https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

Performance evaluation of MOMA - a remote network calibration technique for $PM_{2.5}$ and PM_{10} sensors

Lena Francesca Weissert[1)], Geoff Steven Henshaw[1)], David Edward Williams[2)], Brandon Feenstra[3)], Randy Lam[3)], Ashley Collier-Oxandale[3)], Vasileios Papapostolou[3)] and Andrea Polidori[3)]

[1]Aeroqual Ltd, 460 Rosebank Road, Avondale, Auckland, 1026, New Zealand
[2]School of Chemical Sciences and MacDiramid Institute for Advanced Materials and Nanotechnology, University of Auckland, Private Bag 92019, Auckland, 1142, New Zealand
[3]South Coast Air Quality Management District, 21865 Copley Drive, Diamond Bar, CA, 91765, USA

Correspondence to: Lena Weissert (lena.weissert@aeroqual.com)

Abstract. We evaluate the potential of using a previously developed remote calibration framework we name MOMA to improve the data quality in PM sensors deployed in hierarchical networks. MOMA assumes that a network of reference instruments can be used as 'proxies' to calibrate the sensors given that the probability distribution of the data at the proxy site is similar to that at a sensor site. We use the reference network to test the suitability of proxies selected based on distance versus proxies selected based on land use similarity. The performance of MOMA for PM sensors is tested with sensors collocated with reference instruments across three Southern California regions, representing a range of land uses, topography, and meteorology, and calibrated against a distant proxy reference. We compare two calibration approaches, one where calibration parameters get calculated and applied at monthly intervals and one which uses a drift detection framework for calibration. We demonstrate that MOMA improves the accuracy of the data when compared against the collocated reference data. The improvement was more visible for PM_{10} and when using the drift detection approach. We also highlight that sensor drift was associated with variations in particle composition rather than instrumental factors explaining the better performance of the drift detection approach if wind conditions and associated PM sources varied within a month.

1 Introduction

Particulate matter (PM) is a major air pollutant with adverse cardiovascular and respiratory health effects. Elevated PM concentrations are linked to natural (e.g., volcanoes, wild fires, dust storms, sea salt) and anthropogenic emissions (e.g., transport, industrial, agricultural and household fuel combustion) (Anderson et al., 2012). $PM_{2.5}$ (particles <2.5 µm in aerodynamic diameter) and PM_{10} (particles <10 µm in aerodynamic diameter) are routinely measured by government and research organisations using reference-grade equipment that is either filter-based Federal Reference Method (FRM) or continuous Federal Equivalence Method (FEM). Reference monitoring networks are designed to measure regional air pollution to determine attainment of national ambient air quality standards but are often sparsely sited across a region due to high instrument and operational costs. The last decade has seen a rapid increase in the availability of PM sensors offering

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

opportunities to measure PM with much denser networks and making them popular choices for citizen projects and community monitoring (Giordano et al., 2021; Liang, 2021; Snyder et al., 2013).

Most PM sensors are optical sensors that utilize the light scattered by particles to derive the particle size and number based on the Mie theory (Alfano et al., 2020; Liang, 2021). The relationship between scattered light, particle size and number, and the PM mass is dependent on the properties of the particles, which include size, shape, refractive index, and composition (Chen et al., 2019; Johnson et al., 2018). This poses a major challenge for calibrating PM sensors as calibration factors may change with particle type or properties changes over time. More frequent field calibrations may be required if aerosol properties vary significantly over time (Liang, 2021; Johnson et al., 2018; Badura et al., 2018). Regular calibration and maintenance are therefore critical to ensure reliable data from PM sensor networks (Giordano et al., 2021; Hofman et al., 2022; Williams, 2019). Given the costs and feasibility related to individual site visits and calibrations by collocation, new approaches are required for large scale sensor networks to be viable. Recent studies (Liang, 2021; De Vito et al., 2020; Loh and Choi, 2019) have used Machine learning (ML) approaches to train calibration models with enough collocation data to cover various meteorological and environmental conditions and make them more robust for long-term sensor deployments. However, if conditions (e.g., different traffic conditions, different PM sources) at the calibration site are different from the conditions at the site of interest the model may no longer be suitable (De Vito et al., 2020; Liang, 2021). In addition, while being more robust and effective, ML may still suffer from challenges related to sensor degradation when sensors are deployed in a long-term fashion (Liang, 2021).

In previous publications, we demonstrated that a hierarchical network, consisting of well-maintained reference-grade instruments (referred to as 'proxies') and gas-phase (O_3, NO_2) sensors can be used to correct sensors remotely (Miskell et al., 2018, 2019; Weissert et al., 2020). The correction framework is based on the assumption that the probability distribution at a proxy site, which can be selected based on proximity for O_3 measurements or similar land use for NO_2 measurements is similar to that of the sensor site (Miskell et al., 2018, 2019; Weissert et al., 2020). We have demonstrated that this approach is able to successfully correct for sensor drift without the need of collocation.

In this paper, we examine how this remote calibration methodology named MOMA (from moment matching) performs for PM sensors deployed in Southern California, including the City of LA, the Inland Empire (IE), and a desert region of Riverside County (RC Desert). These three regions differ in terms of land use, terrain and meteorology offering an opportunity to test MOMA under different seasonal conditions and PM sources. We also test proxy selection strategies by examining the comparability of reference sites as a function of distance and land use. Two approaches were investigated, the first one calculated and applied the calibration at monthly intervals while the second approach used a drift detection approach (Miskell et al., 2018, 2019; Weissert et al., 2020) to apply the calibration when drift between a sensor and the proxy site was detected.

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

2 Materials and Methods

2.1 Study area

This study was performed in Southern California in a region that is under the jurisdiction of the South Coast Air Quality Management District (South Coast AQMD). South Coast AQMD manages a network of regulatory-grade $PM_{2.5}$ and PM_{10} monitors. Non-regulatory air quality sensors measuring PM were collocated at two air monitoring sites (AMS) in the City of LA, two AMS in the IE, and two AMS in the RC Desert. Each reference site is equipped with a reference-grade instrument and data was obtained either from AirNow (https://www.airnow.gov/) or directly from South Coast AQMD. Refer to Table S1 for instrumentation at each site.

Elevated PM levels, mostly driven by vehicular emissions and freight activities, pose a serious health risk in Los Angeles (Ault et al., 2009; Habre et al., 2021; Kim and Kwan, 2021). While westerly winds dominate in Southern California meteorology for most of the year, north-easterly Santa Ana Winds (SAW) become more frequent during the fall and winter months. SAW are associated with very dry air and good visibility in the absence of wildfires as urban pollutants are blown offshore (Aguilera et al., 2020). However, they are also key drivers of large wildfires enabling them to spread faster and transporting smoke PM from inland areas to the more populated regions.

2.2 Air Quality Sensors

This study uses a network of AQY v1.0 (AQY) sensor systems from Aeroqual Ltd, Auckland, New Zealand. The AQY measures O_3, NO_2, $PM_{2.5}$, PM_{10}, Temperature, and Relative Humidity. Detailed description about the AQY sensor system is available in Weissert et al. (2020) and Miskell et al. (2019). The focus of this paper is the PM sensor (model SDS011, Nova Fitness Co., Ltd, Jinan City/China) inside the AQY sensor system. The SDS011 is an optical light scattering device which outputs $PM_{2.5}$ and PM_{10} mass concentration (μg m^{-3}) measurements. Previous studies of this sensor have shown high $PM_{2.5}$ correlation with reference instruments (Badura et al., 2018; Liu et al., 2019) but PM_{10} values may be underestimated (Budde et al., 2018; Kuula et al., 2020). Nevertheless, we use both $PM_{2.5}$ and PM_{10} measurements to evaluate the performance of our network calibration technique applied to PM data. The SDS011 sensor was factory calibrated against a Met One 9722 8 channel optical particle counter (Met One Instruments, Inc., Grants Pass, Oregon, US) using 1 μm latex microspheres. The AQY performs a humidity correction on the $PM_{2.5}$ and PM_{10} measurements from the SDS011 using an empirical humidity algorithm developed by Aeroqual Ltd.

The AQY PM measurements were evaluated by South Coast AQMD's Air Quality Sensor Performance Evaluation Centre (AQ-SPEC) (http://www.aqmd.gov/aq-spec/sensordetail/aeroqual-aqy-v1.0).

2.3 Remote Network Calibration

The remote network calibration technique, called MOMA, was developed for hierarchical networks that consist of a network of well-calibrated reference grade instruments acting as "proxies" which are used to calibrate the sensors deployed in the field. The technique is described in detail in Miskell et al. (2016, 2018). We tested two approaches to calibrate the $PM_{2.5}$ and PM_{10} sensors in this study over the period August 2020 to February 2022.

The first approach was a monthly MOMA calibration using the last two weeks of each month to select a suitable seven-day calibration window to calculate the calibration parameters which were then applied to the following month. The last two weeks of the month were selected to ensure most recent data were used to determine calibration gains and offsets. A calibration window was considered suitable if the data completeness for both proxy and sensor was greater than 85%. In addition, we excluded periods with fog from the calibration (Budde et al., 2018). Fog can frequently be present between October and February in the regions, drive by lower inversion levels (Qin et al., 2012; Witiw and LaDochy, 2008) (Fig. S1). Periods when fog was detected were also removed from the analysis in this paper. The MOMA gain and offset were calculated as described in Miskell et al. (2016, 2018) and the new gains and offsets were uploaded to each AQY instrument.

The second approach used a previously described drift detection framework (Miskell et al., 2016) to trigger a MOMA calibration. The drift detection framework uses three statistical tests to detect sensor drift, a Kolmogorov-Smirnov (K-S) test, the MOMA slope, $\boldsymbol{\hat{a}_1}$ and the MOMA offset, $\boldsymbol{\hat{a}_0}$. The statistical tests are run over a three-day period and an alarm is triggered when any of the tests exceeds the predetermined threshold for a period of 5 days. The following thresholds were used to determine if a sensor drifted: $\boldsymbol{p_{KS}}$ = 0.05, $\boldsymbol{\hat{a}_1}$ = 1 $\pm$ 0.25, $\boldsymbol{\hat{a}_0}$ = 0 $\pm$ 5 µg m^{-3} These thresholds can be adjusted to explore test sensitivity to drift detection. This framework was run from August 2020 – January 2021 to compare with the output from the monthly calibrations.

