for the Atmotube Pro sensors. More research is required to identify the minimum number of sensors needed for a performance evaluation assessment. ~~These commercially available Atmotube Pro sensors are factory calibrated and it is possible that some sensors were not calibrated as precisely as others resulting in the variation in their measurements and contributing to high CoV~~ There were inconsistencies observed among Atmotube Pro sensors leading to varying readings under same conditions thus contributing to high CoV. There is also the possibility of environmental factors such as relative humidity and temperature measurements, which may have influenced the $PM_{2.5}$ values of these sensors differently.

To further investigate reproducibility of the sensors, hourly time-step of the $PM_{2.5}$ average ($Avg_h$) of all 8 sensors was derived over the study period. For each sensor, the ratio between the sensor value (per hour time step) and the multi-sensor mean was calculated as in Eq. (2). The temporal distribution of these ratios for each sensor was illustrated using box and whisker diagrams as in Fig. 3 to provide an indication of the sensor-sensor precision.

$$\text{Sensor/Average Ratio} = \frac{x_{hi}}{\mu_h} \qquad (2)$$

Where; $x_{hi}$ = Atmotube Pro $PM_{2.5}$ hourly data ($\mu gm^{-3}$) for sensor (i) where i=1,8;

$\mu h$ = Mean PM2.5 $\mu gm^{-3}$ concentration hour (for all 8 sensors).

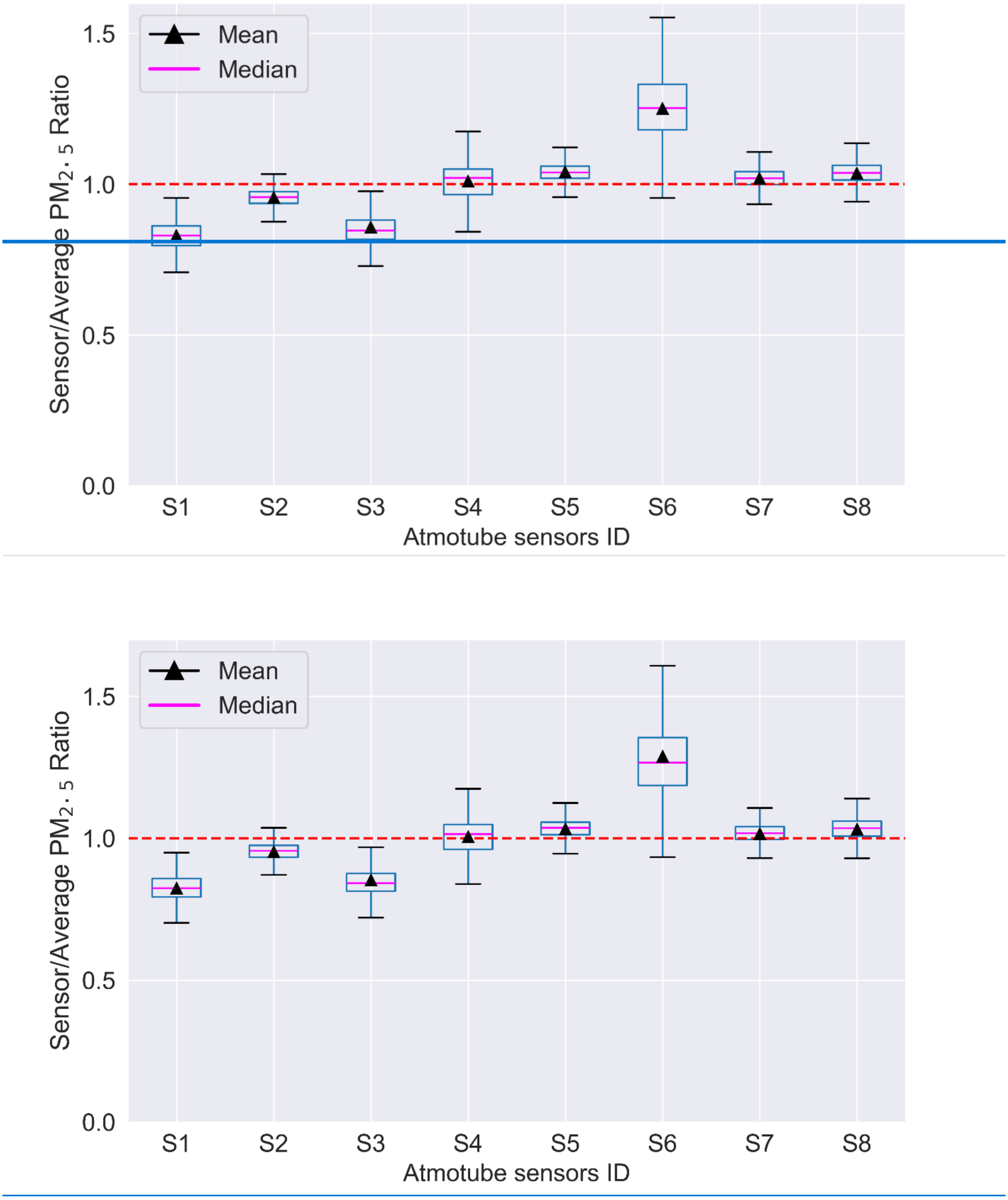
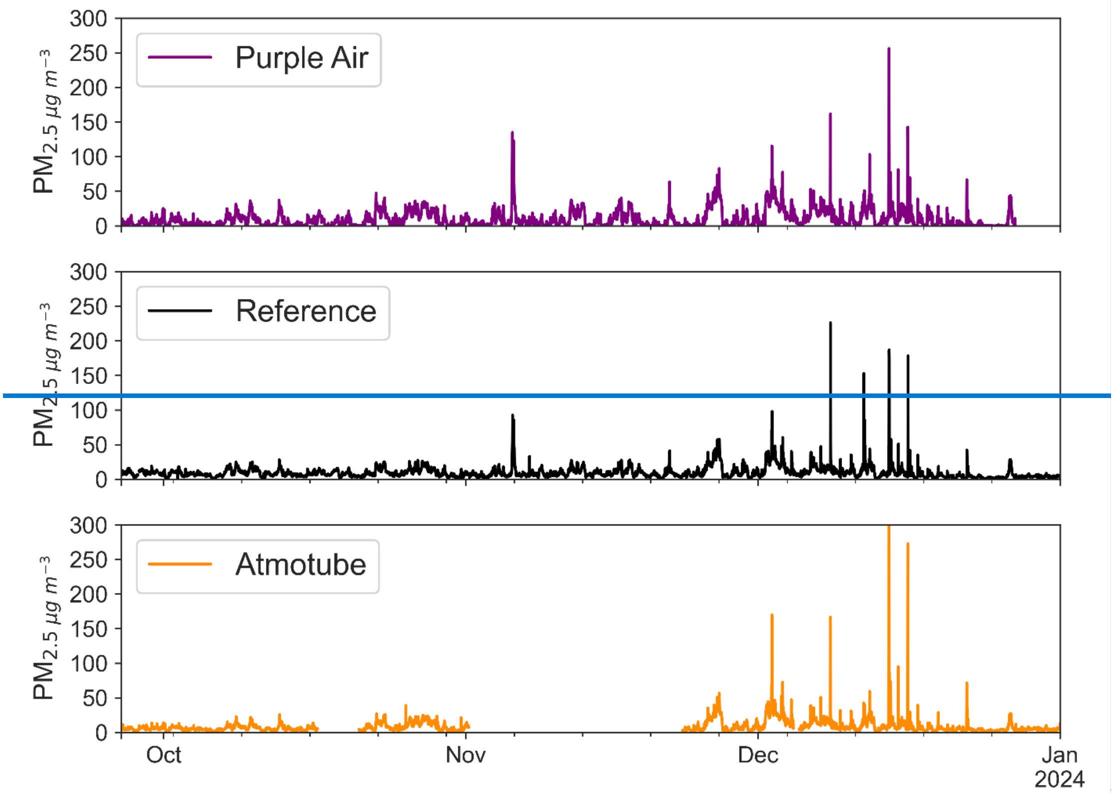



Figure 3: Sensor-sensor precision comparing the ratio of sensor hourly $PM_{2.5}$ values to the 8 Atmotube Pro multi-sensor average as a reference. The whiskers represent the 5$^{th}$ and the 95$^{th}$ percentile. Red dashed line indicates sensor/average ratio of 1 where <1 represents low bias and vice versa.

