

Article

Not peer-reviewed version

---

# Varying Performance of Low-Cost Sensors During Seasonal Smog Events in Moravian-Silesian Region

---

[Václav Nevrlý](#)<sup>\*</sup>, Michal Dostál, Petr Bitala, Vit Klecka, Jiří Sléžka, Pavel Polach, Katarína Nevrlá, [Melánie Barabášová](#), Růžena Langová, [Sarka Bernatikova](#), Barbora Martiníková, [Michal Vašínek](#), [Adam Nevrlý](#), [Milan Lazecký](#), [Jan Suchánek](#), [Hana Chaloupecká](#), David Kiča, Jan Wild

Posted Date: 3 October 2024

doi: 10.20944/preprints202410.0258.v1

Keywords: air quality; low-cost sensors; Polish smog; Saharan dust storm; linear regression



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Varying Performance of Low-Cost Sensors During Seasonal Smog Events in Moravian-Silesian Region

Václav Nevrlý <sup>1,2</sup> , Michal Dostál <sup>1,2</sup> , Petr Bitala <sup>1</sup> , Vít Klečka <sup>1</sup> , Jiří Slézka <sup>1</sup>, Pavel Polách <sup>1</sup>, Katarína Nevrlá <sup>1</sup>, Melánie Barabášová <sup>1</sup> , Růžena Langová <sup>1</sup>, Šárka Bernatíková <sup>1</sup> , Barbora Martiníková <sup>1</sup> , Michal Vašínek <sup>3</sup> , Adam Nevrlý <sup>3</sup>, Milan Lazecký <sup>4</sup> , Jan Suchánek <sup>2</sup> , Hana Chaloupecká <sup>5</sup> , David Kiča <sup>6</sup> and Jan Wild <sup>1</sup> 

<sup>1</sup> Faculty of Safety Engineering, VSB-Technical University of Ostrava, Lumírova 630/13, 700 30 Ostrava-Výškovice, Czech Republic

<sup>2</sup> J. Heyrovský Institute of Physical Chemistry, Czech Academy of Sciences, Dolejškova 2155/3, 182 23 Praha 8, Czech Republic

<sup>3</sup> Faculty of Electrical Engineering and Computer Science, VSB-Technical University of Ostrava, 17.listopadu 15/2172, 708 00 Ostrava-Poruba, Czech Republic

<sup>4</sup> IT4Innovations, VSB—Technical University of Ostrava, 17. listopadu 15/2172, 708 00 Ostrava-Poruba, Czech Republic

<sup>5</sup> Institute of Thermomechanics of the Czech Academy of Sciences, Dolejškova 5, 182 00 Praha 8, Czech Republic

<sup>6</sup> Public Health Institute Ostrava, Partyzánské Náměstí 2633/7, Moravská Ostrava, Ostrava, 702 00, Czech Republic

\* Correspondence: vaclav.nevrlý@vsb.cz; Tel.: +420-59-732-2872 (V.N.)

**Abstract:** This study examines the varying performance of low-cost sensors during seasonal smog events, particularly those resulting from solid fuel combustion linked to residential or industrial heating and other pollution sources. A prototype low-cost sensor node developed within this work was collocated for several months with reference equipment at an air quality monitoring station in the city of Ostrava (Moravian-Silesian region, Czech Republic). Deterioration of air quality during a pair of smog alert events, namely wintertime "Polish smog" episode and springtime "Saharan dust storm", which occurred through the extended period of co-location, was characterized quantitatively based on particulate matter and carbon monoxide concentrations measured by the low-cost sensors. These datasets were compared with reference data and model predictions provided by the Copernicus Atmospheric Monitoring Service. Selected low-cost sensors demonstrated linear response during Polish smog episode with highly correlated time series of carbon monoxide and particulate matter concentrations. However, when compared to reference data, the Sensirion SPS30 sensor is found to overestimate coarse particle (PM<sub>10</sub>) concentrations by the factor of two during winter evaluation period. On the other hand, oppositely biased PM<sub>10</sub> response (underestimation by the factor of five) was determined from co-location measurement during Saharan dust storm event. This observation was attributed to combination of specific detection efficiency of the given low-cost sensor and shifts in particle size distribution due to seasonal change of ambient conditions.

**Keywords:** air quality; low-cost sensors; Polish smog; Saharan dust storm; linear regression

## 1. Introduction

The Moravian-Silesian Region in the Czech Republic, particularly the city of Ostrava, is recognized as a significant air pollution hot spot in Europe. This situation arises historically from a combination of industrial activity, high population density, and geographical factors that worsen air quality levels. The topography of the Moravian-Silesian Region contributes to the accumulation of pollutants, especially during winter months when temperature inversions are common. This leads to poor dispersion, resulting in elevated concentrations of air pollutants. Additionally, air pollution from neighboring areas, particularly from the Silesian Voivodeship (Poland), also plays a significant role in the given context [1–3].

Recently, "Polish smog" was proposed as a specific type of air pollution that occurs in Poland, particularly in the winter months. It is characterized by high concentrations of particulate matter (PM), such as PM<sub>2.5</sub> and PM<sub>10</sub>, as well as polycyclic aromatic hydrocarbons like benzo(a)pyrene [4]. This type of smog is particularly prevalent in Eastern Europe, where it arises from the burning of coal

and other solid fuels for heating purposes, especially during winter months. When compared with photo-chemical smog found in industrialized urban areas, which is driven mainly by volatile organic compounds (VOCs) and nitrogen oxides leading to high ozone levels, Polish smog is more closely linked to residential heating practices and industrial emissions.

Yet, another type of smog event, which is further considered through this paper is caused by particulate matter originating from Sahara desert. Dust storms can transport particles over thousands of kilometers, affecting air quality far from their source. The transport of Saharan dust poses significant challenges for air quality management and public health in Europe. Although this phenomenon is more common in the Mediterranean region, it can occasionally cause a significant deterioration in air quality for Central European countries.

Low-cost sensor (LCS) networks have emerged as a promising solution for monitoring air pollution and providing smog alerts as it can supplement data from regulatory-grade reference instruments [5]. By filling in spatial and temporal gaps in air quality monitoring such information can provide a more comprehensive understanding of pollution patterns both at local and regional level. Citizen science projects involving the public in sensor deployment can further expand the reach of these networks [6–8]. However proper calibration of LCS is crucial to ensure data accuracy and reliability [9].

In principle, uncertainties of factory calibrated LCS response are studied by experiments in controlled (laboratory) or uncontrolled (field) ambient conditions [10–12]. Such calibration procedure is followed by selection of suitable numerical correction method and estimation of parameters using reference datasets for model training or testing purposes [13]. Co-location of the LCS node with reference instruments in real outdoor atmosphere is usually performed over the period of several weeks in order to accumulate appropriately large datasets required for the reliable outputs of the calibration process. Prolonged period of co-location enables to investigate seasonal variability of LCS performance across a wide range of environmental conditions relevant for the target locality.

Following questions are mainly addressed through the paper: Which information (e.g. in relevance to public smog alerts) can be extracted from data provided by LCS and CAMS model predictions and what is their case-specific reliability compared to reference data from air quality monitoring (AQM) station?

## 2. Materials and Methods

Our prototype LCS sensor node (more details described in Appendix A) was mounted on the roof of AQM station, see Figure 1. This setup allows for direct comparison between the outputs of the LCS node and data from reference-grade instruments.



**Figure 1.** LCS sensor node placed on the roof of the reference air quality monitoring station of the Health Institute Ostrava located in the Mariánské Hory district.

### 2.1. Selection of Low-Cost Sensors

Following our goals, LCS node is equipped with a set of low-cost sensors suitable for monitoring of target air pollutants for the specific area, which is particulate matter (PM), carbon monoxide (CO). Selection of the gas and PM sensors was mainly based on the extensive literature survey and experience of previous investigators. Availability of the sensors (distribution in EU) were also taken into account as well as affordability of the entire LCS node setup and level of complexity relevant to requirements for its integration and further development.

For particulate matter (PM) measurement, the node utilizes two sensors: Sensirion SPS30 and Alphasense OPC-N3. The SPS30 is a laser-based optical sensor well-suited for measuring fine particulate matter (especially  $PM_{10}$ ) mass concentration based on the principle of light scattering. The OPC-N3 also uses optical particle counting, when detecting a wider range of particle sizes, from  $0.35\ \mu m$  to  $40\ \mu m$ , across 24 size bins enabling particle size distribution to be determined, which is critical for understanding the composition of atmospheric aerosols and determination of their origin in relevance to source apportionment.

For carbon monoxide monitoring, the LCS node incorporates the Alphasense CO-B4 electrochemical sensor. This sensor can detect CO concentrations between 0 to 1,000 ppm with a resolution of 0.1 ppm. Its sensitivity ranges from 55 to 85 nA/ppm, providing accurate detection of small changes in CO levels.