The statistical analysis was performed in R (v.4.1.3) using tidyverse (Wickham and RStudio, 2022), lubridate (Spinu et al., 2022), zoo (Zeileis et al., 2022), ggrepel (Slowikowski et al., 2022), openair (Carslaw and Ropkins, 2022), RAQSAPI (Mccrowey et al., 2022), ggplot2 (Wickham et al., 2022b), dplyr (Wickham et al., 2022a), ggmap (Kahle and Wickham, 2013) and ggpmisc (Aphalo et al., 2022).

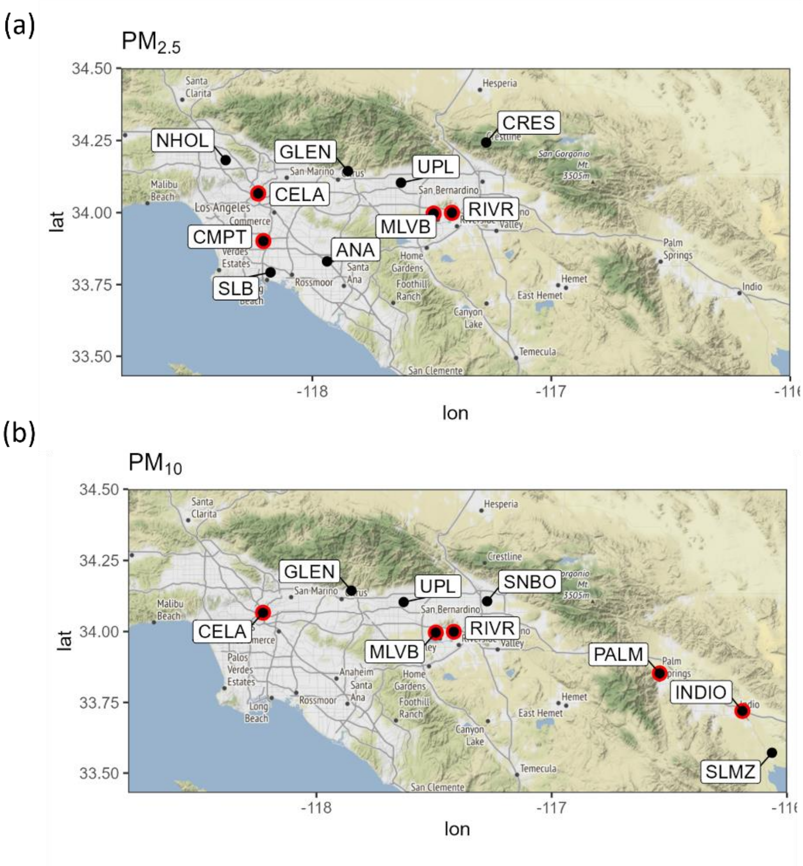
2.4 Proxy selection

A key part of MOMA is the identification of a suitable proxy site for each sensor in the sensor network. In previous publications, we showed that the distance between the sensor and the reference proxy was a suitable selection criterion for O_3 sensors (Miskell et al., 2019). However, land use similarity was a more suitable selection criterion for NO_2 sensors due to the pollutant being more variable spatially and temporally depending on the dominant land use surrounding the site (Weissert et al., 2020). In the case of PM, its spatial and temporal variability is driven by multiple factors including local emissions of primary PM such as traffic, construction, and residential heating as well as regional transport and the formation of secondary PM. There is evidence that while $PM_{2.5}$ levels tend to be relatively homogeneous across an urban region, PM_{10} can be spatially more variable due to the shorter lifetime and more variable sources (Pinto et al., 2004; Sardar, 2005). In this paper, we explore the suitability of a proxy site based on distance versus land use similarity. Land use variables used for the

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

analysis were a) road length within a 1 km buffer around the site, b) distance of the site from a motorway and c) elevation. These are simple and widely available variables and have also been identified as good predictors for PM in land use regression studies in the US (Kloog et al., 2012; Lee et al., 2016) and Europe (Eeftens et al., 2012).

For the proxy selection test, we use $PM_{2.5}$ and PM_{10} data from the South Coast AQMD regulatory monitoring network. Los Angeles, N. Main Street (CELA), Compton (CMPT), Mira Loma – Van Buren (MLVB) and Rubidoux (RIVR) were used as test locations for which a suitable $PM_{2.5}$ proxy is found. CELA, MLVB, RIVR, Palm Springs (PALM) and Indio-29 Palms (INDIO) are used as test locations to identify suitable for PM_{10} proxies (Fig. 1). The Mean Absolute Error (MAE) and coefficient of determination (R^2) were calculated from daily averaged reference data for the year 2020 (LA, IE) and 2021 (RC Desert) for different proxies against the distance and land use similarity for each test location.



(c)

ID	Name
MLVB	Mira Loma - Van Buren
RIVR	Riverside - Rubidoux
SNBO	San Bernadino
CELA	Los Angeles - N.Main Street
CMPT	Compton
NHOL	North Hollywood
ANA	Anaheim
SLB	South Long Beach
GLEN	Glendora - Laurel
CRES	Crestline - Lake Gregory
UPL	Upland
PALM	Palm Springs
INDIO	Indio-29 Palms
SLMZ	Saul Martinez

Figure 1: a) $PM_{2.5}$ and b) PM_{10} South Coast AQMD reference Air Monitoring Network (Los Angeles Region). Red circles highlight test locations for which a suitable proxy is found from network options (black circles,) c) Table of the site Names for the IDs shown in a) and b). The map was created using ggmap (Kahle and Wickham, 2013).

Given that MOMA is based on matching probability distributions rather than regression models we also compare probability distributions of hourly $PM_{2.5}$ and PM_{10} at the test locations (CELA, CMPT, MLVB, RIVR, INDIO, PALM) to those of different proxies using the K-S (Kolmogorov-Smirnov) test statistic as a measure of similarity across probability distributions.

2.5 Evaluating the performance of MOMA

Six AQYs collocated at South Coast AQMD AMS sites were used to test the calibration framework (Table 1, Fig. 1). For each AQY, a proxy reference (other than the collocated reference) was selected for the monthly and the drift detection triggered MOMA calibrations. The sensor measurements were then compared to the collocated reference data. The six sensors were deployed between April 2020 and January 2021 (SI Table 2). They were used to evaluate the performance of the proxy selection and the two different network calibration approaches. $PM_{2.5}$ data from the collocated AQYs were available from August 2020 – December 2021 (and ongoing) while the majority of the PM_{10} data were added at the start of 2021.

2.6 Speciation data

Speciation data were obtained using the RAQSAPI package (Mccrowey et al., 2022), which enables downloading monitoring data from the US Environmental Protection Agency's Air Quality System service. We focused on parameters representing crustal material, trace ions, secondary ions, elemental carbon (EC) and organic carbon (OC) and followed the classification described in Daher et al. (2013) (Table S3). Samples were taken every third day over a 24 -hour period. Organic carbon (OC) and elemental carbon (EC) were collected via an URG 3000N with a Pall Quartz filter and Cyclone Inlet and the total amount was used in this analysis. The remaining parameters were collected using a Met One Speciation Air Sampling System (SASS) (Met One Instruments, Inc., Grants Pass, Oregon, US).

3 Results and Discussion

3.1 General characteristics of the data

Figure 2 shows the seasonal $PM_{2.5}$, PM_{10} variations observed at the reference sites as well as the relative humidity and temperature recorded at these sites from January – December 2021. While $PM_{2.5}$ levels seem to be comparable across the sites and networks in LA and the IE, lower levels were observed in the RC Desert. There are also distinct differences in the PM_{10} concentrations with higher levels observed in the IE (RIVR, MLVB). $PM_{2.5}$ concentrations are highest in autumn and generally more variable over the autumn/winter period. The RC Desert region (PALM, INDIO, SLMZ) is drier and hotter compared to LA and the IE. The timeseries shown in Fig. 3 show that while short-term local effects are visible (particularly for PM_{10} in the IE and RC Desert), overall diurnal $PM_{2.5}$ and PM_{10} variations across sites within the same region are similar. This suggests that MOMA could be an effective calibration framework for PM since the underlying requirement, that the diurnal patterns of pollutants at the proxy site and at the site to be calibrated are similar, seems to be met, particularly for

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

$PM_{2.5}$. For PM_{10}, a more careful selection of a suitable calibration window may be required, given the short-term local differences.

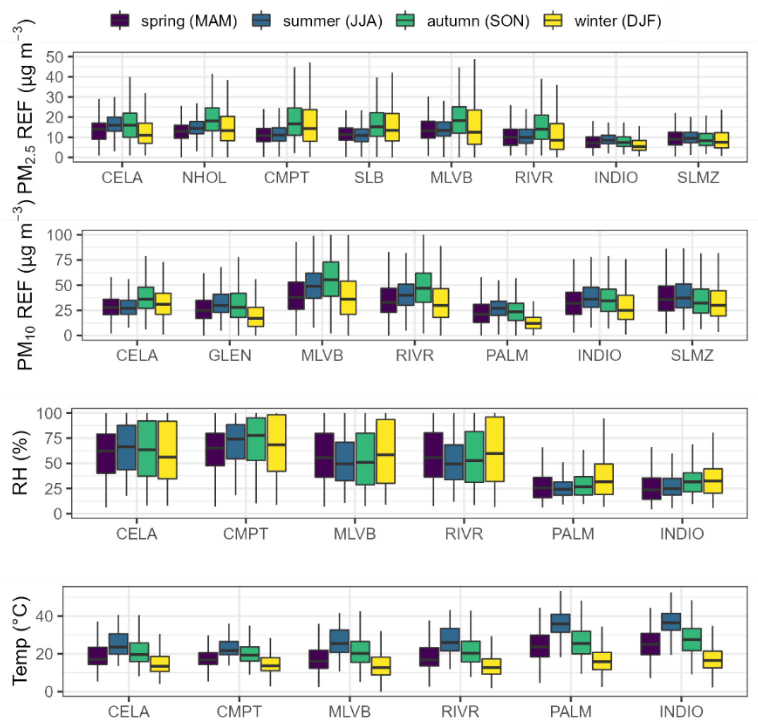


Figure 2: Boxplots showing the seasonal $PM_{2.5}$ and PM_{10} variations in 2021 at collocated sites and the RH and Temperature observed at the South Coast AQMD AMS sites with collocated AQY. The boxplots are coloured by season (spring (MAM): March, April, May, summer (JJA): June, July, August, autumn (SON): September, October, November, winter (DJF): December, January, February).