Although there is no standard on what the sensor precision should be, this investigation makes it clear that sensor S6 can be termed to have large deviation from the average. S6 had a median ratio of ~1.25 and 25$^{th}$-75$^{th}$ range of 1.2-1.3. Sensors S1 and S3 have a small deviation from the average $PM_{2.5}$ values (that is, median of ~0.8 and 25$^{th}$-75$^{th}$ range of 0.7-0.95). Note that the hourly time averages were used for Fig. 3. For inter-sensor quality assurance check where a reference grade instrument is far away or totally lacking, comparing against the mean $PM_{2.5}$ value for all the sensors may prove useful to identify faulty sensors within a network of sensors as shown in Eq. (2). Where the sensor $PM_{2.5}$ median ratio value tends to 1.0, it indicates the sensor measurements are consistent with the majority of the other sensors in the network (Figure 3). Sensors (S2, S4, S5, S7 and S8) may be used as a "supplementary reference" to identify potential anomalous sensors. In summary, 62.5% of the sensors used for the study exhibited greater precision in their measurements.

### ~~3.3~~3.2 *Comparison with a reference-grade monitor*

Sensor performance was investigated further by comparing the $PM_{2.5}$ Atmotube Pro sensor data to measurements from the Local Authority reference monitor data at the Leeds city centre air quality monitoring site. Atmotube 15-minute averaged data were used for this comparison as the reference monitor logs $PM_{2.5}$ data at this temporal resolution.



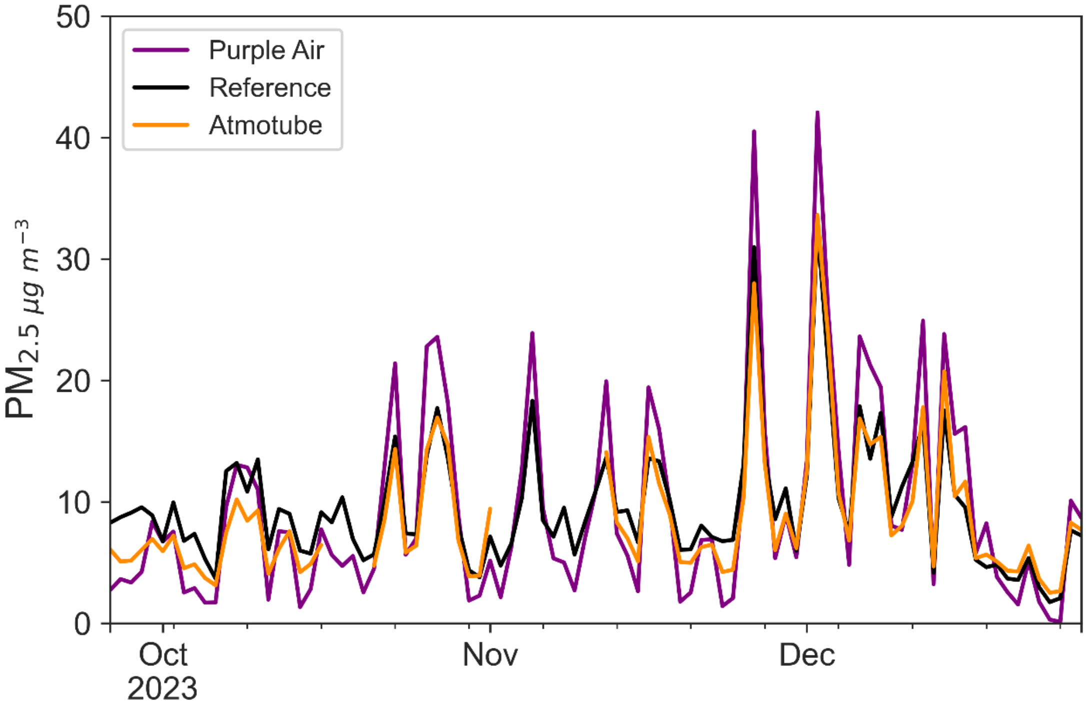



Figure 4: Time series of $PM_{2.5}$ concentration reported by Purple Air, average Atmotube Pro sensors and the reference monitor. The data has been averaged to ~~15~~daily data ~~minutes~~.

The time series in Fig. 4. shows the Atmotube Pro sensors and the purple air sensor captured the reference monitor $PM_{2.5}$ temporal variability and the low-cost sensor $PM_{2.5}$ values are of the same order of magnitude ~~at lower concentrations (<50 µgm$^{-3}$)~~. However, during some high concentration episodes, the ~~Atmotube Pro~~ Purple Air sensors typically overestimated $PM_{2.5}$ values in comparison with the reference.

We calculated the $R^2$, RMSE, NRMSE, MNB, slope and intercept of the relationship between the Atmotube Pro sensor data and the reference monitor. The results for fifteen minutes and hourly averages are summarized in Table 2. For the coarser time resolution, Atmotube Pro sensors had $R^2$ (>0.7) and RMSE (< 7 µgm$^{-3}$) in comparison to the fifteen minutes averages where $R^2$ ~~(0.42-0.54)~~ (0.48-0.53) and RMSE (> 7 µgm$^{-3}$). Using hourly $PM_{2.5}$ averages, the Atmotube Pro and purple air sensors performed well with evaluation metrics within the US-EPA guideline values, with RMSE values of ~~4.19~~3.4 µgm$^{-3}$ and 4.8~~6~~ µgm$^{-3}$, respectively.

Results show that the hourly averaged data of the Atmotube Pro sensors performed better than the higher time resolution data. In comparison with the AQ-SPEC evaluation, Atmotube Pro sensors had $R^2$ ~0.79 and 0.89 using BAM and GRIMM reference monitors respectively (AQMD, 2020).