Knowledge of atmospheric pressure, temperature and humidity is essential when aiming at corrections of LCS response in various ambient conditions. For this purpose, digital module including Bosch Sensortech BME280 sensor was integrated in LCS node. This sensor should operate with pressure accuracy of  $\pm 1\ hPa$ , temperature accuracy of  $\pm 1^\circ C$ , and humidity accuracy of  $\pm 3\% RH$ .

### 2.2. Co-Location Site and Reference Instruments

The co-location measurements of the LCS node were conducted at an air quality monitoring (AQM) station of Health Institute located in the municipal area of the city of Ostrava, which is close to various industrial sites (e.g. metallurgical, chemical, etc.). The AQM station provides the information on meteorological conditions, i.e. wind speed, wind direction as well as reference air quality data in hourly intervals. Atmospheric pressure, temperature, and humidity are measured using the COMET T3113D sensor for temperature and humidity, and the NXP Semiconductor MPX4115A for pressure.

For CO measurement, the HORIBA APMA-370 analyzer is used, which operates on the nondispersive infrared (NDIR) spectroscopy principle. The TEOM 1400 analyzer was utilized as PM reference during entire winter evaluation period including S1 episode. This instrument measures  $PM_{10}$  by

drawing air through a filter mounted on an oscillating balance, with the mass of collected particles altering the oscillation frequency to calculate PM concentration. At the end of March 2024, the TEOM 1400 was replaced by Palas FIDAS 200 analyzer, which employs optical scattering technology for continuous real-time reference measurement of size distribution and following quantification of PM<sub>1</sub>, PM<sub>2.5</sub>, and PM<sub>10</sub>.

### 2.3. Overview of Co-Location Measurement

Low-cost sensors are often collocated with reference instruments in laboratory or field conditions for period of several weeks in order to improve their performance. In our case, an evaluation measurement campaign lasting for three months (from mid-November 2023 to mid-February 2024) was initially planned to be carried out in order to verify and validate the performance of individual LCS and their variation for seasonal meteorological conditions typical in the Moravian-Silesian region.

Accidentally, smog alert events, which occurred in Ostrava during December 2023 and April 2024 were also recorded during extended period of co-location. These datasets enabled us to focus our attention on performance of LCS and CAMS data versus reference measurements during the episodes of serious air quality deterioration, see Table 1.

**Table 1.** Smog episodes with relative humidity, temperature and pressure characterized by their mean and standard deviation values (in brackets).

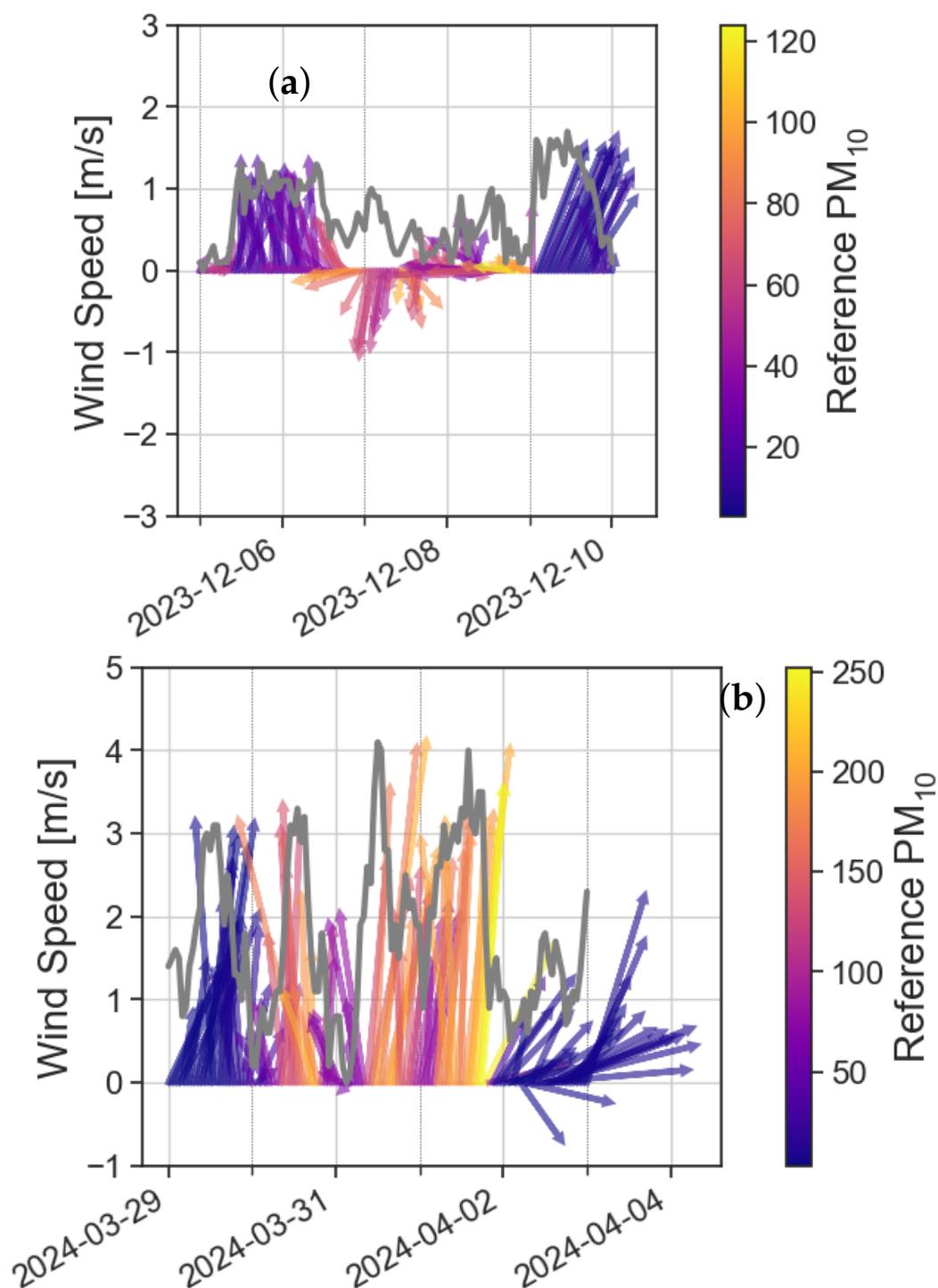
Period	Start	End	Hum. (%)	Temp. (°C)	Pres. (hPa)
S1	2023-12-05	2023-12-09	98(±2)	-2(±2)	1016(±3)
S2	2024-03-29	2024-04-03	45(±15)	15(±10)	1004(±4)

It is worth to note that atmospheric conditions for S1 episode is typical for above mentioned "Polish smog". In the given case the highest concentrations of particulate matter are generally recorded at low temperatures, specifically between -10°C and 0°C. Additionally, higher atmospheric pressure correlates with increased PM concentrations, as stable air masses inhibit vertical mixing and allow pollutants to accumulate near the surface. [4]. Evolution of above described meteorological situation can be identified from Figure ?? based on reference data from AQM station.

Different atmospheric conditions for S2 episode corresponds to seasonal Saharan dust storm over Central Europe<sup>1</sup>. Moderate speed of wind blowing from south-west direction can be clearly identified during PM<sub>10</sub> maxima from the Figure 2.

---

<sup>1</sup> For more detailed information on the given topic, see <<https://atmosphere.copernicus.eu/climate-atmosphere-podcast-understanding-impact-saharan-dust-storms>>



**Figure 2.** Temporal evolution of wind speed and wind orientation during S1 (a) and S2 (b) episodes respectively. Wind speed is plotted by grey solid line in positive scale. Wind orientation vectors are represented by arrows have length proportional to wind speed and color relevant to  $[PM_{10}]_{REF}$  concentration.

#### 2.4. Datasets and Preparatory Analysis

Reference 1-hour averaged data were extracted from monthly datasets provided by the Health Institute (AQM station) after verification involving and replacement of values of measurement below limit of detection ( $LoD$ ) by the value equal to  $LoD/2$ .