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

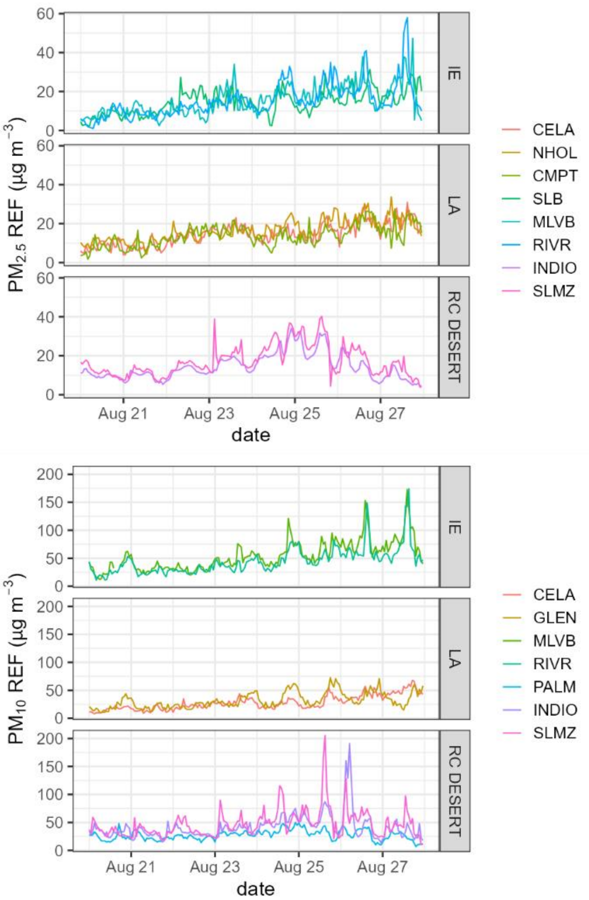


Figure 3: $PM_{2.5}$ and PM_{10} reference timeseries for a 7-day period grouped by Networks (i.e., IE, LA, RC Desert).

3.2 Proxy selection criteria

To assess the influence of distance versus land use on proxy selection we used three metrics, MAE, R^2 and the K-S test statistic, to evaluate each proxy option at four South Coast AQMD AMS across the regions. By using data from the reference network any uncertainties related to sensor performance are eliminated. Figure 4 illustrates that in most cases the nearest proxy site rather than the site with the most similar land use is the most suitable proxy resulting in the lowest and highest R^2 across the whole year. Using the K-S test statistic as a measure of similarity across probability distributions reveals a slightly different pattern suggesting that $PM_{2.5}$ CMPT or SLB may be more suitable proxies for CELA and that $PM_{2.5}$ CELA could be a suitable proxy for MLVB or RIVR when upwind from MLVB or RIVR.

However, there are exceptions to this observation suggesting that other factors, such as particle sources associated with the surrounding land use, terrain, or prevailing wind direction, likely also contribute to the suitability of a proxy. For example, a proxy further away (CELA) seems to perform similarly to a nearby proxy (UPL) for $PM_{2.5}$ at Mira Loma (MLVB). Mira

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

Loma is downwind from CELA for most of the year, possibly explaining the low MAE against MLVB. The CRES site also seems to be a poorer $PM_{2.5}$ proxy for MVLB and RIVR, which may be due to its location at higher altitudes as well as being separated from MVLB and RIVR by the San Bernardino mountains (1200+ meters high).

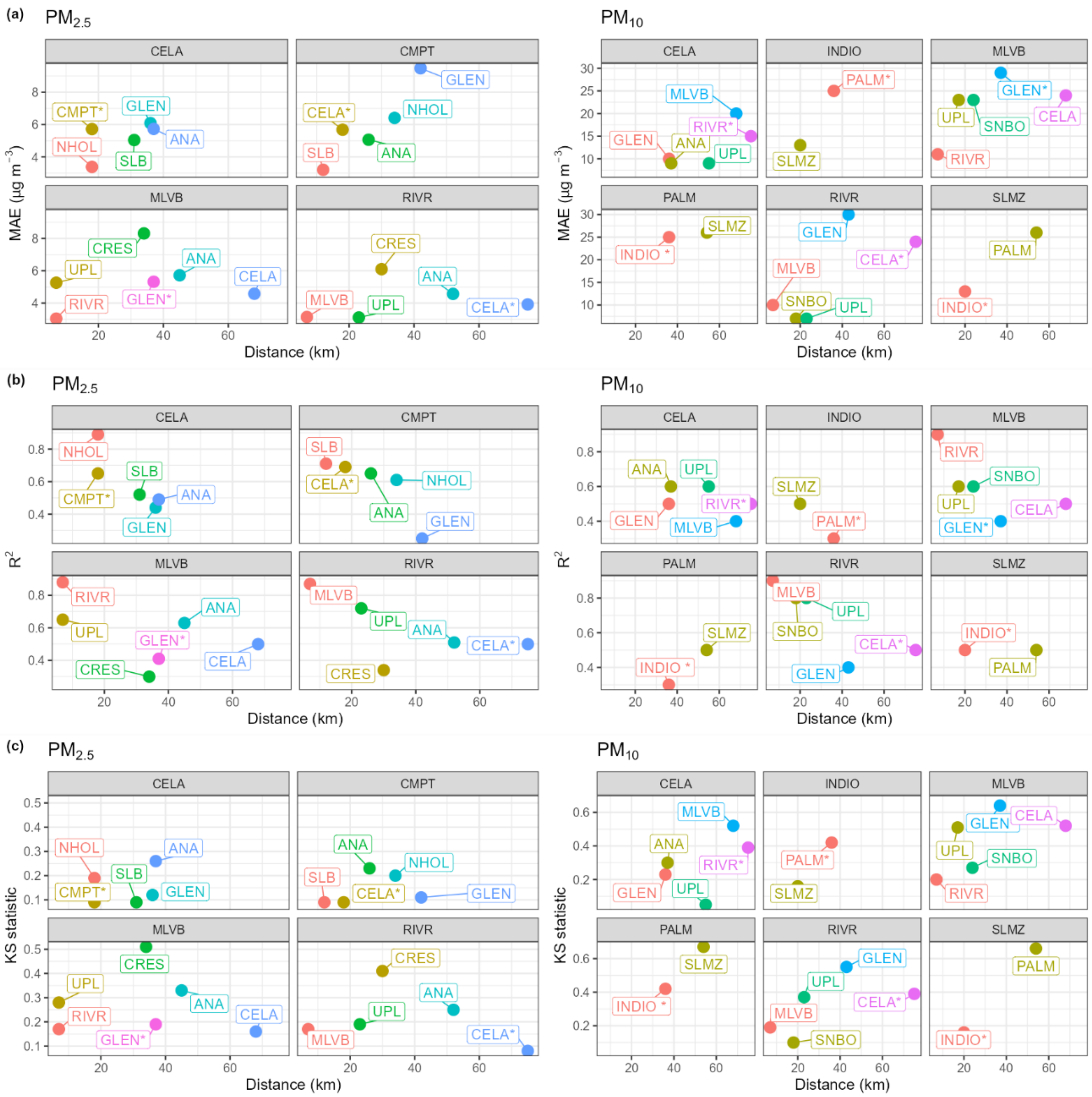


Figure 4: a) MAE, b) R^2 and c) K-S statistic calculated from daily averaged reference data for the year 2020 for different proxies against distance to site of interest for $PM_{2.5}$: CELA, CMPT, MLVB, RIVR, and PM_{10}: CELA, MLVB, RIVR. The site with the most similar land use is labelled with a '*'. The proxy site is labelled in each facet. The full site names are shown in Fig. 1 c).

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

Overall, the nearest proxy generally resulted in the most similar distribution with the smallest MAE and largest R^2, and therefore the nearest proxy was selected to calibrate the sensors in the following sections.

3.3 MOMA Calibration performance

The performance of MOMA was evaluated using sensors that were collocated at a regulatory site. Each sensor was mapped to its nearest proxy (Table 1), calibrated using the MOMA technique and compared to its collocated South Coast AQMD AMS using the metrics MAE and R^2.

Table 1. List of AQYs collocated at South Coast AQMD AMS sites with their proxy reference sites.

Site	Network	PM2.5 Proxy	PM10 Proxy	Distance to $PM_{2.5}$ Proxy	Distance to PM_{10} Proxy
RIVR	IE	MLVB	MLVB	7	7
MLVB	IE	RIVR	RIVR	7	7
CELA	LA	SLB	GLEN	12	36
CMPT	LA	NHOL	*	18	
PALM	RC Desert	*	INDIO		36
INDIO	RC Desert	SLMZ	PALM	21	36

* There is no PM_{10} data available from CMPT and no $PM_{2.5}$ measurement available from PALM

3.3.1 $PM_{2.5}$

Figure 5 shows hourly uncalibrated (gain = 1, offset = 0) + RH correction, monthly calibrated and drift calibrated sensor data and proxy data against the collocated reference data over the 12-month period Jan 2021 to Dec 2021. The sensors at CELA and CMPT show a clear improvement with both the monthly and drift calibration applied as indicated by a better agreement with the 1:1 line (Fig. 5) for most of the data and a reduction in the MAE (Fig. 6). However, the monthly and drift calibrations did not improve the R^2 or slope for the sensors in the IE at MLVB and RIVR. The uncalibrated sensors displayed a good R^2, slope and MAE which indicated the standard factory sensor calibration transferred to the field well at MLVB and RIVR. Also, the PM sensor does not exhibit significant instrumental drift over the 12-month period. Calibrating the sensor data against the proxy however, seemed to have introduced errors. There are several reasons for this.

Firstly, Fig. 6 shows that the MAE between the collocated reference data and the proxy data is larger at RIVR then the MAE for the uncalibrated data against the collocated reference data indicating that the MLVB proxy was not always suitable for MOMA calibration of the RIVR sensor. This is also supported by the differing probability distributions from the two sites (Fig. S2) which suggests the sites are exposed to different PM levels. On the other hand, the probability distributions for

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

CELA and NHOL $PM_{2.5}$ data and that for CMPT and SLB are very similar (Fig. S2) and hence the MOMA calibration process produces improved accuracy.