The $R^2$ values can be further improved by calibration methods as reported in the literature with different calibration and correction methods (Badura et al., 2019; Giordano et al., 2021; Hong et al., 2021; Pawar and Sinha, 2020) to improve the $PM_{2.5}$ data quality of different low- cost sensor models. ~~The focus of this paper is not to investigate correction or calibration as this has been well established in other similar low-cost sensors, but to show the overall skill of the Atmotube Pro sensors.~~

Table 2: Accuracy metrics using Atmotube Pro and Purple Air sensors in comparison with reference data at 15-minutes and hourly averaging time. $R^2$ = correlation of determination, "RMSE" = root mean square error; "NRMSE" = normalized root mean square error, "a" = slope; "b" = intercept, "S1-S8" = Atmotube Pro sensors, "Mean" = Atmotube Pro sensor average, "PA" = Purple Air sensor.

| | 15-minutes Average $PM_{2.5}$ $\mu gm^{-3}$ | | | | | Hourly Averaged $PM_{2.5}$ $\mu gm^{-3}$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **$R^2$** | **RMSE $\mu gm^{-3}$** | **NRMSE (%)** | **a** | **b** | **$R^2$** | **RMSE $\mu gm^{-3}$** | **NRMSE (%)** | **a** | **b** |
| S1 | 0.43 | 7.93 | 2.54 | 0.72 | 0.31 | 0.74 | 4.1 | 2.19 | 0.94 | -1.86 |
| S2 | 0.42 | 9.31 | 2.46 | 0.83 | 0.4 | 0.78 | 4.78 | 2.16 | 1.08 | -2.13 |
| S3 | 0.43 | 7.69 | 2.55 | 0.7 | 0.64 | 0.79 | 3.89 | 2.14 | 0.91 | -1.48 |
| S4 | 0.46 | 8.76 | 2,57 | 0.85 | 0.7 | 0.82 | 4.29 | 2.09 | 1.1 | -1.87 |
| S5 | 0.44 | 9.32 | 2.78 | 0.87 | 0.81 | 0.8 | 4.7 | 2.33 | 1.13 | -1.89 |
| S6 | 0.54 | 7.62 | 2.57 | 0.87 | 2.23 | 0.79 | 4.52 | 3.45 | 1.07 | 0.19 |
| S7 | 0.48 | 8.45 | 2.36 | 0.63 | 0.63 | 0.83 | 4.18 | 2.02 | 1.11 | 1.88 |
| S8 | 0.44 | 9.54 | 2.37 | 0.89 | 0.58 | 0.8 | 4.86 | 2.10 | 1.16 | -2.15 |
| Mean | 0.47 | 8.31 | 2.55 | 0.82 | 0.79 | 0.82 | 4.19 | 2.17 | 1.06 | -1.63 |
| PA | 0.56 | 8.78 | 3.43 | 1.05 | -0.11 | 0.85 | 4.86 | 3.25 | 1.37 | -3.49 |

| | 15-minutes Average $PM_{2.5}$ $\mu gm^{-3}$ | | | | | Hourly Averaged $PM_{2.5}$ $\mu gm^{-3}$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **$R^2$** | **RMSE $\mu gm^{-3}$** | **NRMSE (%)** | **a** | **b** | **$R^2$** | **RMSE $\mu gm^{-3}$** | **NRMSE (%)** | **a** | **b** |
| S1 | 0.51 | 6.05 | 0.02 | 0.65 | 0.92 | 0.85 | 2.94 | 0.02 | 0.85 | -1.01 |
| S2 | 0.51 | 7.02 | 0.02 | 0.75 | 1.13 | 0.85 | 3.42 | 0.02 | 0.97 | -1.08 |
| S3 | 0.50 | 5.91 | 0.02 | 0.64 | 1.23 | 0.85 | 2.85 | 0.02 | 0.83 | -0.63 |
| S4 | 0.53 | 6.86 | 0.02 | 0.78 | 1.35 | 0.87 | 3.23 | 0.02 | 1.01 | -0.91 |
| S5 | 0.51 | 7.45 | 0.02 | 0.81 | 1.27 | 0.85 | 3.63 | 0.02 | 1.05 | -1.11 |
| S6 | 0.48 | 9.73 | 0.02 | 0.99 | 1.42 | 0.77 | 5.20 | 0.03 | 1.15 | -0.15 |
| S7 | 0.52 | 7.11 | 0.02 | 0.79 | 1.27 | 0.86 | 3.42 | 0.02 | 1.03 | -1.04 |
| S8 | 0.51 | 7.56 | 0.02 | 0.82 | 1.22 | 0.85 | 3.64 | 0.02 | 1.07 | -1.21 |
| Mean | 0.54 | 6.74 | 0.02 | 0.78 | 1.23 | 0.86 | 3.38 | 0.02 | 0.99 | -1.63 |
| PA | 0.58 | 8.46 | 0.03 | 1.07 | -0.45 | 0.85 | 4.79 | 0.03 | 1.37 | -3.47 |

A measure of correlation is necessary when assessing performance of low-cost sensors, but alone is not sufficient as the error bias should also be reported (Giordano et al., 2021). The AQ-SPEC, however, did not mention the error bias of the Atmotube Pro sensors in its report. The RMSE describes the difference between sensors $PM_{2.5}$ measurements and the true value (reference instruments). The NRMSE accounts for testing in conditions where high $PM_{2.5}$ concentrations were recorded and the RMSE is normalized using the average of the reference $PM_{2.5}$ measurements over the testing period (Duvall et al., 2021; Zimmerman, 2022). The RMSE and NRMSE values as shown in Table 2 were within the recommended US-EPA guidelines of RMSE <7 $\mu gm^{-3}$ and <30% respectively using hourly averaged $PM_{2.5}$ data. The MNB is a model evaluation metric which helps to quantify the accuracy of the measurements over the collocation period (Giordano et al., 2021). The MNB values for the 15-minute average Atmotube Pro and the purple air sensors were -0.~~0~~ 17 and ~~0.56~~-2.7 respectively. For one-hour averages, MNB for the average Atmotube Pro and the purple air sensors were 0.01~~17~~ and ~~2.73~~ -0.29 respectively.

The US-EPA guideline also recommends a target slope and intercept range 1.0±0.35 and -5 to +5, respectively. The slope and intercept of the Atmotube Pro sensors had an average of ~~1.06~~ 0.99 and -1.63 respectively while the values for the purple air sensor were 1.37 and -3.47~~9~~, respectively. The overall performance of the 8 Atmotube Pro sensors is summarized in Table 3. The Atmotube Pro sensors met the USEPA base testing criteria (precision, bias, linearity, and error) at coarser resolution averages (one-hour). However, the linearity and the error did not meet these criteria at lower resolution averages (fifteen-minute). Also, at $PM_{2.5}$ concentration below 100 µgm$^{-3}$ for lower resolution averages, the criteria were met for these metrics indicating the Atmotube sensors perform better at lower concentrations.