Similarly, co-location site-specific timeseries of concentrations for selected pollutant (CO, O<sub>3</sub>, NO<sub>2</sub>, PM<sub>10</sub>, PM<sub>2.5</sub> and Dust) based on the forecast of the CAMS ENSAMBLE model [14] with 11 km spatial resolution were downloaded in the form of comma-separated value (CSV) files from [Open Meteo](#) webpage. Monthly datasets extracted from LCS node (10-minute averages of sensor readings available from [SOASENSE](#) project webpage) were converted into hourly timeseries using pandas (Python library) resample method and ordered with reference and CAMS model data according to relevant GMT timestamps. Multilinear fitting (MLR) of the carbon monoxide LCS voltage were converted into concentration using linear model form `scikit-learn` (Python library). Entire dataset of winter measurements, i.e. from November 2023 to February 2024 was assumed as representative for the given step. Temperature reading from LCS node was converted into Kelvins in order to avoid numerical issues relevant to negative values. Entire dataset was splitted into training and testing subsets having 1232 and 2392 datapoints respectively. Finally, MLR model having high coefficient of determination, i.e.  $R^2 = 0.89$  and acceptable Mean Average Error ( $MAE < 50 \mu\text{g}/\text{m}^3$ ) for predicted CO concentration was estimated assuming only T[K] and CO-B4 sensor voltage values as predictors with presumed parametrization of the Equation 1:

$$[CO]_{LCS}(\tau) = c_1 \times [U_{WE}(\tau) - U_{AE}(\tau)] + c_2 \times T(\tau) + c_3 \quad (1)$$

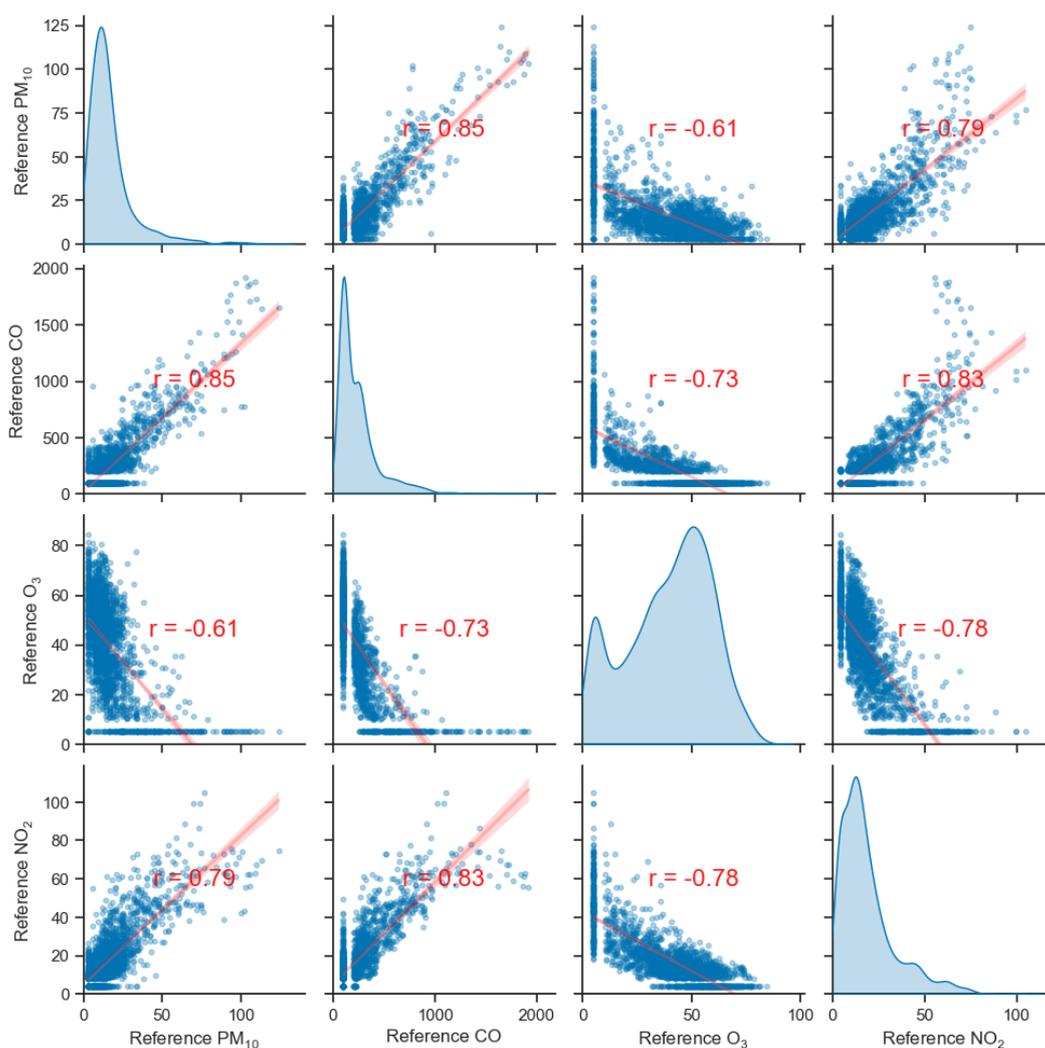
where  $U_{WE}(\tau)$  is a working electrode voltage [V],  $U_{AE}(\tau)$  is an auxillary electrode voltage [V] at a given Greenwich (i.e. GMT) time  $\tau$  and case-specific values of relevant MLR coefficients are as follows:  $c_1 = 4394$ ,  $c_2 = -0.693$ ,  $c_3 = 0$ .

Exploratory data analysis including Correlation Matrix and Kernel Density Estimation (KDE) of dataset pairs was performed using `seaborn` (Python library). Subsequently, simple linear regression (SLR) and plots of diurnal variations of air pollutant concentrations were performed employing relevant methods implemented in `atmospy` (Python library). Particle size distribution measured by Alphasense OPC-N3 sensor were analyzed using the `smps-py` (Python library).

### 3. Results

#### 3.1. Key Findings from Winter Evaluation Period

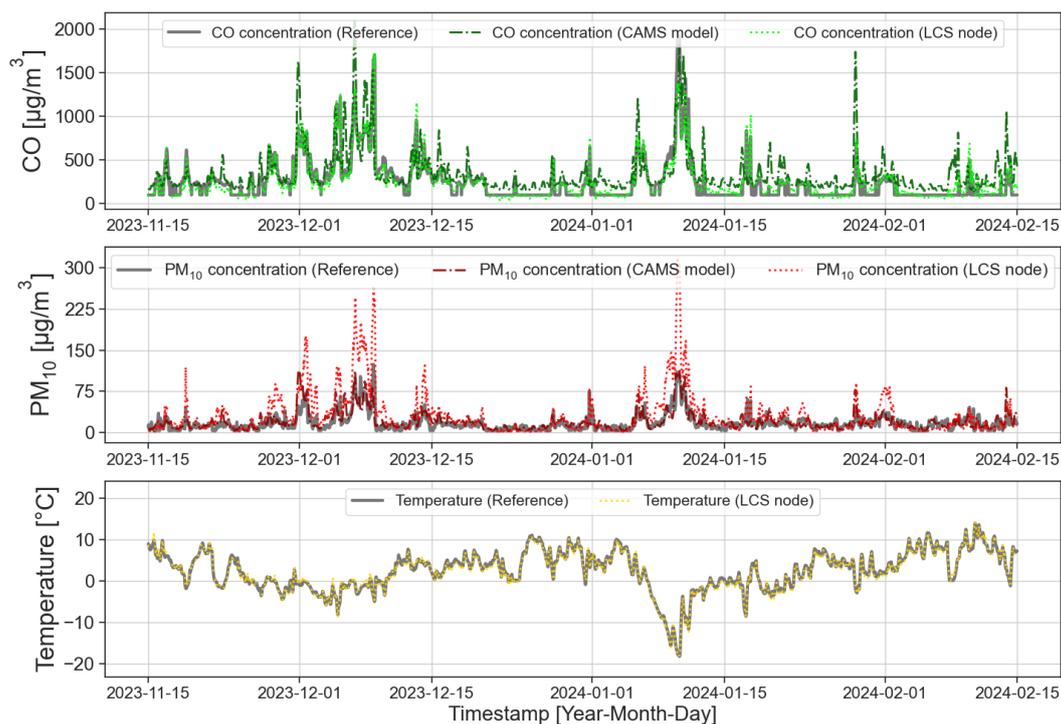
In winter, the time series for key air pollutants are expected to be significantly correlated, which was confirmed from analysis of the reference data (PM<sub>10</sub>, CO, NO<sub>2</sub>, O<sub>3</sub>). Carbon monoxide shows strong linear correlation with PM<sub>10</sub> having a Pearson coefficient value of  $r = 0.85$ , see Figure 3.



**Figure 3.** Correlation matrix with scatter plots, linear regression and histograms for reference pollutant concentrations from winter evaluation period. All scales are depicted in  $[\mu\text{g}/\text{m}^3]$  units. Pearson correlation coefficients  $r$  are depicted in red for non-diagonal subplots.

At the same time, the  $\text{NO}_2$  concentration values are also correlated with  $\text{CO}$  and  $\text{PM}_{10}$ , which can be attributed to a similar source of these pollutants or to a related mechanism. This observation is in line with an assumption, that incomplete combustion of fossil fuels and subsequent atmospheric dispersion of smoke plumes is dominant contributor to winter air pollution in the given location. As can be expected, due to the lower solar (UV) radiation intensity, ozone concentrations only rarely reached elevated values (above  $75 \mu\text{g}/\text{m}^3$ ) during this period. However, the values of Pearson coefficient for the negative correlation of  $\text{O}_3$  with other pollutants suggest that the ozone reactivity influence atmospheric chemistry even in the colder part of the year, mostly in the case of  $\text{NO}_2$ .