Secondly, monthly variability in particle source and composition will impact the reliability of the MOMA calibration particularly for those performed at monthly intervals. For example, the very high monthly MOMA MAE for February at CELA, MVLB and RIVR suggests the January particle composition was not representative of that observed in February at these sites. Particle composition is known to vary with different wind directions (desert vs. marine/urban particles) and impact the sensor reading as observed in previous studies (Castell et al., 2017; Gao et al., 2015; Giordano et al., 2021; Kelly et al., 2017). The effect of this phenomenon is particularly visible between November and February when wind is more variable. This is supported by Fig. 6, which shows that for both the LA and IE regions the MAE tended to be higher in November/December and January for uncalibrated as well as calibrated data. The difference between the proxy and the collocated reference data also tended to be larger during these months.

A similar month-to-month variability in the MAE can be observed when comparing the regulatory monitor (BAM 1020, Met One Instruments, Inc., Grants Pass, Oregon, US) at RIVR against the reference grade optical instruments T640 (Teledyne API, San Diego, US) and the GRIMM optical particle counter (EDM 180, GRIMM Aerosol Technik GmbH & Co., Airing, Germany), also located at the RIVR site. The T640 and GRIMM are both optical particle counter instruments that determine the aerosol particle size distribution from which they estimate the PM concentration. The BAM-1020 samples aerosols through a PM_{10} inlet and uses a Very Sharp Cut Cyclone (VSCC) to classify it into $PM_{2.5}$ before collecting it on a filter tape and determining the $PM_{2.5}$ concentration by the aerosol's attenuation of a C_{14} beta radiation source (Hagler et al., 2022). Due to the differences in the measurement principles, the instruments can give different results depending on the properties of the measured particles.

The T640 and GRIMM match each other consistently across the year (similar technologies) but the BAM/T640 and BAM/GRIMM MAE are higher in general and highest during the November/December months. This further shows how differences between measurement technologies will be exacerbated when particle composition is variable. This is discussed in more detail in sect. 3.5.

Thirdly, measurement noise in the hourly reference data from the beta attenuation monitors deployed at the sites may be too high to reliably calibrate low-cost sensors when concentrations are low (< 40 µg m^{-3}) as often the case in the RC Desert (Hagler et al., 2022; Johnson et al., 2018; Zheng et al., 2018). The calibration improved the data most during the summer months with the MAE equal or below 5 µg m^{-3}.

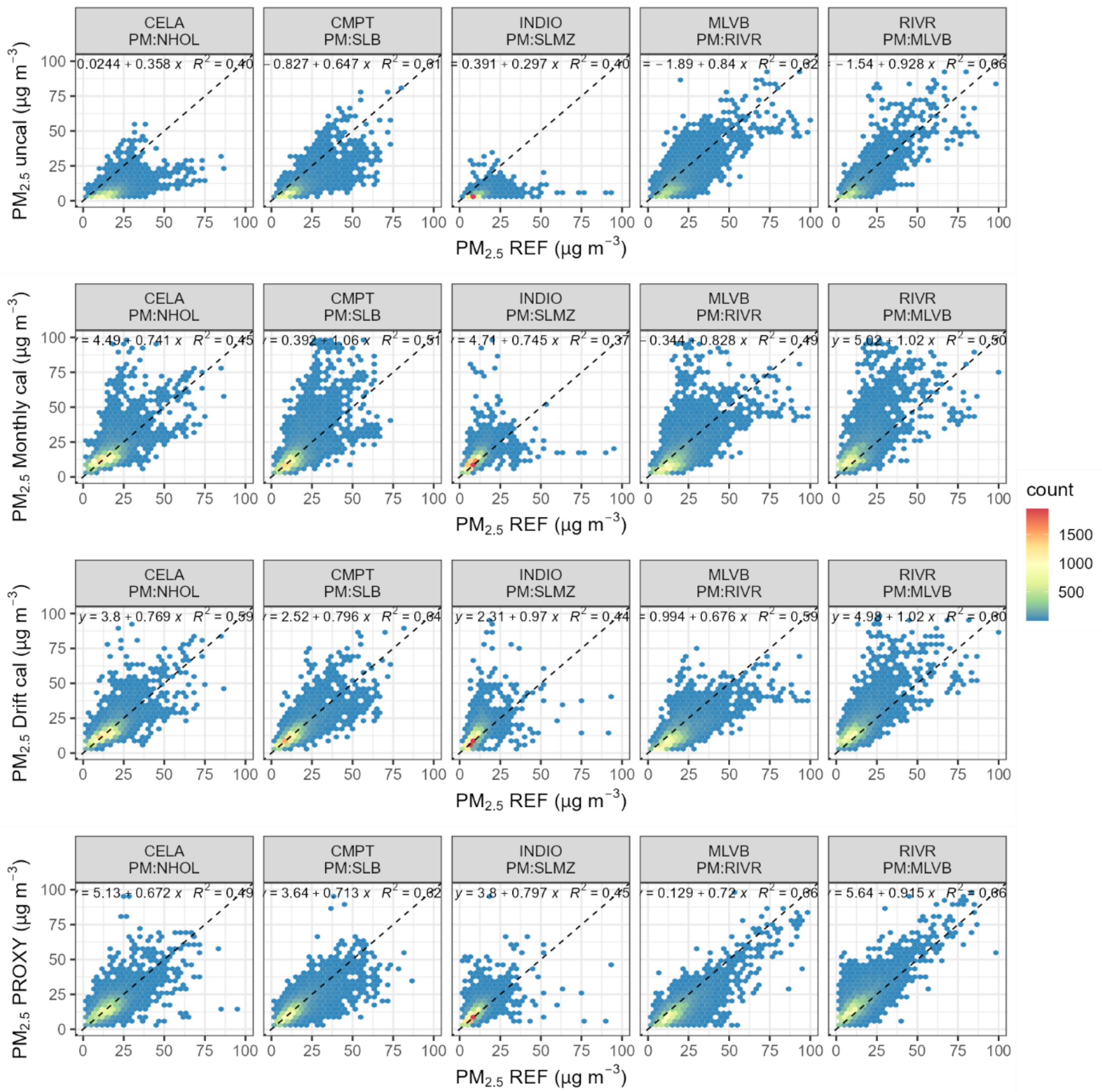


Figure 5: Hexbin scatterplots showing hourly uncalibrated + RH corrected, monthly calibrated and drift calibrated $PM_{2.5}$ measured by the AQY and calibrated against the nearest proxy site vs. the collocated South Coast AQMD AMS over 12 months. The colours refer to the number of points within each bin.

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

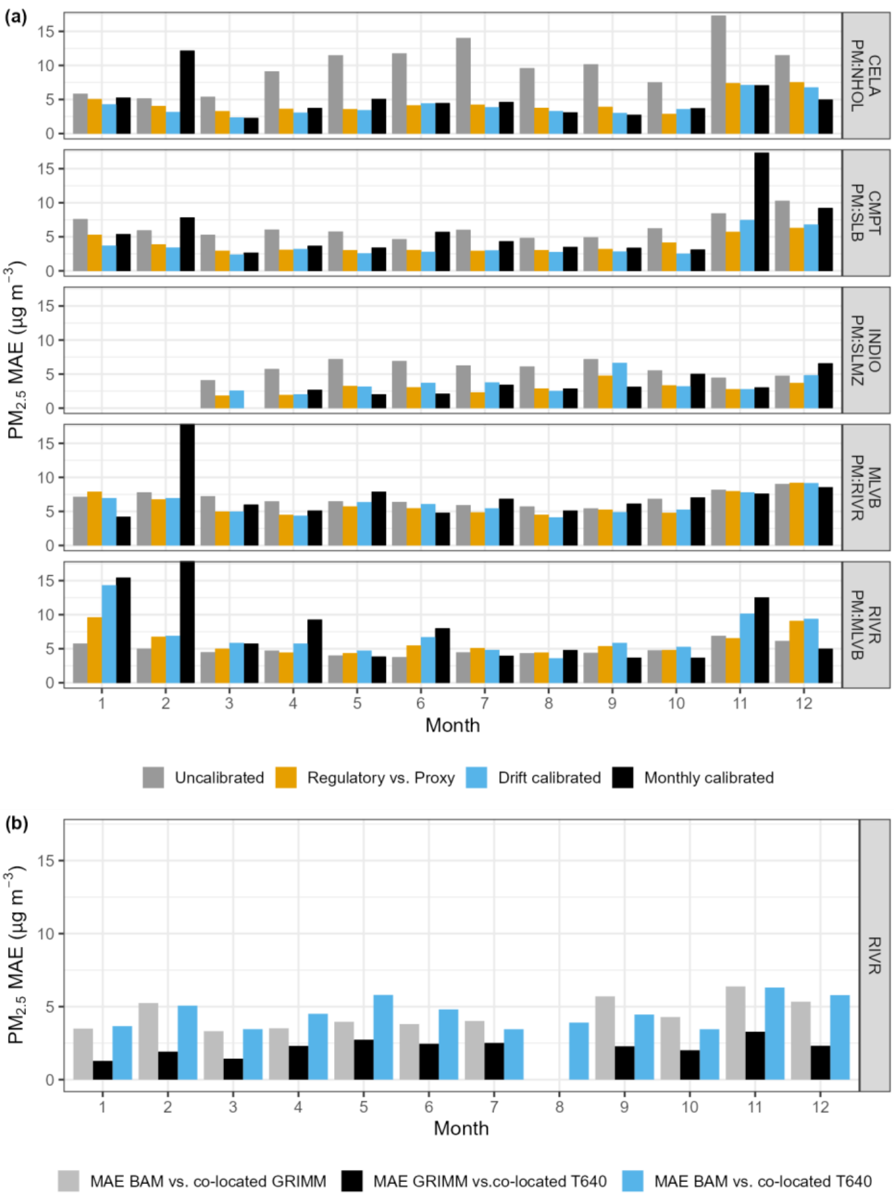


Figure 6: a) monthly calibrated and drift calibrated $PM_{2.5}$ data as well as for the collocated reference data versus the proxy reference. b) MAE between the BAM and the collocated T640 and GRIMM across different months using hourly averaged data.

3.3.2 PM_{10}

As expected, the PM_{10} data from the sensors generally showed a poorer agreement with the collocated Reference with relatively large MAE and low R^2 (Fig. 7) for uncalibrated data. The uncalibrated data were also considerably underestimating, particularly in the RC Desert (INDIO, PALM). This is in agreement with previous work which showed that

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

the SDS011 underestimates PM_{10}, particularly for particles greater than 5 µm (Budde et al., 2018; Kuula et al., 2020). The monthly and drift triggered MOMA calibrations had a clear positive impact on PM_{10} and improved the accuracy as indicated by a better fit around the 1:1 line. However, the scatter remained resulted in no improvement in the R^2. Examination of the Proxy/REF scatterplots (Fig. 7) and probability distributions (SI Fig. 2) show there are considerable discrepancies between sites. To some extent this is expected since the PM coarse fraction (PM_{10} – $PM_{2.5}$) is more dominated by local sources than $PM_{2.5}$ (Pinto et al., 2004). Similar to $PM_{2.5}$, there was month-to-month variability in the calibration performance, with better improvements during summer and poor performance in November, particularly in the IE and RC Desert (Fig. 8). Potential factors that contribute to the large MAE in November are further discussed in sect. 3.5.