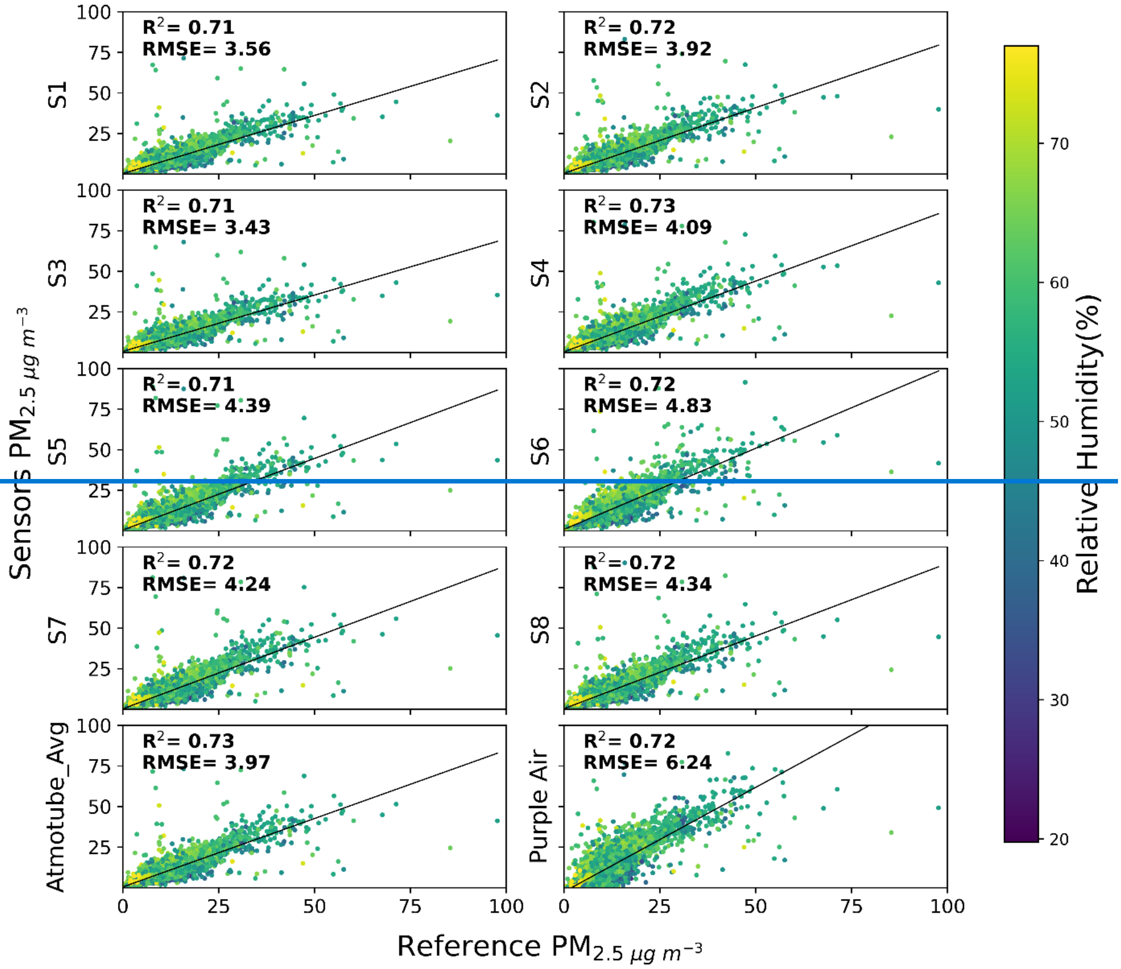
Table 3: Overview Performance Summary of reproducibility and accuracy among identical Atmotube Pro sensors using US-EPA guidelines.

| **Performance metrics (US-EPA)** | | **Target values** | **Atmotube Pro sensors ($PM_{2.5}$ values)** | |
|---|---|---|---|---|
| Base Testing | | | 15-minutes | 1-hour average |
| Precision | SD | <5 µgm$^{-3}$ | **Failed** | **Failed** |
| | CoV | <30% | **Passed** | **Passed** |
| Bias | Slope | 1 ± 0.35 | **Passed** | **Passed** |
| | Intercept | -5 ≤ b ≤ +5 | **Passed** | **Passed** |
| Linearity | $R^2$ | ≥ 0.7 | **Failed** using full dataset (~~$R^2$ 0.44-0.56)~~($R^2$ 0.48-0.53)<br><br>**Passed** at $PM_{2.5}$ values <100 µgm$^{-3}$ ($R^2$ 0.72 - 0.75) | **Passed** |
| Error | RMSE | ≤7 µgm$^{-3}$ | **Failed** using full dataset (RMSE 7.6-9.2 µgm$^{-3}$) (RMSE 5.9-9.7 µgm$^{-3}$)<br><br>**Passed** at $PM_{2.5}$ values <100 µgm$^{-3}$ (RMSE 3.3 - 4.6 µgm$^{-3}$ | **Passed** |
| | NRMSE | ≤ 30% | **Passed** | **Passed** |

### 3.2~~3~~.1 Separating high concentration events.

The performance of the 8 Atmotube Pros showed the $R^2$ using 15-minute averaged $PM_{2.5}$ data were well correlated at $PM_{2.5}$ concentration below 100 µgm$^{-3}$. $R^2$ > 0.7 for all 8 Atmotube Pro sensors and the purple air sensor as shown in Fig. 5. Although, correlation was low using the full dataset ($R^2$ range 0.42 to 0.56) for 15–minute averaged as seen in Table 2, this is indicative of poorer performance at higher concentrations above 100 µgm$^{-3}$. At higher averaging time the $R^2$ improved significantly, and this is in line with a report by (Hong et al., 2021) using Sensirion, Plantower and Honeywell sensors.
Comparing the error bias in the regression analysis of the 15-minutes averaged data of the full $PM_{2.5}$ dataset (1-300 µgm$^{-3}$) and $PM_{2.5}$ dataset below 100 µgm$^{-3}$ only, the RMSE range was 7.6-9.5 µgm$^{-3}$ and 3.56-4.83 µgm$^{-3}$ respectively. This shows a general lower bias in error at lower concentrations between the Atmotube Pro sensors and the reference. The same applies to the PA sensor, as there was also a reduction in RMSE values from 8.8 to 6.2 µgm$^{-3}$. The plot in Fig. 5. was coloured by individual RH data

logged by each sensor. Section 3.4 highlights the influence of RH and temperature on the sensor data.