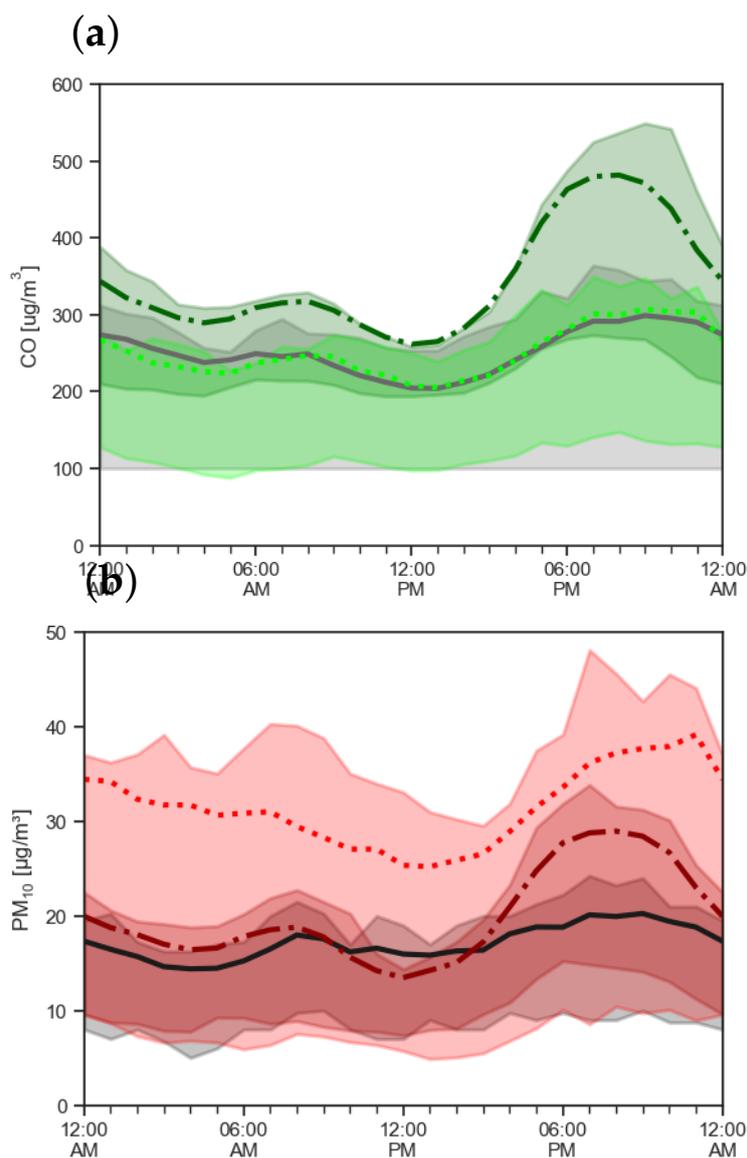
Most prominent qualitative features of the LCS response and selected CAMS model predictions can be recognised from the plots of selected hourly data (Figure 4) when compared to reference measurements. Firstly, we can spot excellent fit of  $[\text{CO}]_{\text{LCS}}$  to  $[\text{CO}]_{\text{REF}}$  together with systematic shift of  $[\text{CO}]_{\text{CAMS}}$  data towards slightly higher values. On the contrary,  $\text{PM}_{10}$  response provided by SPS30 sensor is considerably overestimated (especially for  $\text{PM}_{10} > 50 \mu\text{g}/\text{m}^3$ ) compared to  $\text{PM}_{10}$  prediction by CAMS model which is very close to reference data.



**Figure 4.** Comparison of the selected hourly data series from winter evaluation period.

Negative correlation of major pollutant concentration ambient temperature is also evident from the plot depicted above. Numerous spikes in CO and PM<sub>10</sub> timeseries follows sharp decrease of ambient temperature (mostly below 0°C). We can also conclude excellent agreement of temperature from LCS node with reference, which is important due to requirement of correcting CO-B4 sensor according to Equation 1.

The overall quality of the LCS node data and the qualitative indicators of the CAMS model compared to the reference measurements are also illustrated by the plots shown in Figure 5. The average reference CO concentration over the evaluation period shows the same trend of diurnal variation as PM<sub>10</sub>.



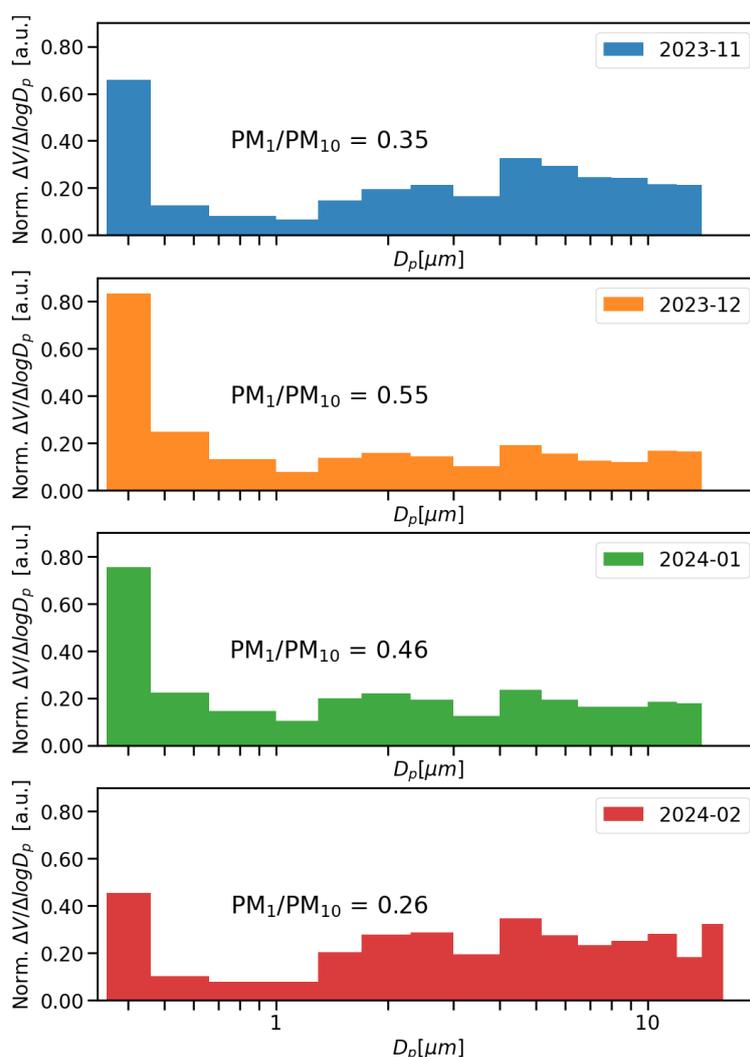
**Figure 5.** Plot of diurnal variations of CO concentration (a) and PM<sub>10</sub> (b) during winter evaluation period extracted from reference instrument (solid line), LCS node (dotted line) and CAMS model (dash-dotted line) data with mean value (thick lines) and the interquartile range (shaded regions).

In the case of carbon monoxide LCS data shows nearly perfect coincidence with the reference diurnal trend. There is obvious increase of CO and PM<sub>10</sub> during morning rush hours (at about 8 AM) and second peak around 8 PM. While for the morning peak the  $[\text{PM}_{10}]_{\text{CAMS}}$  value (predicted by CAMS) coincides with the reference measurement of PM<sub>10</sub> ( $[\text{PM}_{10}]_{\text{REF}}$ ).

As in the case of the CO diurnal variation plot, the peak PM<sub>10</sub> concentration observed in the late evening is significantly overestimated by the CAMS model, which may be related to uncertainty in emission factors or data related to industrial and domestic heating, which is dominant in the winter. For example, the Czech Republic is currently undergoing technological improvements based on government support for the replacement of domestic heating systems, which may have already been effective but has not yet been reflected in the model inputs. The plot of diurnal SPS30 response, i.e.,  $[\text{PM}_{10}]_{\text{LCS}}$ , shows a large overestimation with a more pronounced deviation during the night hours.

In the following sections of this paper, we anticipate that this change is due to a combination of the specific sensitivity of the SPS30 sensor and the changing ratio of fine to coarse particles in the air over the course of the day and year due to different intensities of domestic and industrial combustion and smoke dispersion, see Section 4.

Prior to the start of this study, we assumed that the Alphasense OPC-N3 sensor integrated in the LCS node would be used to accurately determine PM concentrations as complementary measurement to the TEOM reference. However, it has been found that OPC-N3 response do not show adequate agreement with the reference instrument ( $R^2 < 0.3$  for  $PM_{10}$  from the evaluation period). Nevertheless, the particle size distribution was continuously measured by this sensor throughout the co-location period. Thus we used these data to determine the particle size-resolved spectra. Particle volume concentrations for each of 24 bins were determined from measured values (i.e. number of particles counted per sensor bin) and their daily mean values were recalculated. Despite considerable noise and uncertainty in the response of few bins (e.g. between 3-4  $\mu\text{m}$ ), some seasonal trend of fine and coarse particle concentration is evident from data depicted in the Figure 6.