A comparison of the PM_{10} data from the reference instruments at RIVR (BAM, GRIMM, T640) shows that the MAE across different instrument types can be as high as ~15 µg/m^3 and interestingly the GRIMM and T640 PM_{10} MAE is the highest – the opposite of the $PM_{2.5}$ result.

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

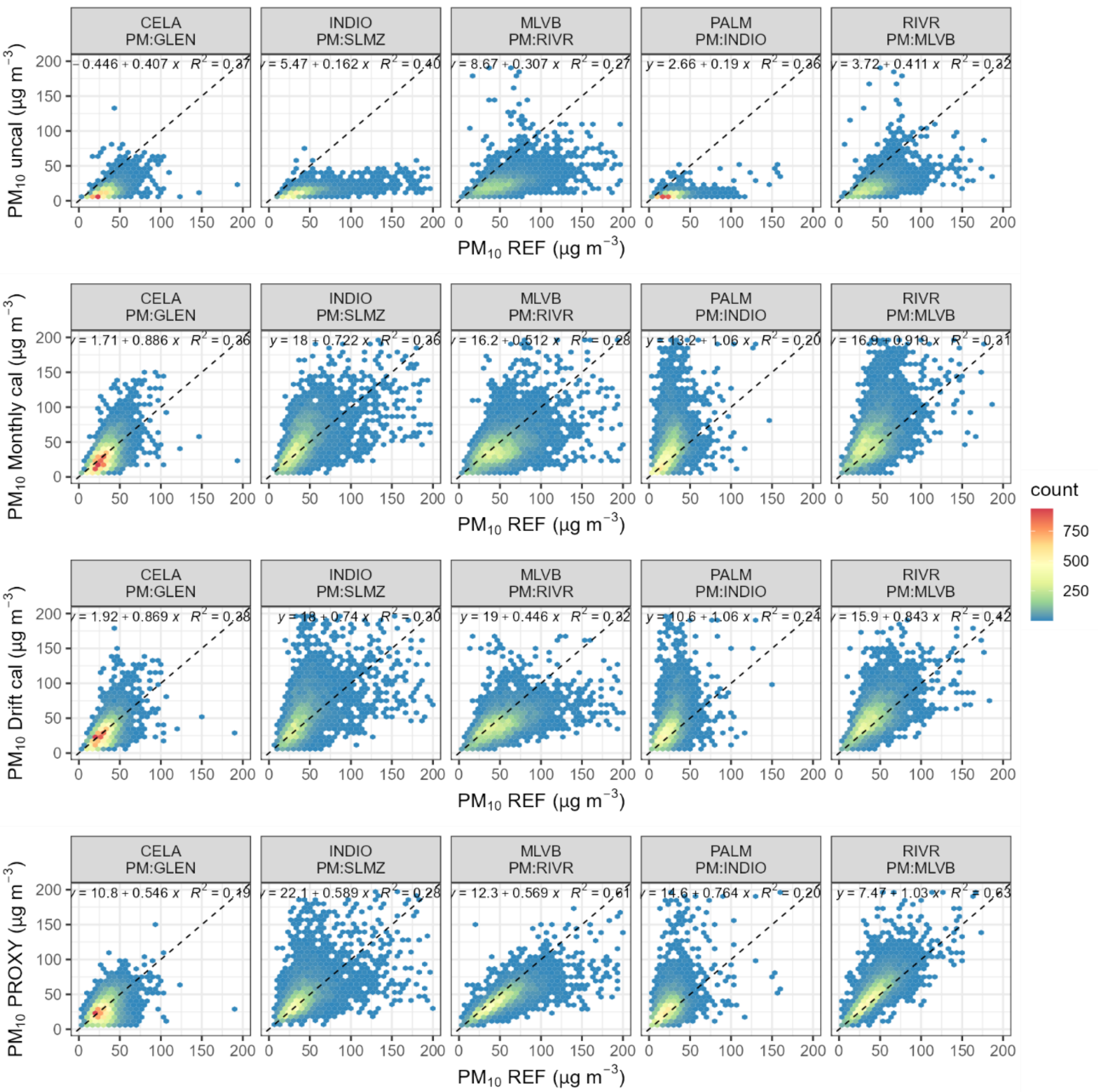


Figure 7: Hexbin scatterplots showing hourly uncalibrated + RH corrected, monthly calibrated and drift calibrated PM_{10} measured by the AQY and calibrated against the nearest proxy site vs. the collocated Reference. The colours refer to the number of points within each bin.

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

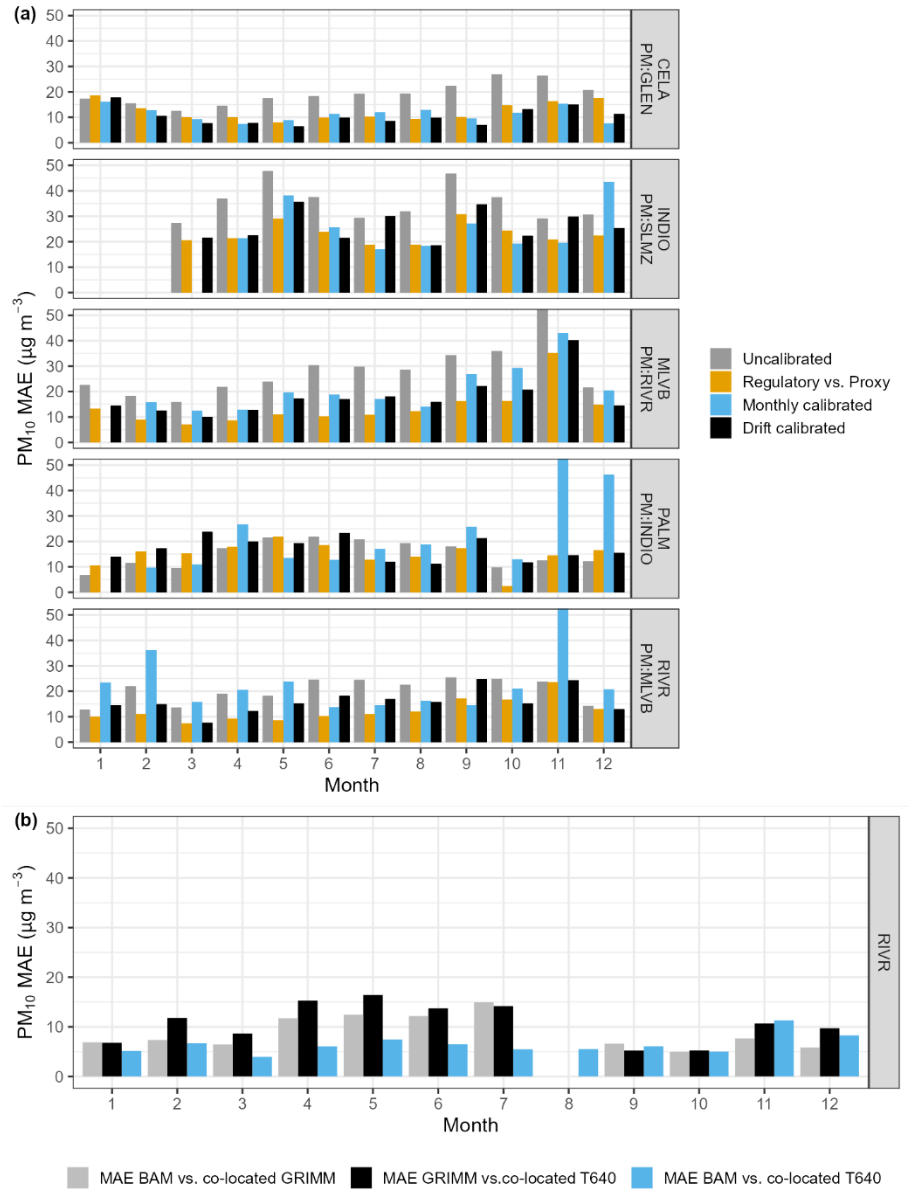


Figure 8: a) monthly calibrated and drift calibrated PM_{10} data as well as for the collocated reference data versus the proxy reference. b) MAE between the BAM and the collocated T640 and GRIMM across different months using hourly averaged data.

3.4 Drift detection triggers

The results from the drift detection framework tests are shown in Fig. 9 (K-S test, MV-intercept test and MV-slope test) for $PM_{2.5}$ and PM_{10} measured by a PM sensor deployed in the LA region and one in the IE region. The black points indicate when the framework triggered a drift alarm and calibration. It is evident that most alarms were raised due to significant

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

differences in the probability distributions (K-S test), followed by a change in the slope between the proxy and sensor (MV-slope test). In the IE (RIVR: AQY BD-1146) alarms were related to changes in the MV-slope and clustered around February, May, and September/October indicating more frequent changes in environmental conditions (e.g., RH) or particle composition and size during these months (discussed in sect. 3.5). The AQY sensor (BD-1069) installed at the CELA AMS sent off alarms that were more spread across the whole year suggesting that sensor drift at this site was not related to seasons. The figure also shows that there are more frequent slope adjustments within a month likely due to within month changes in meteorological and environmental conditions (discussed in sect. 3.5). This partly explains the better performance of the drift calibrated data compared to the monthly calibrated data, which triggered more frequent calibrations within a month.