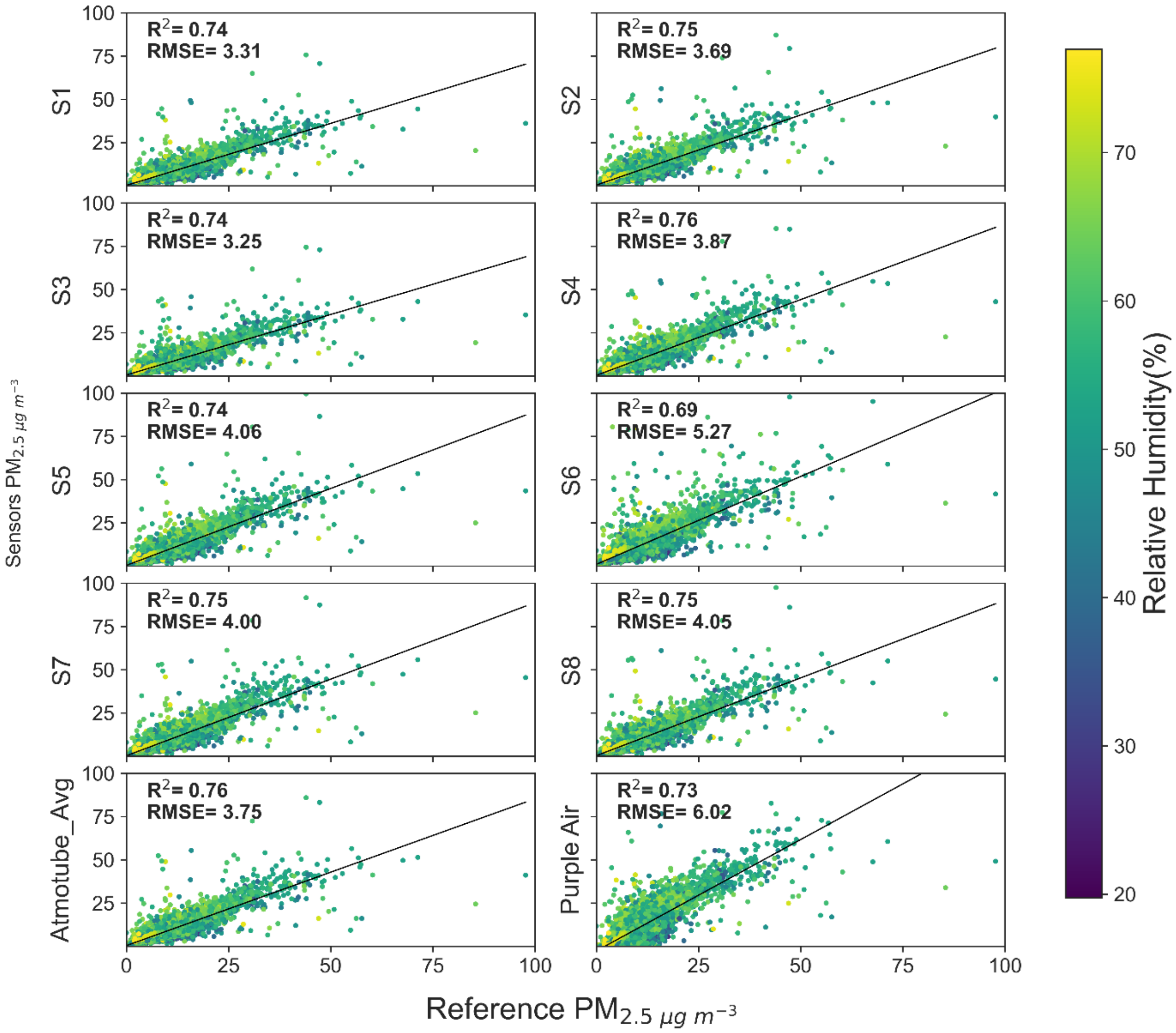



Figure 5: Summary of comparison metrics of each Atmotube Pro sensor, Atmotube Pro Average, Purple Air sensor and reference (15 minutes averaged data) showing $PM_{2.5}$ concentration below 100 $\mu gm^{-3}$.

### ~~3.4~~3.3 Influence of Temperature and Relative Humidity

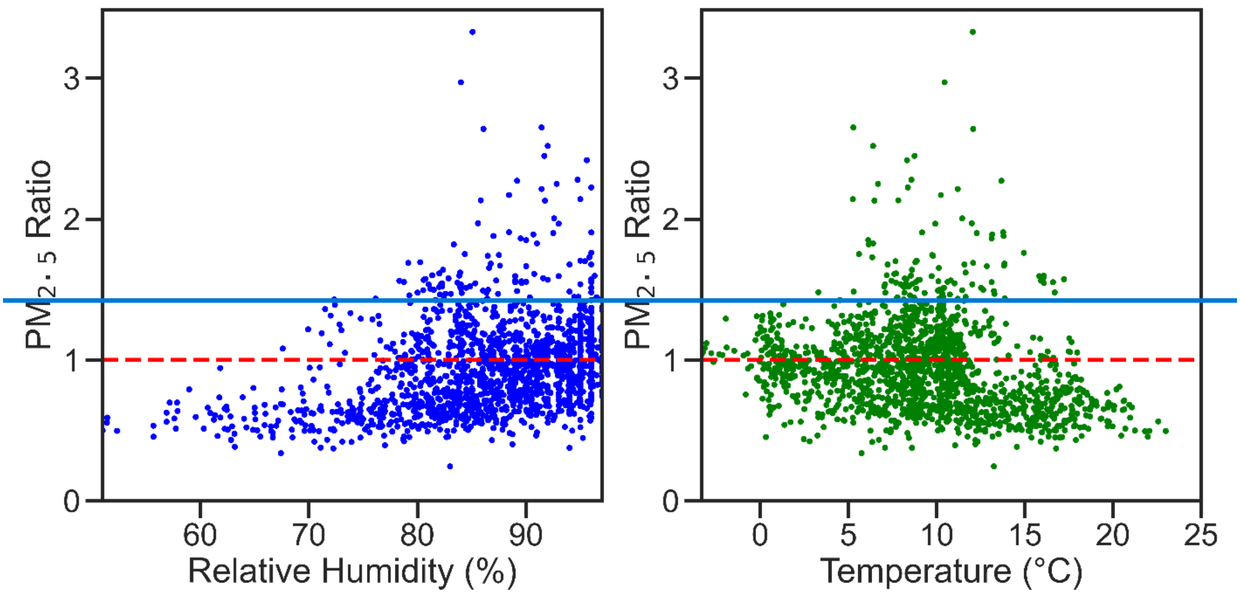
The ratio of the average of all 8 Atmotube Pro sensors and the Reference $PM_{2.5}$ data for hourly averages were calculated. Scatter plots of the $PM_{2.5}$ ratio (defined as Average Atmotube Pro sensor $PM_{2.5}$ / Reference $PM_{2.5}$) as calculated in Eq. (32) were plotted as a function of RH and temperature reported by a nearby weather station as shown in Fig. 6. Data were collected from a local weather station rather than from the Atmotube Pro sensor themselves, since, Zimmerman, (2022) reported that ~~because the~~ RH and temperature sensors in the ~~Atmotube Pro~~ low-cost devices~~sensors~~ can be influenced by sensor heating when connected to power. The nearest meteorological station set up on the rooftop of the School of Earth and Environment building at the University of Leeds (53$^{0}$ 49' 38'' N, 1$^{0}$ 34' 19'' W) about 0.6 miles away from where the colocation experiment of the Atmotube Pro sensors, purple air sensor and reference monitors took place (53$^{0}$ 47' 51'' N, 1$^{0}$ 33' 8'' W).

$$\text{PMRatio} = \frac{\mu_h}{R_h} \qquad (3)$$

Where; $\mu_h$ = Mean PM2.5 concentration hourly average for hour (h) (for all 8 sensors) ($\mu gm^{-3}$)

$R_h$ = Reference $PM_{2.5}$ concentration hourly average for hour (h) ($\mu gm^{-3}$)

For RH, there is a clear relationship with the $PM_{2.5}$ sensor/reference ratio, which increases sharply at RH>80% while at low RH the ratio was below 1.0 indicating the sensors were underestimating $PM_{2.5}$ value relative to the reference monitor. There was no clear influence observed for the $PM_{2.5}$ ratio relative to the temperature, however, there was a general low bias at all temperatures apart from mid-temperature range of 5-15°C. This agrees with results as reported by (Zimmerman, 2022) using Purple Air sensors where a clear influence at 80% RH was also observed and no influence from temperature. Implementing a statistical correction using RH values for these sensors could improve the accuracy of the measured $PM_{2.5}$ values. We recommend further exploration on correction methods using RH in future research investigation.



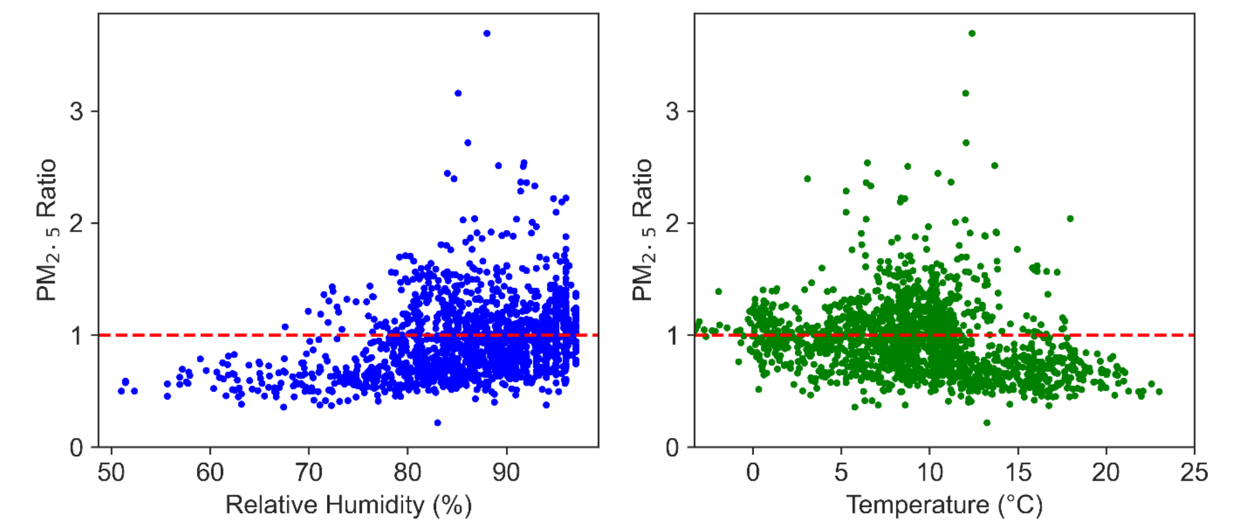




Figure 6: Relationships between (a) relative humidity (RH) and (b) temperature (T) and average Atmotube Pro Sensor/Reference $PM_{2.5}$ ratio.