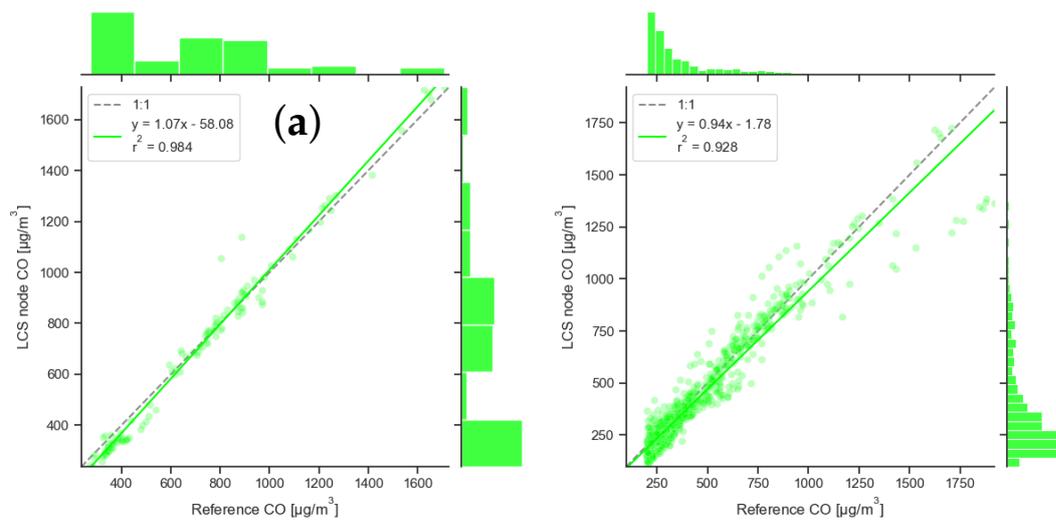


**Figure 6.** Seasonal variation of size distribution depicted as the normalized particle volume by bin of the Alphasense OPC-N3 sensor. The value of mass-weighted  $PM_1/PM_{10}$  ratio was estimated for each co-location month based on median value of the relevant 24-h averages.

### 3.2. LCS Performace During Polish Smog

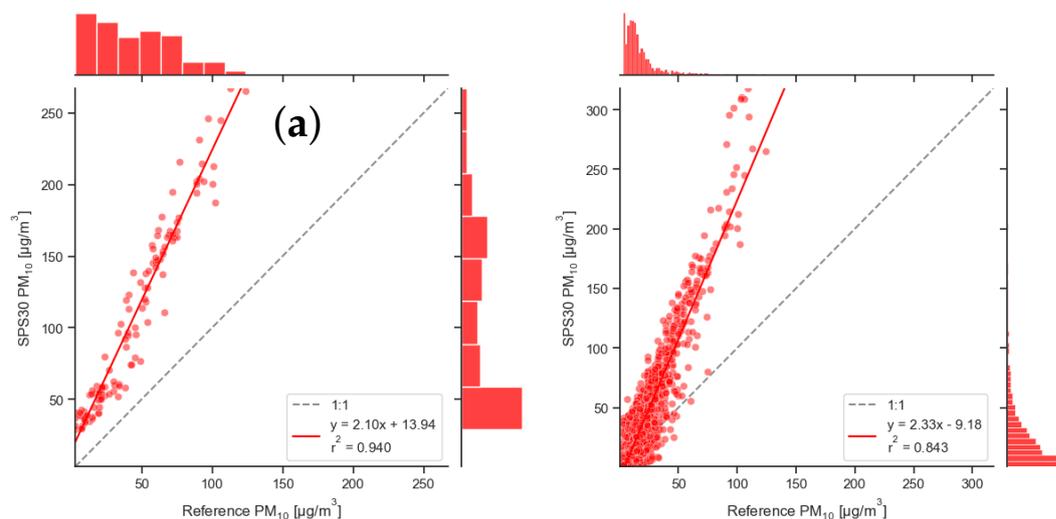
Figure 7 shows highly linear response of the CO-B4 sensor during S1 episode. The slope parameter is nearly equal to its value obtained for entire evaluation period. This fact can be expected due to presence of high concentrations datapoints recorded during the smog episode. However, we can find close agreement of the data with initial  $[CO]_{LCS}$  calibration according to Equation 1. Note that data

points recorded, when the values measured by the reference instrument were below the detection limit (i.e.  $[CO]_{REF} < 200 \mu\text{g}/\text{m}^3$ ) were removed before the SLR analysis.



**Figure 7.** Simple linear regression of Alphasense CO-B4 sensor response versus reference instrument (HORIBA) for "Polish smog" episode S1 (a) compared with data for winter evaluation period (b). The histograms displayed adjacent to the axes illustrate the normalized frequency of the measured concentration ranges within the respective dataset.

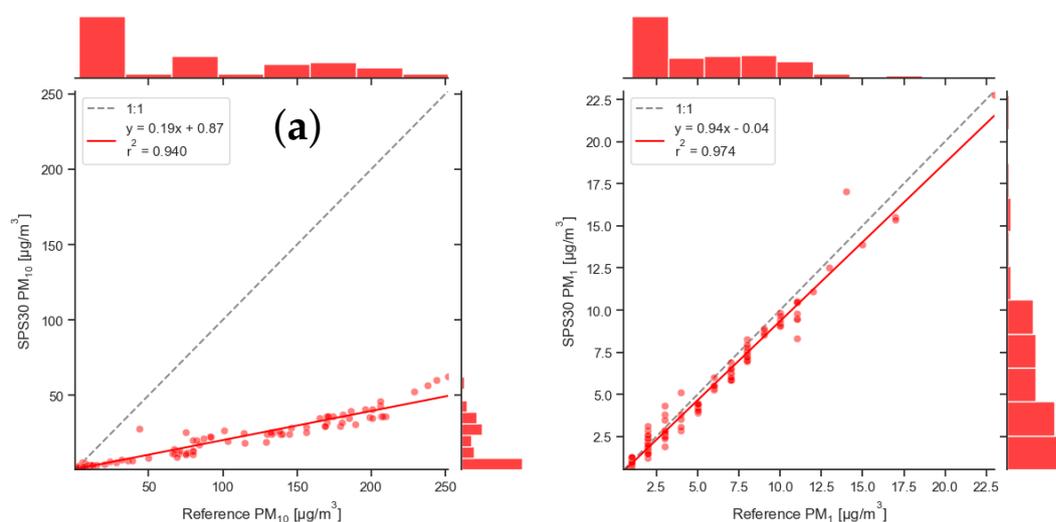
Systematically over-predicted  $PM_{10}$  response was obtained from SPS30 sensor compared to TEOM reference instrument during S1 episode. Nevertheless, similar value of slope parameter  $a$  was obtained from SLR fit (where  $a$  is the slope of  $y = ax + b$  regression line) for the entire winter evaluation period, see Figure 8). These results indicate consistency of datasets obtained from the SPS30 sensor response together with relatively high coefficient of determination ( $R^2 > 0.8$ ) for  $PM_{10}$  over entire co-location period.



**Figure 8.** Simple linear regression of Sensirion SPS30 sensor response versus reference instrument (TEOM) for "Polish smog" episode S1 (a) compared with data for winter evaluation period (b). The histograms displayed adjacent to the axes illustrate the normalized frequency of the measured concentration ranges within the respective dataset.

### 3.3. LCS Performance During Saharan Dust Storm

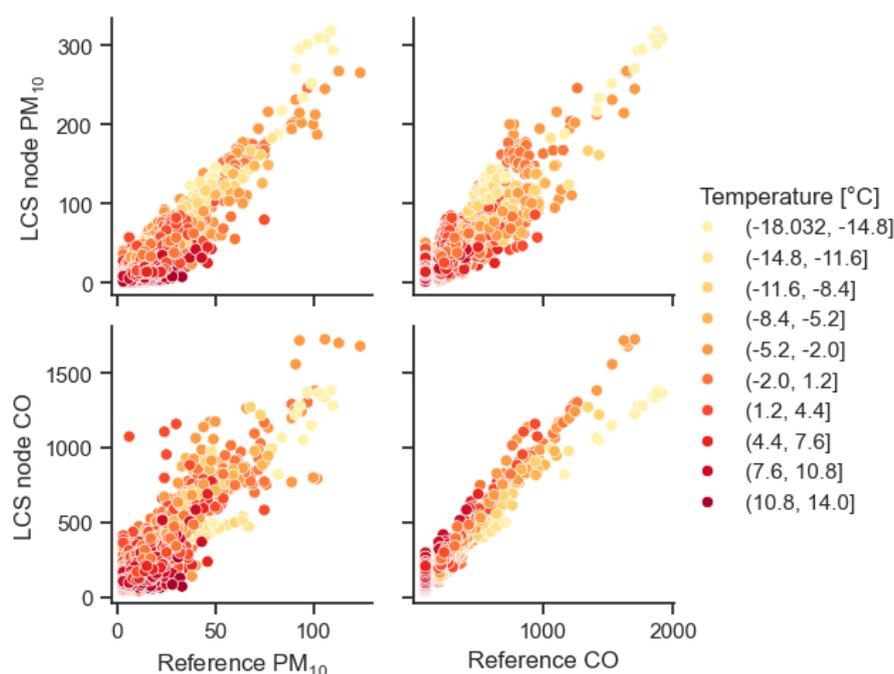
Dramatic difference in SPS30 response to coarse particles was observed during the S2 episode. Nevertheless, thanks to availability of innovated reference instruments (FIDAS) we were able to determine simultaneously SLR fit for fine particulate matter. Nearly perfect agreement of SPS30 response was found for  $PM_{10}$ , see Figure 9. It is worth to note that the concentrations of carbon monoxide during S2 episode were steadily below the limit of detection for reference instrument ( $[CO]_{REF} < 200 \mu\text{g}/\text{m}^3$ ).