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

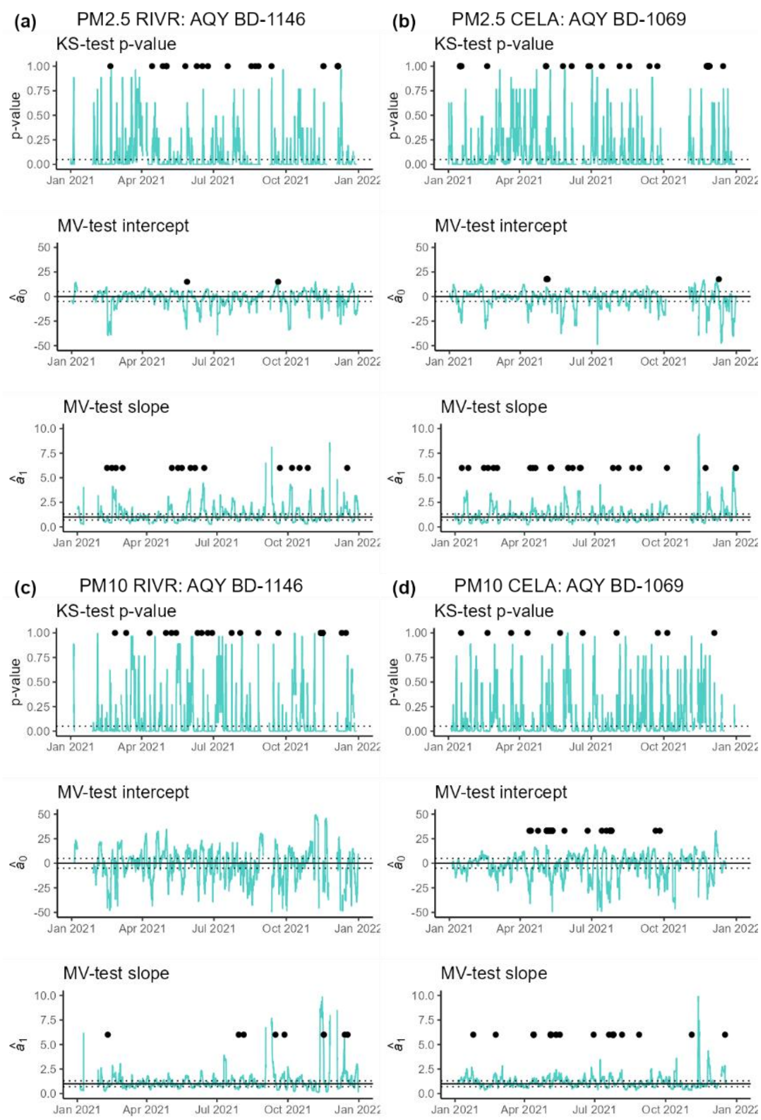


Figure 9: Test statistics from drift detection framework for a site in the IE (a) / (c) and one in the City of LA region (b) / (d) for $PM_{2.5}$ and PM_{10}, respectively. The black points show when the drift detection framework resulted in an alarm and triggered a calibration.

Figure 10 shows the temporal variability of monthly and drift calculated gains for sensors in the LA, IE and RC Desert Region. Interestingly, the temporal variation of the $PM_{2.5}$ and PM_{10} gains calculated the monthly calibrations (Fig. 10. (b)/(d) show a distinct seasonal pattern with larger gains (~2-3) during autumn and winter and lower gains (~1) during the summer months, particularly in the IE region. An opposite pattern is visible in the RC Desert. The gains from the drift detection framework were more variable but also showed some seasonal variability. These results suggest that unlike calibrating for

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

sensor drift (which would be shown as an increase in the slope over time) PM sensors are calibrated for different conditions, which can vary frequently as shown by the drift gains.

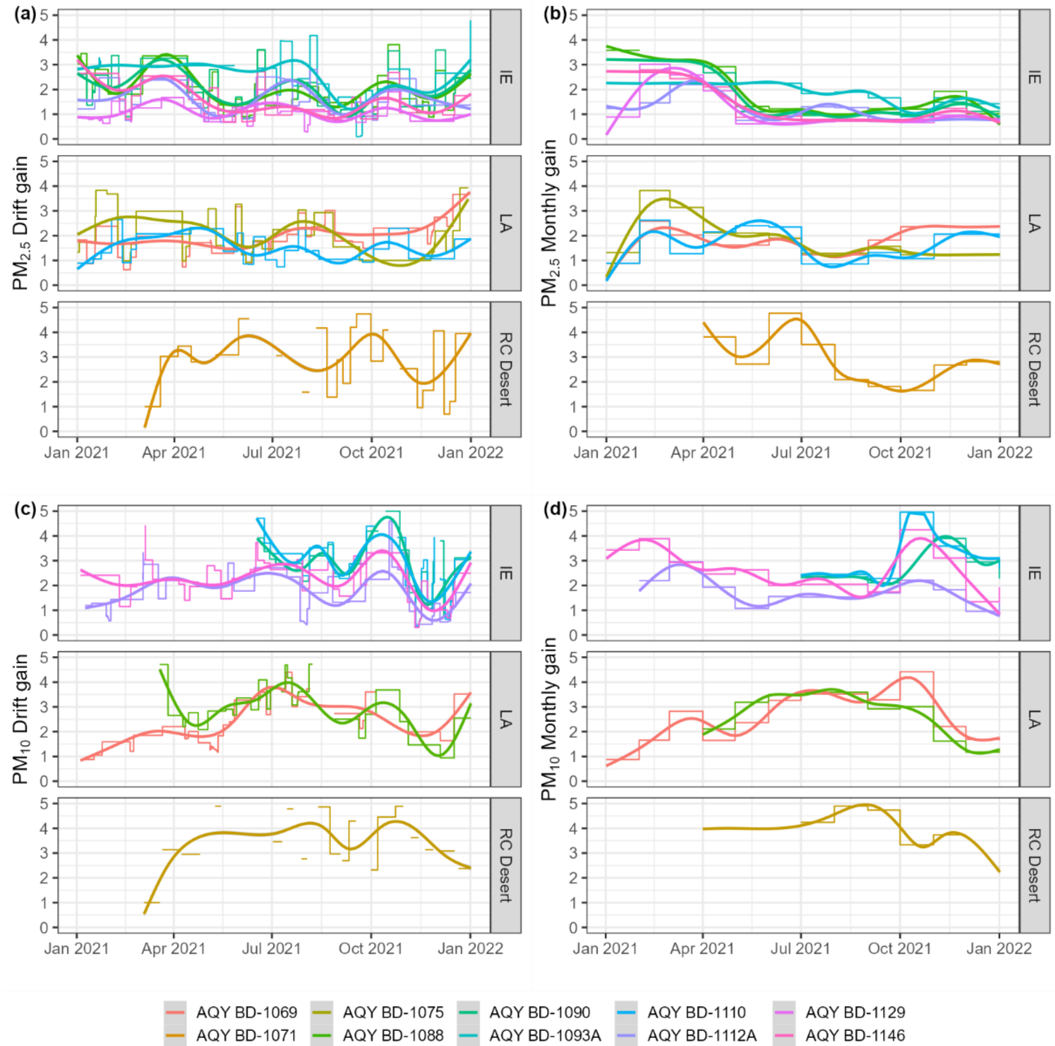


Figure 10: Temporal variation of the gains as calculated from the drift detection framework and the monthly calibrations for $PM_{2.5}$ a) and b), respectively and PM_{10} c) and d), respectively.

3.5 Particle composition variability

As observed in the previous sections, calibrating PM sensors can be challenging in complex areas where particle composition, size and physical properties (i.e., shape and refractive index) vary spatially and temporally (Kuula et al., 2020). In this section, we discuss some of the origins for the variations in particle composition with a specific focus on the Riverside area (RIVR AMS).

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

The RIVR AMS wind data shown in Fig. 11, clearly indicates the seasonal variation in the wind direction with N/NE winds dominating during the late autumn/winter months and W winds dominating during the rest of the year. The N/NE winds likely correspond to the Santa Ana Winds which are associated with very dry downslope air flow from the northeast and common between October and April, with a peak in December and January (Aguilera et al., 2020). Typically, PM concentrations during SAW conditions are dominated by coarse particles of crustal components (Guazzotti et al., 2001; Qin et al., 2012).

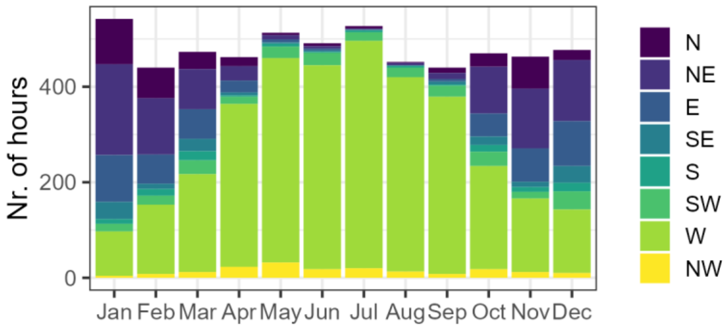


Figure 11: Nr. of hours dominated by different wind direction during each month. Surface Meteorological Data was downloaded from the NOAA Integrated Surface Database (ISD) via the worldmet Package in R (Carslaw, 2022).

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

This is in agreement with observations from Fig. 12 which shows higher concentrations of Crustal Material and Elemental Carbon during N/NE and NW, reaching a maximum in November. Trace ions (Chloride, Sodium and Potassium ion) and secondary ions (Nitrate, Sulfate, Ammonium), on the other hand, are highest downwind from the City of LA reaching a maximum in spring/summer due to increased photochemical activity and a larger contribution of sulfate sources and its precursor (fuel/ship emissions) upwind of the City of LA (Daher et al., 2013).