### 3.4 Correction factor development

Many studies have used multiple linear regression (MLR) calibration models that include temperature, RH and dew point to improve the $PM_{2.5}$ data recorded by low-cost sensors (Badura et al., 2019; Barkjohn et al., 2021; Karaoghlanian et al., 2022; Malings et al., 2019; Raheja et al., 2023). In this section we explored the use of MLR using RH and temperature values to improve Atmotube Pro $PM_{2.5}$ data. We tested using 15 minutes and 1-hour time resolution for the calibration model and assessed the model performance using $R^2$ and RMSE. Given the results of investigating performance of $PM_{2.5}$ concentrations < 100 µgm$^{-3}$ for the 15 minutes average data, we got improved performance as shown in Fig.5.

Table 4: Correction equation forms, the $R^2$ and the RMSE. Best performing calibration equation is indicated as (*)~~,~~. a = slope; i = intercept;

| | Equation | $R^2$ | RMSE |
|---|---|---|---|
| | Hourly averaged (full dataset) | | |
| Linear | $S = a_1 \times PM_{2.5} + i$ | 0.86 | 3.38 |
| +RH | $S = a_1 \times PM_{2.5} + a_2 \times RH + i$ | 0.88 | 3.05 |
| +T | $S = a_1 \times PM_{2.5} + a_2 \times T + i$ | 0.87 | 3.20 |
| | 15 minutes average ($PM_{2.5}$ <100 µgm$^{-3}$) | | |
| Linear | $S = a_1 \times PM_{2.5} + i$ | 0.73 | 3.97 |
| +RH | $S = a_1 \times PM_{2.5} + a_2 \times RH + i$ | 0.79 | 3.48 |
| +T | $S = a_1 \times PM_{2.5} + a_2 \times T + i$ | 0.78 | 3.53 |

The addition RH and temperature values to the model improved the $R^2$ value and decreased the RMSE value. However, the addition of T values only resulted in smaller improvement in the $R^2$ and RMSE relative to using RH values. Similar improvement was also gained in higher resolution data at concentrations < 100 µg/m$^{-3}$. We note that this result cannot be generalised, since the calibration is done at a single location in an urban background during the winter months. It is possible that warmer seasons or different influences on aerosol composition would require different calibration factors.

## 4.0 CONCLUSION

We have conducted comprehensive inter-sensor and reference data comparisons for a set of 8 Atmotube Pro sensors in order to characterise their precision and bias at different levels of $PM_{2.5}$ exposure. The research also explored the potential of identifying underperforming sensors within a network of low-cost sensors, particularly in situations where no reference-grade monitors are available. The study revealed the $PM_{2.5}$ values from the Atmotube Pro sensors had reasonably good precision (CoV of ~~15~~18%) indicating low inter-sensor variability of the sensors. The data cleaning method was successful in improving the inter-sensor variability among the Atmotube Pro sensors. The sensor measurements also replicated measurements from a reference monitor well, with accuracy metrics ranging from; $R^2$ ~~(0.74 to 0.83)~~ (0.77-0.87), Slope ~~(0.9 to 1.2)~~(0.99 to 1.15), Intercept ~~(-2.2 to +0.19)~~ (-1.2 to -0.15) and error biases ranged below the recommended limits for low-cost sensors; RMSE ~~(3.9 to 4.9 µgm$^{-3}$)~~ (2.85to 5.2 µgm$^{-3}$) and NRMSE ~~(2.0 to 3.5%)~~ (0.02 to 0.03%) based on the routinely used US-EPA guidelines. The sensors also showed a strong correlation with purple air sensor where $R^2$ average value was 0.88~~6~~ and an error bias (RMSE) of 2.9 µgm$^{-3}$. Performance of Atmotube Pro sensors was also observed to have deteriorated at higher $PM_{2.5}$ concentrations and improved at a coarser temporal resolution.

Out of the 8 Atmotube Pro sensors used for the assessment, one sensor showed poorer performance with an r value range of 0.57-0.59 while the other sensors reported values above 0.9. ~~The poor sensor had improved the r value range of 0.74-0.77 after applying a data filtering threshold.~~ The overall performance of the 8 Atmotube Pros used for the colocation study is summarized in Table 3. This study observed a precision uncertainty (SD) of ~~9.1~~8.8 µgm$^{-3}$ and an accuracy (RMSE) error of ~~4.4±0.4~~ 3.7 ± 0.8 µgm$^{-3}$ for hourly Atmotube Pro $PM_{2.5}$ data and the chance of having a less reliable sensor in a group of sensors is ~10% (12.5% as the case in this study) and overall gives a useful information for local monitoring or citizen science use. Calibration using multiple linear regression model improved the performance of Atmotube Pro sensors. $R^2$ improved from 0.86 to0.88 and RMSE decreased from 3.38 to 3.05 µg/m$^{-3}$ when accounting for RH values. Future work may look at using multiple models in a longer-term colocation study and in multiple colocation sites to achieve a more robust calibration. It is worthwhile to note that Atmotube Pro sensors (used for both static and non-static $PM_{2.5}$ measurements) are not "plug-and-play" as they require close monitoring and frequent data download to achieve good data recovery.

**Data Availability**

The data downloaded from the Atmotube Pro sensors, data from the FIDAS sensor (reference) and Purple Air sensor data collected from Leeds City Council were used for the analysis and have been uploaded to Zenodo and can be accessed via https://zenodo.org/records/11059054

**Author Contributions**

Aishah Shittu undertook the research study and prepared the manuscript with contributions from all co-authors.

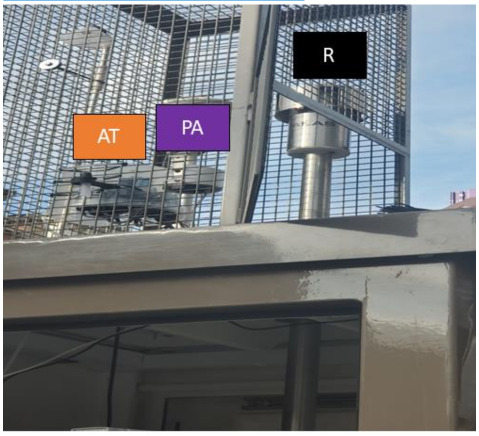
**Competing Interests**

The authors declare that they have no conflict of interest.