**Figure 9.** Simple linear regression of Sensirion SPS30 sensor response versus reference instrument (FIDAS) during spring "Saharan dust storm" episode S2. Performance shown for  $PM_{10}$  (a) and  $PM_1$  (b). The histograms displayed adjacent to the axes illustrate the normalized frequency of the measured concentration ranges within the respective dataset.

## 4. Discussion

We first discuss our results regarding the response of the CO-B4 sensor and compare them with the observations of previous researchers. In the work of Camprodon et al. [8], a very high correlation ( $R^2 > 0.8$ ) and low error ( $RMSE < 0.1$  ppm) of CO measurements were observed during more than two months of CO-B4 sensor deployment. The sensor was found to behave linearly with respect to the CO concentrations and its decrease during the co-location period was negligible, which is quite consistent with our measurements. Our data obtained during the winter evaluation period (3 months) show slightly higher coefficient of determination  $R^2 \approx 0.9$  when re-calibration according to the Equation 1 is evaluated and compared with the reference data. Similar performance of this sensor is reported in Han et al. [11], evaluating an almost identical season with similar ranges of air pollutants but with temperatures in the range 0-20°C. Our co-location was carried out at much lower temperatures, while temperature correction following Equation 1 seems to be less effective at extremely low ambient temperatures (below -10°C) and high CO levels, see Figure 10. Conversely, slightly overestimated values of  $[CO]_{LCS}$  were observed during warmer days (with  $t > 5^\circ\text{C}$ ). In the given case, we can attribute biased  $[CO]_{LCS}$  values to direct temperature effect on sensing mechanism, i.e. reduced rate of (electro-)chemical reactions, and corresponding non-linearities.



**Figure 10.** Correlation of reference measurements and LCS response during the winter evaluation period and the effect of ambient temperature on sensor performance (shown by the colour of the data point).

On the other hand, as far as the influence of temperature on the response of the SPS30 sensor in our local conditions is concerned, we anticipate rather an indirect effect consisting in the change of particle size distribution due to the increased need for domestic and industrial heating at lower ambient temperatures.

This hypothesis is consistent with a number of previous publications, e.g. [15,16], mentioning in particular the work of Zareba et al. [16], which show a negative correlation between ambient temperature and the air pollution in an area close to our co-location site. Their study confirms that in moderate climate zones with coal burning as the primary source of air pollution, temperature is the most significant factor influencing monthly average PM<sub>10</sub> concentrations.

As in the case of our co-location site, many AQM stations in the region covered by this paper are not yet equipped with reference instruments for measuring fine PM fractions. Moreover, the air quality criteria recommended by WHO, EU, or local authorities for declaring a smog alert situation usually take into account PM<sub>10</sub> concentrations or rarely PM<sub>2.5</sub>. Therefore, our main motivation was to find a solution to reliably determine PM<sub>10</sub> values based on LCS data.

Below we briefly summarize some of the findings from previous studies on the performance of SPS30 sensors, in particular on their reliability in measuring fine and coarse PM concentration.

In a study of Roberts et al. [17] co-locating the SPS30 with regulatory methods, it achieved an average bias adjusted  $R^2 = 0.75$  for 24-hour averages and 0.57 for 1-hour averages, suggesting reasonable accuracy in real-time monitoring. The mean bias error was minimal, indicating that the SPS30 provides reliable data for PM<sub>2.5</sub> levels.

According to Kuula et al. [18] the SPS30 sensor is suitable to be used for measuring PM<sub>1</sub> particles when  $R^2 = 0.91$ , indicating high accuracy and consistency. Whereas for PM<sub>2.5</sub> particles this value was 0.83 and for PM<sub>10</sub> particles it was 0.12, which is characterized by low measurement reliability and this sensor is not suitable for larger particle sizes.

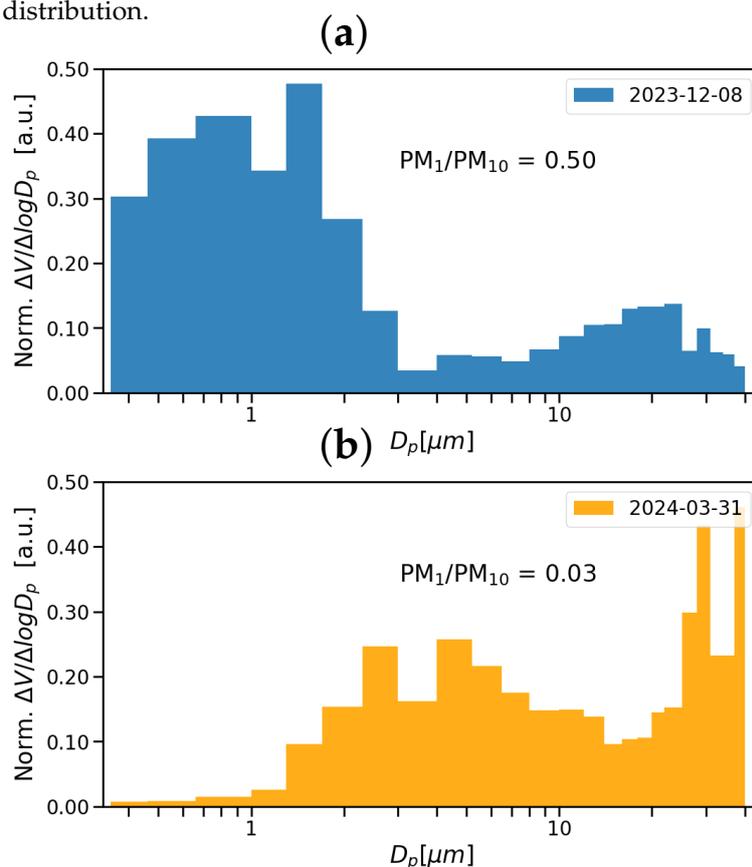
Vogt et al. [19] also confirm that the SPS30 sensor is the mostly accurate and reliable for PM<sub>1</sub> particles with  $R^2 = 0.94$ . For PM<sub>2.5</sub> particles the  $R^2$  value was around 0.73. The results for PM<sub>10</sub>

particles indicate a higher value ( $R^2 = 0.46$ ) compared to the results of Kuula et al. [18], yet still not suitable for practical AQM applications.

Molino Ruada et al. [20] confirmed the trend of the SPS30 sensor being able to measure  $PM_{10}$  particles with a high accuracy of  $R^2 = 0.93$ . As particle size increases, the accuracy of the sensor decreases yielding  $R^2 = 0.72$  for  $PM_{2.5}$  and  $R^2 = 0.23$  for  $PM_{10}$ , respectively.

The physical explanation for the unreliable measurement of larger particles is related to design of optically based LCS and principle of their operation (i.e. light-scattering). Above all, shortened viewing angles, losses occurring during particle intake and also differences in particle shape and refractive index needs to be taken into account as well as the effect of humidity and sensor aging, when these LCS are exposed to realistic outdoor conditions.

Considering these findings together with the results of our LCS node measurements against the reference data, we can conclude that the SPS30 provides a reliable response to fine dust particles, especially  $PM_{10}$ , even under Saharan dust storm conditions. The  $PM_{10}$  readings from the SPS30 sensor according to its original calibration (i.e. factory setting) are burdened with a systematic bias whose trend (negative or positive) depends on the type of smog situation. Therefore, to conclude this discussion, let us take a closer look at the size-resolved histogram of the PM volume concentration distribution obtained from the OPC-N3 sensor on days with maximum  $PM_{10}$  concentration in the case of S1 and S2 episodes, see Figure 11. The difference in particle size resolution is noticeable, with both data showing significant bimodality. In the case of the Polish smog (S1), the total volume is clearly dominated by  $PM_{10}$ . On the other hand, in the case of the Saharan dust storm (S2), particles with aerodynamic diameter  $D_p \approx 4\mu m$  have the highest volume concentration from the total  $PM_{10}$  found in the size-resolved distribution.