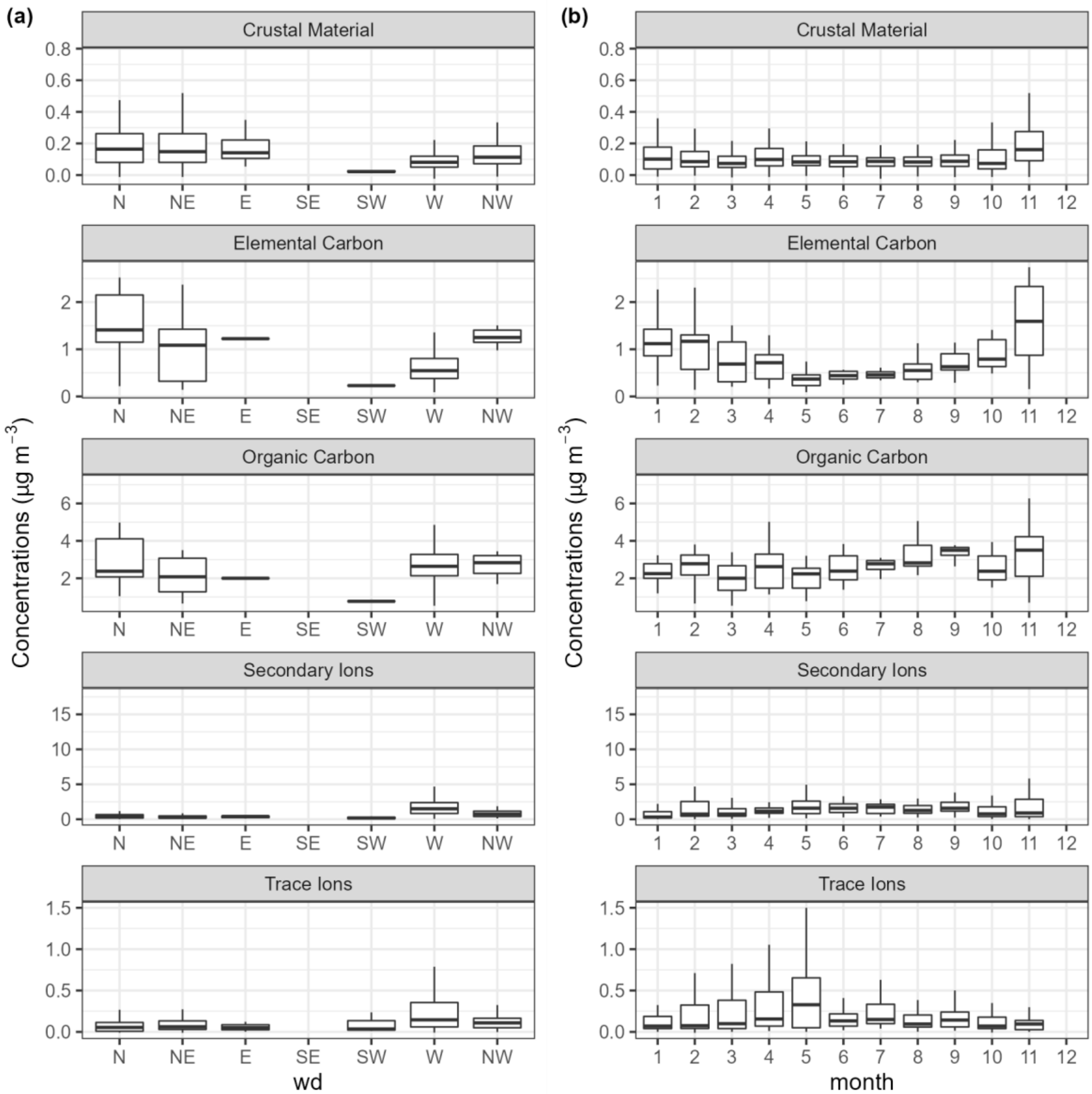


Figure 12: Speciation concentrations collected at AMS – Rubidoux (RIVR) grouped into 5 categories (Panels) plotted against wind direction (wd) (a) and for each month of the year (b).

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

Figure 13 illustrates the relationship between the BAM and collocated sensor data coloured by wind direction and course fraction (1 – $PM_{2.5}/PM_{10}$). The figure reveals a clear slope dependence on the wind direction (<1 when wind was from a northeast origin and >1 when wind from a western origin dominated), suggesting that it underestimates $PM_{2.5}$ levels during north-eastern wind (SAW conditions). It is also visible from Fig. 11 that wind is more variable in late fall/winter possibly explaining the more frequent alarms observed for these months at Riverside (Fig. 9).

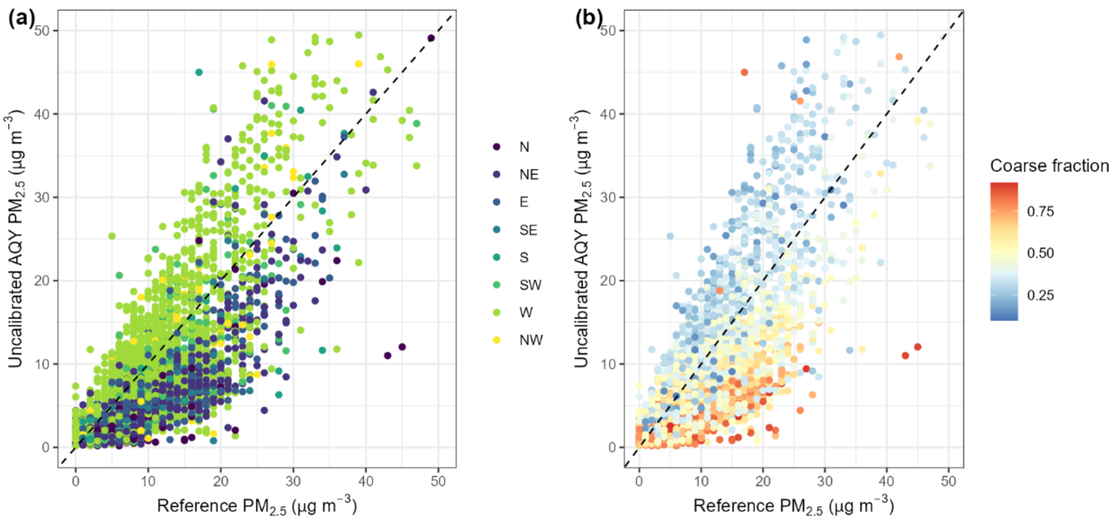


Figure 13: Hourly uncalibrated low-cost sensor data against hourly collocated reference data at AMS - Rubidoux (RIVR) during 2021, a) coloured by wind direction, b) coloured by the AQY PM coarse fraction: 1 – $PM_{2.5}/PM_{10}$.

4 Conclusions

This work is part of a large study set out to determine if a remote calibration framework (MOMA), previously developed for the correction of drift in O_3 and NO_2 sensors (Miskell et al., 2018, 2019; Weissert et al., 2020) can be applied for $PM_{2.5}$ and PM_{10} data from PM sensors. We identified suitable reference proxies based on distance and presented two approaches to remotely calibrate data from sensor networks, 1) at monthly intervals and 2) using a drift detection framework that triggers a calibration when drift is detected. Our results show that while both approaches were able to improve the data as indicated by a better fit around the 1:1 line when compared to the collocated reference data, the drift triggered MOMA approach performed better. Overall, the improvement due to the MOMA calibration was more obvious for PM_{10} data, which were considerably underestimating prior to calibration. We note that sensor drift was less associated with monitor operational

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

factors and more affected by variations in particle composition which exacerbated differences in response between the regulatory BAM instruments and the PM sensors.

Calibrating at monthly intervals was not always sufficient, particularly if wind conditions were variable within a month. This was clearly visible in the IE where particle composition varied from desert dust (N/NE) and marine/urban aerosol (W) during the winter months. This highlights the need for reference instruments to be deployed at sites representing different land use and PM source types which would allow a more flexible choice of proxies depending on dominant wind direction and particle source.

Appendix

Supplementary Information

Data and code availability

10-min, 1-hr, and 24-hr averaged data from the SCAQMD sensor network can be exported from https://aqportal.aqmd.gov/. The code is not publicly accessible due to intellectual property.

Author contributions

G.S.H., D.E.W, V.P. formulation of overarching research goals and aims.; D.E.W., L.F.W. and G.S.H. developed the methodology; B.F., A.C.-O. and R.L. managed and maintained the sensor network, L.F.W. developed the software and performed the data analysis, L.F.W. prepared the manuscripts with contributions from all co-authors. V.P., A.P. and G.S.H. supervised the project.

Competing interests

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: L.F.W. and G.S.H. are employees of Aeroqual Ltd, manufacturer of the sensor nodes used in these studies. G.S.H. and D.E.W. are founders and shareholders in Aeroqual Ltd.

References

Aguilera, R., Gershunov, A., Ilango, S. D., Guzman-Morales, J., and Benmarhnia, T.: Santa Ana Winds of Southern California Impact $PM_{2.5}$ With and Without Smoke From Wildfires, GeoHealth, 4, https://doi.org/10.1029/2019GH000225, 2020.

© Author(s) 2023. CC BY 4.0 License.

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.

Anderson, J. O., Thundiyil, J. G., and Stolbach, A.: Clearing the Air: A Review of the Effects of Particulate Matter Air Pollution on Human Health, J. Med. Toxicol., 8, 166–175, https://doi.org/10.1007/s13181-011-0203-1, 2012.

Aphalo, P. J., Slowikowski, K., and Mouksassi, S.: ggpmisc: Miscellaneous Extensions to "ggplot2," 2022.

Ault, A. P., Moore, M. J., Furutani, H., and Prather, K. A.: Impact of Emissions from the Los Angeles Port Region on San Diego Air Quality during Regional Transport Events, Environ. Sci. Technol., 43, 3500–3506, https://doi.org/10.1021/es8018918, 2009.

Badura, M., Batog, P., Drzeniecka-Osiadacz, A., and Modzel, P.: Evaluation of Low-Cost Sensors for Ambient $PM_{2.5}$ Monitoring, J. Sens., 2018, 1–16, https://doi.org/10.1155/2018/5096540, 2018.

Budde, M., Schwarz, A. D., Mueller, T., Laquai, B., Streibl, N., Schindler, G., Koepke, M., Riedel, T., Dittler, A., and Beigl, M.: Potential and Limitations of the Low-Cost SDS011 Particle Sensor for Monitoring Urban Air Quality | ProScience, in: ProScience, 3rd International Conference on Atmospheric Dust - DUST2018, 6–12, 2018.

Carslaw, D.: worldmet: Import Surface Meteorological Data from NOAA Integrated Surface Database (ISD), 2022.

Carslaw, D. and Ropkins, K.: openair: Tools for the Analysis of Air Pollution Data, 2022.

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, https://doi.org/10.1016/j.envint.2016.12.007, 2017.

Chen, D., Liu, X., Han, J., Jiang, M., Wang, Z., and Qi, J.: A New Angular Light Scattering Measurement of Particulate Matter Mass Concentration for Homogeneous Spherical Particles, Sensors, 19, 2243, https://doi.org/10.3390/s19102243, 2019.

Daher, N., Hasheminassab, S., Shafer, M. M., Schauer, J. J., and Sioutas, C.: Seasonal and spatial variability in chemical composition and mass closure of ambient ultrafine particles in the megacity of Los Angeles, Env. Sci Process. Impacts, 15, 283–295, https://doi.org/10.1039/C2EM30615H, 2013.

De Vito, S., Esposito, E., Castell, N., Schneider, P., and Bartonova, A.: On the robustness of field calibration for smart air quality monitors, Sens. Actuators B Chem., 310, 127869, https://doi.org/10.1016/j.snb.2020.127869, 2020.