**Acknowledgements**

This work is funded by Tertiary Education Trust Fund (TETFund) Nigeria through a postgraduate research scholarship for AS [award reference TETF/ES/POLY/ILARO/TSAS/2020]. This work was supported by the Natural Environment Research Council [grant numbers NE/T010401/1 and NE/R016518/1]. We acknowledge Richard Crowther (Leeds City Council) for providing access to their data used for reference comparison and granting permission for the colocation of sensors at their air quality monitoring site. We express our gratitude to the Born-in-Bradford (BiB) Breathes project [grant number NIHR128833] for providing the Atmotube Pro sensors used for the research. Ben Silver provided valuable advice throughout this project.

**Supplementary Material**




Supp Fig. 1: AT= Atmotube Pro sensors in a makeshift cover to shield rain; PA= Purple Air sensor; R= FIDAS reference monitor inlet

## REFERENCES


Alfano, B., Barretta, L., Del Giudice, A., De Vito, S., Di Francia, G., Esposito, E., Formisano, F., Massera, E., Miglietta, M. L., and Polichetti, T.: A Review of Low-Cost Particulate Matter Sensors from the Developers' Perspectives, Sensors, 20, 6819, https://doi.org/10.3390/s20236819, 2020.

AQMD, 2020: Field Evaluation Atmotube Pro, Air Quality Sensor performance Evaluation Centre, 2020.

Atmotube: https://atmotube.com/atmotube-support/atmotube-technical-specification, last access: 24 July 2023.

Badura, M., Batog, P., Drzeniecka-Osiadacz, A., and Modzel, P.: Regression methods in the calibration of low-cost sensors for ambient particulate matter measurements, SN Appl. Sci., 1, 622, https://doi.org/10.1007/s42452-019-0630-1, 2019.

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

Barkjohn, K. K., Holder, A. L., Frederick, S. G., and Clements, A. L.: Correction and Accuracy of PurpleAir PM2.5 Measurements for Extreme Wildfire Smoke, Sensors, 22, 9669, https://doi.org/10.3390/s22249669, 2022.

Borghi, F., Spinazzè, A., Rovelli, S., Campagnolo, D., Del Buono, L., Cattaneo, A., and Cavallo, D.: Miniaturized Monitors for Assessment of Exposure to Air Pollutants: A Review, IJERPH, 14, 909, https://doi.org/10.3390/ijerph14080909, 2017.

Bousiotis, D., Allison, G., Beddows, D. C. S., Harrison, R. M., and Pope, F. D.: Towards comprehensive air quality management using low-cost sensors for pollution source apportionment, npj Clim Atmos Sci, 6, 122, https://doi.org/10.1038/s41612-023-00424-0, 2023.

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

Carvlin, G. N., Lugo, H., Olmedo, L., Bejarano, E., Wilkie, A., Meltzer, D., Wong, M., King, G., Northcross, A., Jerrett, M., English, P. B., Hammond, D., and Seto, E.: Development and field validation of a community-engaged particulate matter air quality monitoring network in Imperial, California, USA, Journal of the Air & Waste Management Association, 67, 1342–1352, https://doi.org/10.1080/10962247.2017.1369471, 2017.

Chatzidiakou, L., Krause, A., Popoola, O. A. M., Di Antonio, A., Kellaway, M., Han, Y., Squires, F. A., Wang, T., Zhang, H., Wang, Q., Fan, Y., Chen, S., Hu, M., Quint, J. K., Barratt, B., Kelly, F. J., Zhu, T., and Jones, R. L.: Characterising low-cost sensors in highly portable platforms to quantify personal exposure in diverse environments, Atmos. Meas. Tech., 12, 4643–4657, https://doi.org/10.5194/amt-12-4643-2019, 2019.

Diez, S., Lacy, S., Coe, H., Urquiza, J., Priestman, M., Flynn, M., Marsden, N., Martin, N. A., Gillott, S., Bannan, T., and Edwards, P. M.: Long-term evaluation of commercial air quality sensors: an overview from the QUANT (Quantification of Utility of Atmospheric Network Technologies) study, Atmos. Meas. Tech., 17, 3809–3827, https://doi.org/10.5194/amt-17-3809-2024, 2024.

Duvall, Clements, A., Hagler, G., Kamal, A., Kilaru, V., Goodman, L., and Frederick, S.: Performance testing Protocols, Metrics, and Target Values for Fine Particulate Matter Air Sensors, US Environmental Protection Agency, Office of Research and Development, 79, 2021.

Feenstra, B., Papapostolou, V., Hasheminassab, S., Zhang, H., Boghossian, B. D., Cocker, D., and

Polidori, A.: Performance evaluation of twelve low-cost PM2.5 sensors at an ambient air monitoring site, Atmospheric Environment, 216, 116946, https://doi.org/10.1016/j.atmosenv.2019.116946, 2019.

FIDAS: https://www.acoem.co.uk/product/palas/fidas-200/, last access: 6 February 2024.

Giordano, M. R., Malings, C., Pandis, S. N., Presto, A. A., McNeill, V. F., Westervelt, D. M., Beekmann, M., and Subramanian, R.: From low-cost sensors to high-quality data: A summary of challenges and best practices for effectively calibrating low-cost particulate matter mass sensors, Journal of Aerosol Science, 158, 105833, https://doi.org/10.1016/j.jaerosci.2021.105833, 2021.

Hagan, D. H. and Kroll, J. H.: Assessing the accuracy of low-cost optical particle sensors using a physics-based approach, Atmos. Meas. Tech., 13, 6343–6355, https://doi.org/10.5194/amt-13-6343-2020, 2020.

Hagan, D. H., Gani, S., Bhandari, S., Patel, K., Habib, G., Apte, J. S., Hildebrandt Ruiz, L., and Kroll, J. H.: Inferring Aerosol Sources from Low-Cost Air Quality Sensor Measurements: A Case Study in Delhi, India, Environ. Sci. Technol. Lett., 6, 467–472, https://doi.org/10.1021/acs.estlett.9b00393, 2019.

Hong, G.-H., Le, T.-C., Tu, J.-W., Wang, C., Chang, S.-C., Yu, J.-Y., Lin, G.-Y., Aggarwal, S. G., and Tsai, C.-J.: Long-term evaluation and calibration of three types of low-cost PM2.5 sensors at different air quality monitoring stations, Journal of Aerosol Science, 157, 105829, https://doi.org/10.1016/j.jaerosci.2021.105829, 2021.

Jovašević-Stojanović, M., Bartonova, A., Topalović, D., Lazović, I., Pokrić, B., and Ristovski, Z.: On the use of small and cheaper sensors and devices for indicative citizen-based monitoring of respirable particulate matter, Environmental Pollution, 206, 696–704, https://doi.org/10.1016/j.envpol.2015.08.035, 2015.

Kang, Y., Aye, L., Ngo, T. D., and Zhou, J.: Performance evaluation of low-cost air quality sensors: A review, Science of The Total Environment, 818, 151769, https://doi.org/10.1016/j.scitotenv.2021.151769, 2022.

Karaoghlanian, N., Noureddine, B., Saliba, N., Shihadeh, A., and Lakkis, I.: Low cost air quality sensors “PurpleAir” calibration and inter-calibration dataset in the context of Beirut, Lebanon, Data in Brief, 41, 108008, https://doi.org/10.1016/j.dib.2022.108008, 2022.

Keswani, A., Akselrod, H., and Anenberg, S. C.: Health and Clinical Impacts of Air Pollution and Linkages with Climate Change, NEJM Evidence, 1, https://doi.org/10.1056/EVIDra2200068, 2022.