**Figure 11.** Size distribution of the normalized particle volume by bin of the Alphasense OPC-N3 sensor. The value of mass-weighted  $PM_{10}/PM_{10}$  ratio estimated from 24-hour average on selected days during S1 (a) and S2 (b) episodes.

In analogy to recent work by Kaur and Kelly [21], we propose a strategy to derive  $PM_{10}$  concentrations from the biased PM-LCS response based on correction factors obtained from the OPC-N3 sensor working in concert. Further assume an Equation 2 and Equation 3, which can be used to adjust the slope  $a_{SLR,S1}$  and  $a_{SLR,S2}$  respectively to ideal value  $a_{COR} \approx 1$ . Then we can use the inverse estimation in order to determine that the calibration of SPS30 is presumably carried out with the aerosol mixture having  $[PM_1/PM_{10}]_{SPS,calibr} \approx 0.2 \pm 0.05$ , which corresponds to a common (traffic-related) air pollution in urban areas.

$$\frac{[PM_1/PM_{10}]_{OPC,S1}}{a_{SPS,S1}} = \frac{0.5}{2.1} = 0.23 \approx [PM_1/PM_{10}]_{SPS,calibr} \quad (2)$$

$$\frac{[PM_1/PM_{10}]_{OPC,S2}}{a_{SPS,S2}} = \frac{0.03}{0.19} = 0.16 \approx [PM_1/PM_{10}]_{SPS,calibr} \quad (3)$$

Therefore, the  $PM_{10}$  values measured by SPS30 are systematically biased if the actual  $[PM_1/PM_{10}]$  values differs significantly from the  $[PM_1/PM_{10}]_{SPS,calibr}$ . In other words, it was proved, that the biased SPS30 reading of  $PM_{10}$  can be roughly corrected using  $[PM_1/PM_{10}]_{OPC}$  divided by a factor  $(0.2 \pm 0.05)$ . More precise corrections will only be possible after further analysis and experimentation.

#### 4.1. Practical applicability

This work represents a significant step towards strengthening the role of citizen science and democratizing environmental data in AQM, and demonstrates the importance of academic support for these efforts as the current state of knowledge and technology is still rather prohibitive to the straightforward deployment of commercially available LCS systems in their default (factory calibrated) setup. Therefore, careful evaluation of LCS performance (in the form of co-location measurements) and consideration of specific conditions of their deployment at local and regional level before their practical application is inevitable.

In the framework of this work, we have been able to explain the seasonal variability of the Sensirion SPS30 sensor response and a correction method increasing the reliability of its  $PM_{10}$  response have been established. According to our findings, we can exploit the strengths of the SPS30 sensor and overcome its previously reported limitations. Correction of biased response can be expressed based on fine-to-coarse particle ratio, e.g. as  $PM_1/PM_{10}$  evaluated from OPC-N3 sensor. It was also found that additional temperature correction needs to be estimated for CO-B4 sensor to account for biased response at extremely low temperatures.

Our future aim is to enhanced reliability of the regional AQM data when combining CAMS model predictions and LCS response by the means of machine-learning approaches employing parameterized (e.g. MLR or HDMR [22]) or non-parameterized methods [23].

#### 4.2. Limitations

This study has several limitations, mainly due to seasonal character and the influence of weather conditions relevant to the location and the winter season. It also specifically focuses only on the response of selected LCS systems integrated into a prototype node that is still under development. In our study, only individual pieces of the selected LCSs were tested and evaluated, thus not including the influence of inter-unit variability. Due to the duration of the co-location measurement, LCS aging factors were neglected.

### 5. Conclusions

Prototype LCS node was successfully developed and co-located at AQM station in Ostrava-Mariánské Hory for the period of several months. Important features of LCS response was observed during seasonal smog episodes when searching for appropriate calibration procedure and methods. It was found that the key question in this context is not how to perform sensor calibration (i.e. which mathematical method is applied) but when and where the co-location measurement is carried out (with

case-specific circumstances). Our results demonstrate the potential of LCS-based networks for reliable monitoring of major pollutants (CO and PM<sub>10</sub>) during smog situations that occur during the winter months in the Moravian-Silesian region.

**Author Contributions:** V.N.: Writing – original draft, Resources, Methodology, Investigation, Data curation, Conceptualization. M.D.: Writing – original draft, Validation, Methodology, Investigation. P.B.: Conceptualization, Investigation, Resources. V.K.: Methodology, Investigation, Data curation. J.S.: Software, Investigation. P.P.: Software, Investigation. K.N.: Investigation. M.B.: Writing – original draft, Visualization, Investigation. R.L.: Investigation. Š.B.: Investigation. B.M.: Investigation. M.V.: Software, Data curation, Investigation. A.N.: Software, Data curation, Investigation. M.L.: Data curation, Investigation. J.Su.: Project administration, Resources, Investigation. H.C.: Investigation. D.K.: Data curation, Investigation. J.W.: Supervision, Funding acquisition.

**Funding:** This research was funded from the project no. SS03010139 of Technology Agency of the Czech Republic and the project No. 22-19812S funded by the Czech Science Foundation

**Data Availability Statement:** The raw datasets and Python interactive notebooks used for preparation of this paper are freely available in the [Zenodo repository](https://zenodo.org/doi/10.5281/zenodo.13863368) (doi.org/10.5281/zenodo.13863368).

**Acknowledgments:** We gratefully acknowledge the institutional support of the VSB-Technical University of Ostrava (Faculty of Safety Engineering) via student project No. SP2024/044.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

PM <sub>1</sub>	Particulate Matter with diameter ≤ 1 micrometer
PM <sub>2.5</sub>	Particulate Matter with diameter ≤ 2.5 micrometers
PM <sub>10</sub>	Particulate Matter with diameter ≤ 10 micrometers
CO	Carbon Monoxide
CAMS	Copernicus Atmospheric Monitoring Service
VOCs	Volatile Organic Compounds
O <sub>3</sub>	Ozone
LCS	Low-Cost Sensors
AQM	Air Quality Monitoring
EU	European Union
RMSE	Root Mean Square Error
RH	Relative Humidity
NDIR	Nondispersive Infrared
TEOM	Tapered Element Oscillating Microbalance
LoD	Limit of Detection
CSV	Comma-Separated Value
MLR	Multilinear Regression
MAE	Mean Average Error
GMT	Greenwich Mean Time
SLR	Simple Linear Regression
atmospy	Python library for atmospheric data analysis
smps-py	Python library for particle size distribution analysis
NO <sub>2</sub>	Nitrogen Dioxide
UV	Ultraviolet Radiation
r	Pearson Correlation Coefficient
[CO] <sub>LCS</sub>	CO concentration measured by LCS
[CO] <sub>REF</sub>	Reference CO concentration
[CO] <sub>CAMS</sub>	CO Concentration from CAMS model
[PM <sub>10</sub> ] <sub>REF</sub>	Reference PM <sub>10</sub> concentration
S1	Smog Episode 1
S2	Smog Episode 2
R <sup>2</sup>	Coefficient of Determination

ANN	Artificial Neural Network
HDMR	High-Dimensional Model Representation
LoRaWAN	Long Range Wide Area Network
ASA	Acrylonitrile Styrene Acrylate (3D printing material)
MQTT	Message Queuing Telemetry Transport

## Appendix A. LCS Node Design and Data Management

### Appendix A.1. Hardware Description

The LCS sensor node is designed as a modular system with three interconnected printed circuit boards (PCBs), each serving a specific function: control and communication, sensor interface, and power management.

The Control and Communication Board houses the LilyGo TTGO LoRa32 T3 v1.6 with 868MHz microcontroller module, featuring the ESP32. This board manages data collection and communication, utilizing Wi-Fi (employed within this work) and LoRaWAN (optional). It operates primarily at 3.3V, with voltage regulators and DC-DC converters ensuring stable power supply for consistent performance.

The Power Management Board distributes power across the system and supports multiple power input options, including DC from the grid (employed within this work), lithium-ion batteries, and 6V lead-acid batteries (optional). This board includes DC-DC converters and voltage regulators to step down input voltages to the necessary levels. It also integrates a slot for NEO-6M GPS module for geolocation, providing real-time position data alongside environmental measurements.

The Sensor Interface Board integrates individual sensors, including particulate matter (Alphasense OPC-N3 and Sensirion SPS30). It also features an Analog-to-Digital Converter ADS1115 to process signals from analog Alphasense type B sensors. Stable voltage of 3.5V is provided for Individual Sensor Boards as required for measurement accuracy.

The interconnection of these PCBs ensures seamless communication and power distribution, contributing to the mechanical stability of the sensor node and simplifying assembly and future modifications.

Additionally, a custom-designed enclosure for the sensor node was developed and manufactured using MDF 3D printing technology. The material chosen for the enclosure is ASA (Acrylonitrile Styrene Acrylate), which offers excellent weather resistance and durability. The enclosure not only protects the sensors from external elements but also ensures proper airflow for accurate measurements. A 3D model of the sensor node's enclosure is shown in Figure A1.