Eeftens, M., Beelen, R., de Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., Declercq, C., Dėdelė, A., Dons, E., de Nazelle, A., Dimakopoulou, K., Eriksen, K., Falq, G., Fischer, P., Galassi, C., Gražulevičienė, R., Heinrich, J., Hoffmann, B., Jerrett, M., Keidel, D., Korek, M., Lanki, T., Lindley, S., Madsen, C., Mölter, A., Nádor, G., Nieuwenhuijsen, M., Nonnemacher, M., Pedeli, X., Raaschou-Nielsen, O., Patelarou, E., Quass, U., Ranzi, A., Schindler, C., Stempfelet, M., Stephanou, E., Sugiri, D., Tsai, M.-Y., Yli-Tuomi, T., Varró, M. J., Vienneau, D., Klot, S. von, Wolf, K., Brunekreef, B., and Hoek, G.: Development of Land Use Regression Models for $PM_{2.5}$, $PM_{2.5}$ Absorbance, PM_{10} and PM_{coarse} in 20

© Author(s) 2023. CC BY 4.0 License.

European Study Areas; Results of the ESCAPE Project, Environ. Sci. Technol., 46, 11195–11205, https://doi.org/10.1021/es301948k, 2012.

Gao, M., Cao, J., and Seto, E.: A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of $PM_{2.5}$ in Xi'an, China, Environ. Pollut., 199, 56–65, https://doi.org/10.1016/j.envpol.2015.01.013, 2015.

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, J. Aerosol Sci., 158, 105833, https://doi.org/10.1016/j.jaerosci.2021.105833, 2021.

Guazzotti, S. A., Whiteaker, J. R., Suess, D., Coffee, K. R., and Prather, K. A.: Real-time measurements of the chemical composition of size-resolved particles during a Santa Ana wind episode, California USA, Atmos. Environ., 35, 3229–3240, https://doi.org/10.1016/S1352-2310(01)00140-6, 2001.

Habre, R., Girguis, M., Urman, R., Fruin, S., Lurmann, F., Shafer, M., Gorski, P., Franklin, M., McConnell, R., Avol, E., and Gilliland, F.: Contribution of tailpipe and non-tailpipe traffic sources to quasi-ultrafine, fine and coarse particulate matter in southern California, J. Air Waste Manag. Assoc., 71, 209–230, https://doi.org/10.1080/10962247.2020.1826366, 2021.

Hagler, G., Hanley, T., Hassett-Sipple, B., Vanderpool, R., Smith, M., Wilbur, J., Wilbur, T., Oliver, T., Shand, D., Vidacek, V., Johnson, C., Allen, R., and D'Angelo, C.: Evaluation of two collocated federal equivalent method $PM_{2.5}$ instruments over a wide range of concentrations in Sarajevo, Bosnia and Herzegovina, Atmospheric Pollut. Res., 13, 101374, https://doi.org/10.1016/j.apr.2022.101374, 2022.

Hofman, J., Nikolaou, M., Shantharam, S. P., Stroobants, C., Weijs, S., and La Manna, V. P.: Distant calibration of low-cost PM and NO_2 sensors; evidence from multiple sensor testbeds, Atmospheric Pollut. Res., 13, 101246, https://doi.org/10.1016/j.apr.2021.101246, 2022.

Johnson, K. K., Bergin, M. H., Russell, A. G., and Hagler, G. S. W.: Field Test of Several Low-Cost Particulate Matter Sensors in High and Low Concentration Urban Environments, Aerosol Air Qual. Res., 18, 565–578, https://doi.org/10.4209/aaqr.2017.10.0418, 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, https://doi.org/10.1016/j.envpol.2016.12.039, 2017.

Kim, J. and Kwan, M.-P.: How Neighborhood Effect Averaging Might Affect Assessment of Individual Exposures to Air Pollution: A Study of Ozone Exposures in Los Angeles, Ann. Am. Assoc. Geogr., 111, 121–140, https://doi.org/10.1080/24694452.2020.1756208, 2021.

Kloog, I., Nordio, F., Coull, B. A., and Schwartz, J.: Incorporating Local Land Use Regression And Satellite Aerosol Optical Depth In A Hybrid Model Of Spatiotemporal $PM_{2.5}$ Exposures In The Mid-Atlantic States, Environ. Sci. Technol., 46, 11913–11921, https://doi.org/10.1021/es302673e, 2012.

© Author(s) 2023. CC BY 4.0 License.

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

Lee, H. J., Chatfield, R. B., and Strawa, A. W.: Enhancing the Applicability of Satellite Remote Sensing for $PM_{2.5}$ Estimation Using MODIS Deep Blue AOD and Land Use Regression in California, United States, Environ. Sci. Technol., 50, 6546–6555, https://doi.org/10.1021/acs.est.6b01438, 2016.

Liang, L.: Calibrating low-cost sensors for ambient air monitoring: Techniques, trends, and challenges, Environ. Res., 197, 111163, https://doi.org/10.1016/j.envres.2021.111163, 2021.

Liu, H.-Y., Schneider, P., Haugen, R., and Vogt, M.: Performance Assessment of a Low-Cost $PM_{2.5}$ Sensor for a near Four-Month Period in Oslo, Norway, Atmosphere, 10, 41, https://doi.org/10.3390/atmos10020041, 2019.

Loh, B. G. and Choi, G. H.: Calibration of Portable Particulate Matter–Monitoring Device using Web Query and Machine Learning, Saf. Health Work, 10, 452–460, https://doi.org/10.1016/j.shaw.2019.08.002, 2019.

Mccrowey, C., Sharac, T., Mangus, N., Jager, D., Brown, R., Garver, D., Wells, B., and Brittingham, H.: RAQSAPI: A Simple Interface to the US EPA Air Quality System Data Mart API, 2022.

Miskell, G., Salmond, J., Alavi-Shoshtari, M., Bart, M., Ainslie, B., Grange, S., McKendry, I. G., Henshaw, G. S., and Williams, D. E.: Data Verification Tools for Minimizing Management Costs of Dense Air-Quality Monitoring Networks, Environ. Sci. Technol., 50, 835–846, https://doi.org/10.1021/acs.est.5b04421, 2016.

Miskell, G., Salmond, J. A., and Williams, D. E.: Solution to the Problem of Calibration of Low-Cost Air Quality Measurement Sensors in Networks, ACS Sens., 3, 832–843, https://doi.org/10.1021/acssensors.8b00074, 2018.

Miskell, G., Alberti, K., Feenstra, B., Henshaw, G. S., Papapostolou, V., Patel, H., Polidori, A., Salmond, J. A., Weissert, L., and Williams, D. E.: Reliable data from low cost ozone sensors in a hierarchical network, Atmos. Environ., 214, 116870, https://doi.org/10.1016/j.atmosenv.2019.116870, 2019.

Pinto, J. P., Lefohn, A. S., and Shadwick, D. S.: Spatial Variability of $PM_{2.5}$ in Urban Areas in the United States, J. Air Waste Manag. Assoc., 54, 440–449, https://doi.org/10.1080/10473289.2004.10470919, 2004.

Qin, X., Pratt, K. A., Shields, L. G., Toner, S. M., and Prather, K. A.: Seasonal comparisons of single-particle chemical mixing state in Riverside, CA, Atmos. Environ., 59, 587–596, https://doi.org/10.1016/j.atmosenv.2012.05.032, 2012.

Sardar, S. B.: Seasonal and spatial variability of the size-resolved chemical composition of particulate matter (PM_{10}) in the Los Angeles Basin, J. Geophys. Res., 110, D07S08, https://doi.org/10.1029/2004JD004627, 2005.

Slowikowski, K., Schep, A., Hughes, S., Dang, T. K., Lukauskas, S., Irisson, J.-O., Kamvar, Z. N., Ryan, T., Christophe, D., Hiroaki, Y., Gramme, P., Abdol, A. M., Barrett, M., Cannoodt, R., Krassowski, M., Chirico, M., and Aphalo, P.: ggrepel: Automatically Position Non-Overlapping Text Labels with "ggplot2," 2022.

Kahle D., Wickham H. "ggmap: Spatial Visualization with ggplot2." *The R Journal*, 5(1), 144–161. https://journal.r-project.org/archive/2013-1/kahle-wickham.pdf, 2013.

https://doi.org/10.5194/egusphere-2023-969
Preprint. Discussion started: 30 May 2023
© Author(s) 2023. CC BY 4.0 License.

Snyder, E. G., Watkins, T. H., Solomon, P. A., Thoma, E. D., Williams, R. W., Hagler, G. S. W., Shelow, D., Hindin, D. A., Kilaru, V. J., and Preuss, P. W.: The Changing Paradigm of Air Pollution Monitoring, Environ. Sci. Technol., 47, 11369–11377, https://doi.org/10.1021/es4022602, 2013.

Spinu, V., Grolemund, G., Wickham, H., Vaughan, D., Lyttle, I., Costigan, I., Law, J., Mitarotonda, D., Larmarange, J., Boiser, J., and Lee, C. H.: lubridate: Make Dealing with Dates a Little Easier, 2022.

Weissert, L., Miles, E., Miskell, G., Alberti, K., Feenstra, B., Henshaw, G. S., Papapostolou, V., Patel, H., Polidori, A., Salmond, J. A., and Williams, D. E.: Hierarchical network design for nitrogen dioxide measurement in urban environments, Atmos. Environ., 228, 117428, https://doi.org/10.1016/j.atmosenv.2020.117428, 2020.

Wickham, H. and RStudio: tidyverse: Easily Install and Load the "Tidyverse," 2022.

Wickham, H., François, R., Henry, L., Müller, K., and RStudio: dplyr: A Grammar of Data Manipulation, 2022a.

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., and RStudio: ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics, 2022b.

Williams, D. E.: Low Cost Sensor Networks: How Do We Know the Data Are Reliable?, ACS Sens., 4, 2558–2565, https://doi.org/10.1021/acssensors.9b01455, 2019.

Witiw, M. R. and LaDochy, S.: Trends in fog frequencies in the Los Angeles Basin, Atmospheric Res., 87, 293–300, https://doi.org/10.1016/j.atmosres.2007.11.010, 2008.

Zeileis, A., Grothendieck, G., Ryan, J. A., Ulrich, J. M., and Andrews, F.: zoo: S3 Infrastructure for Regular and Irregular Time Series (Z's Ordered Observations), 2022.

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, Aerosols/In Situ Measurement/Instruments and Platforms, https://doi.org/10.5194/amt-2018-111, 2018.