Khreis, H., Cirach, M., Mueller, N., De Hoogh, K., Hoek, G., Nieuwenhuijsen, M. J., and Rojas-Rueda, D.: Outdoor air pollution and the burden of childhood asthma across Europe, Eur Respir J, 54, 1802194, https://doi.org/10.1183/13993003.02194-2018, 2019.

Liu, Y., Paciorek, C. J., and Koutrakis, P.: Estimating Regional Spatial and Temporal Variability of $PM_{2.5}$ Concentrations Using Satellite Data, Meteorology, and Land Use Information, Environ Health Perspect, 117, 886–892, https://doi.org/10.1289/ehp.0800123, 2009.

Malings, C., Tanzer, R., Hauryliuk, A., Kumar, S. P. N., Zimmerman, N., Kara, L. B., Presto, A. A., and R. Subramanian: Development of a general calibration model and long-term performance evaluation of low-cost sensors for air pollutant gas monitoring, Atmos. Meas. Tech., 12, 903–920, https://doi.org/10.5194/amt-12-903-2019, 2019.

Malings, C., Tanzer, R., Hauryliuk, A., Saha, P. K., Robinson, A. L., Presto, A. A., and Subramanian, R.: Fine particle mass monitoring with low-cost sensors: Corrections and long-term performance evaluation, Aerosol Science and Technology, 54, 160–174, https://doi.org/10.1080/02786826.2019.1623863, 2020.

Mansourian, M., Javanmard, S., Poursafa, P., and Kelishadi, R.: Air pollution and hospitalization for respiratory diseases among children in Isfahan, Iran, Ghana Medical Journal, 44, https://doi.org/10.4314/gmj.v44i4.68906, 2011.

Masri, S., Rea, J., and Wu, J.: Use of Low-Cost Sensors to Characterize Occupational Exposure to

PM2.5 Concentrations Inside an Industrial Facility in Santa Ana, CA: Results from a Worker- and Community-Led Pilot Study, Atmosphere, 13, 722, https://doi.org/10.3390/atmos13050722, 2022.

Masri, S., Flores, L., Rea, J., and Wu, J.: Race and Street-Level Firework Legalization as Primary Determinants of July 4th Air Pollution across Southern California, Atmosphere, 14, 401, https://doi.org/10.3390/atmos14020401, 2023.

Maynard, R., Carare, R., Grigg, J., Fox, N., Love, S., Mudway, I., Shaddick, G., Delgado-Saborit, J. M., Guercio, V., Earl, N., Mitsakou, C., Doutsi, A., Exley, K., Robertson, S., Douglas, P., and Gowers, A.: 53 Cognitive Decline, Dementia and Air Pollution: A Report by the Committee on the Medical Effects of Air Pollutants, Annals of Work Exposures and Health, 67, i80–i81, https://doi.org/10.1093/annweh/wxac087.196, 2023.

Molina Rueda, E., Carter, E., L'Orange, C., Quinn, C., and Volckens, J.: Size-Resolved Field Performance of Low-Cost Sensors for Particulate Matter Air Pollution, Environ. Sci. Technol. Lett., 10, 247–253, https://doi.org/10.1021/acs.estlett.3c00030, 2023.

Morawska, L., Thai, P. K., Liu, X., Asumadu-Sakyi, A., Ayoko, G., Bartonova, A., Bedini, A., Chai, F., Christensen, B., Dunbabin, M., Gao, J., Hagler, G. S. W., Jayaratne, R., Kumar, P., Lau, A. K. H., Louie, P. K. K., Mazaheri, M., Ning, Z., Motta, N., Mullins, B., Rahman, M. M., Ristovski, Z., Shafiei, M., Tjondronegoro, D., Westerdahl, D., and Williams, R.: Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone?, Environment International, 116, 286–299, https://doi.org/10.1016/j.envint.2018.04.018, 2018.

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

Polidori, A., Papapostolou, V., Feenstra, B., and Zhang, H.: Field Evaluation of Low-Cost Air Quality Sensors, 2017.

Pope, F., Bousiotis, D., Beddows, D., and Allison, G.: Low cost source apportionment of urban air pollution, , https://doi.org/10.5194/egusphere-egu22-4361, 2022.

Raheja, G., Nimo, J., Appoh, E. K.-E., Essien, B., Sunu, M., Nyante, J., Amegah, M., Quansah, R., Arku, R. E., Penn, S. L., Giordano, M. R., Zheng, Z., Jack, D., Chillrud, S., Amegah, K., Subramanian, R., Pinder, R., Appah-Sampong, E., Tetteh, E. N., Borketey, M. A., Hughes, A. F., and Westervelt, D. M.: Low-Cost Sensor Performance Intercomparison, Correction Factor Development, and 2+ Years of Ambient PM $_{2.5}$ Monitoring in Accra, Ghana, Environ. Sci. Technol., 57, 10708–10720, https://doi.org/10.1021/acs.est.2c09264, 2023.

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, Science of The Total Environment, 607–608, 691–705, https://doi.org/10.1016/j.scitotenv.2017.06.266, 2017.

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

Voultsidis, D., Gialelis, J., Protopsaltis, G., Bali, N., and Mountzouris, C.: Utilizing Unobtrusive Portable Electronic Devices for Real-Time Assessment of Indoor PM2.5 and tVOC Exposure and Its Correlation with Heart Rate Variability, Procedia Computer Science, 224, 550–557, https://doi.org/10.1016/j.procs.2023.09.080, 2023.

Wang, X., Zhou, H., Arnott, W. P., Meyer, M. E., Taylor, S., Firouzkouhi, H., Moosmüller, H., Chow, J. C., and Watson, J. G.: Evaluation of gas and particle sensors for detecting spacecraft-relevant fire emissions, Fire Safety Journal, 113, 102977, https://doi.org/10.1016/j.firesaf.2020.102977, 2020.

Westervelt, D. M., Isevulambire, P. K., Yombo Phaka, R., Yang, L. H., Raheja, G., Milly, G., Selenge, J.-L. B., Mulumba, J. P. M., Bousiotis, D., Djibi, B. L., McNeill, V. F., Ng, N. L., Pope, F., Mbela, G. K., and

Konde, J. N.: Low-Cost Investigation into Sources of PM $_{2.5}$ in Kinshasa, Democratic Republic of the Congo, ACS EST Air, 1, 43–51, https://doi.org/10.1021/acsestair.3c00024, 2024.

WHO, 2021: WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide, WHO European Centre for Environment and Health, Bonn, Germany, 2021.

Williams, R., Kilaru, V., Snyder, E., Kaufman, A., Dye, T., Rutter, A., Russell, A., and Hafner, H.: Air Sensor Guidebook, Office of Research and Develoment, 2014.

Yang, L. H., Hagan, D. H., Rivera-Rios, J. C., Kelp, M. M., Cross, E. S., Peng, Y., Kaiser, J., Williams, L. R., Croteau, P. L., Jayne, J. T., and Ng, N. L.: Investigating the Sources of Urban Air Pollution Using Low-Cost Air Quality Sensors at an Urban Atlanta Site, Environ. Sci. Technol., 56, 7063–7073, https://doi.org/10.1021/acs.est.1c07005, 2022.

Zamora, M. L., Rice, J., and Koehler, K.: One year evaluation of three low-cost PM2.5 monitors, Atmospheric Environment, 235, 117615, https://doi.org/10.1016/j.atmosenv.2020.117615, 2020.

Zimmerman, N.: Tutorial: Guidelines for implementing low-cost sensor networks for aerosol monitoring, Journal of Aerosol Science, 159, 105872, https://doi.org/10.1016/j.jaerosci.2021.105872, 2022.