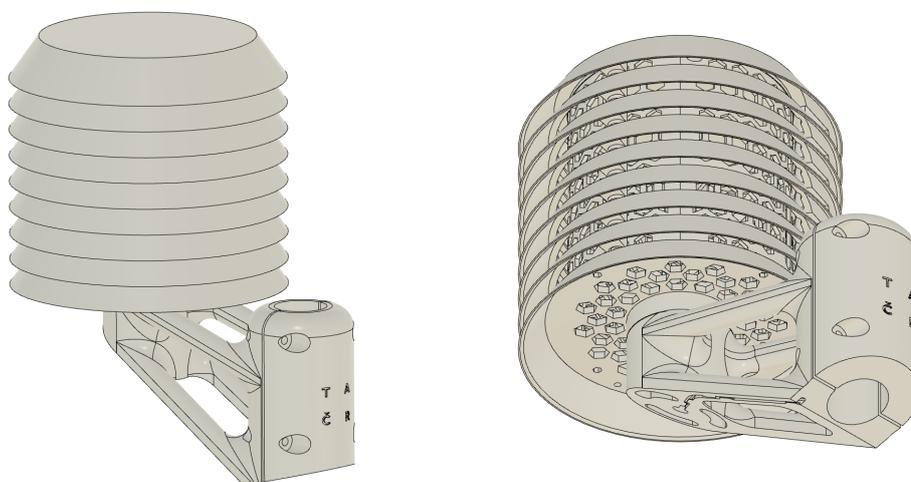


Figure A1. Rendered 3D model of the LCS node enclosure.

## Appendix A.2. Datalogging

Each sensor node is equipped with an ESP32 microcontroller, which gathers data at regular intervals. This data is transmitted using Wi-Fi via the MQTT protocol, specifically through the Mosquitto MQTT broker, which publishes the data to designated topics. Telegraf, configured to listen to these MQTT topics, subscribes to the relevant data streams and stores the incoming information in an InfluxDB database. To make this data accessible and interpretable, Grafana connects to the InfluxDB database and visualizes the data on customizable dashboards. Selected datasets from co-location measurement are resampled with 10-minutes temporal resolution and automatically transferred to CSV format in monthly intervals. These data can be accessed from publicly available web interface [SOASENSE](#).

## References

1. Svozilík, V.; Svozilíková Krakovská, A.; Bitta, J.; Jančík, P. Comparison of the Air Pollution Mathematical Model of PM<sub>10</sub> and Moss Biomonitoring Results in the Triaia Region. *Atmosphere* **2021**, *12*, 656. Number: 6 Publisher: Multidisciplinary Digital Publishing Institute, doi:10.3390/atmos12060656.
2. Volná, V.; Hladký, D.; Seibert, R.; Krejčí, B. Transboundary Air Pollution Transport of PM<sub>10</sub> and Benzo[a]pyrene in the Czech–Polish Border Region. *Atmosphere* **2022**, *13*, 341. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute, doi:10.3390/atmos13020341.
3. Volná, V.; Seibert, R.; Hladký, D.; Krejčí, B. Identification of Causes of Air Pollution in a Specific Industrial Part of the Czech City of Ostrava in Central Europe. *Atmosphere* **2024**, *15*, 177. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute, doi:10.3390/atmos15020177.
4. Wielgosiński, G.; Czerwińska, J. Smog Episodes in Poland. *Atmosphere* **2020**, *11*, 277. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute, doi:10.3390/atmos11030277.
5. Castell, N.; Dauge, F.R.; Schneider, P.; Vogt, M.; Lerner, U.; Fishbain, B.; Broday, D.; Bartonova, A. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? *Environment International* **2017**, *99*, 293–302. doi:10.1016/j.envint.2016.12.007.
6. Gabrys, J. Planetary health in practice: sensing air pollution and transforming urban environments. *Humanities and Social Sciences Communications* **2020**, *7*, 1–11. Publisher: Palgrave, doi:10.1057/s41599-020-00534-7.
7. Mahajan, S.; Chung, M.K.; Martinez, J.; Olaya, Y.; Helbing, D.; Chen, L.J. Translating citizen-generated air quality data into evidence for shaping policy. *Humanities and Social Sciences Communications* **2022**, *9*, 1–18. Publisher: Palgrave, doi:10.1057/s41599-022-01135-2.
8. Camprodon, G.; González, Ó.; Barberán, V.; Pérez, M.; Smári, V.; de Heras, M.Á.; Bizzotto, A. Smart Citizen Kit and Station: An open environmental monitoring system for citizen participation and scientific experimentation. *HardwareX* **2019**, *6*, e00070. doi:10.1016/j.ohx.2019.e00070.
9. Giordano, M.R.; Malings, C.; Pandis, S.N.; Presto, A.A.; McNeill, V.; Westervelt, D.M.; Beekmann, M.; 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* **2021**, *158*, 105833. doi:10.1016/j.jaerosci.2021.105833.
10. Cross, E.S.; Williams, L.R.; Lewis, D.K.; Magoon, G.R.; Onasch, T.B.; Kaminsky, M.L.; Worsnop, D.R.; Jayne, J.T. Use of electrochemical sensors for measurement of air pollution: correcting interference response and validating measurements. *Atmospheric Measurement Techniques* **2017**, *10*, 3575–3588. doi:10.5194/amt-10-3575-2017.
11. Han, P.; Mei, H.; Liu, D.; Zeng, N.; Tang, X.; Wang, Y.; Pan, Y. Calibrations of Low-Cost Air Pollution Monitoring Sensors for CO, NO<sub>2</sub>, O<sub>3</sub>, and SO<sub>2</sub> **2021**.
12. Hagan, D.H.; Isaacman-VanWertz, G.; Franklin, J.P.; Wallace, L.M.M.; Kocar, B.D.; Heald, C.L.; Kroll, J.H. Calibration and assessment of electrochemical air quality sensors by co-location with regulatory-grade instruments. *Atmospheric Measurement Techniques* **2018**, *11*, 315–328. Publisher: Copernicus GmbH, doi:10.5194/amt-11-315-2018.
13. Liang, L.; Daniels, J. What Influences Low-cost Sensor Data Calibration? - A Systematic Assessment of Algorithms, Duration, and Predictor Selection. *Aerosol and Air Quality Research* **2022**, *22*, 220076. Publisher: Taiwan Association for Aerosol Research, doi:10.4209/aaqr.220076.

14. METEO FRANCE.; INSTITUT NATIONAL DE L'ENVIRONNEMENT INDUSTRIEL ET DES RISQUES (INERIS).; AARHUS UNIVERSITY.; NORWEGIAN METEOROLOGICAL INSTITUTE (MET NORWAY).; JÜLICH INSTITUT FÜR ENERGIE- UND KLIMAFORSCHUNG (IEK).; INSTITUTE OF ENVIRONMENTAL PROTECTION – NATIONAL RESEARCH INSTITUTE (IEP-NRI).; KONINKLIJK NEDERLANDS METEOROLOGISCH INSTITUUT (KNMI).; NEDERLANDSE ORGANISATIE VOOR TOEGEPAST-NATUURWETENSCHAPPELIJK ONDERZOEK (TNO).; SWEDISH METEOROLOGICAL AND HYDROLOGICAL INSTITUTE (SMHI).; FINNISH METEOROLOGICAL INSTITUTE (FMI).; ITALIAN NATIONAL AGENCY FOR NEW TECHNOLOGIES, ENERGY AND SUSTAINABLE ECONOMIC DEVELOPMENT (ENEA).; BARCELONA SUPERCOMPUTING CENTER (BSC). CAMS European air quality forecasts, ENSEMBLE data. <https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-europe-air-quality-forecasts?tab=overview>, 2022. Accessed on <01-08-2024>.
15. Zhang, X.; Zhou, B.; Li, Z.; Lin, Y.; Li, L.; Han, Y. Seasonal Distribution of Atmospheric Coarse and Fine Particulate Matter in a Medium-Sized City of Northern China. *Toxics* **2022**, *10*, 216. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute, doi:10.3390/toxics10050216.
16. Zareba, M.; Weglinska, E.; Danek, T. Air pollution seasons in urban moderate climate areas through big data analytics. *Scientific Reports* **2024**, *14*, 3058. Publisher: Nature Publishing Group, doi:10.1038/s41598-024-52733-w.
17. Roberts, F.; Van Valkinburgh, K.; Green, A.; Post, C.J.; Mikhailova, E.A.; Commodore, S.; Pearce, J.L.; Metcalf, A.R. Evaluation of a new low-cost particle sensor as an internet-of-things device for outdoor air quality monitoring. *Journal of the Air & Waste Management Association* **2022**, *72*, 1219–1230. Publisher: Taylor & Francis \_eprint: <https://doi.org/10.1080/10962247.2022.2093293>, doi:10.1080/10962247.2022.2093293.
18. Kuula, J.; Mäkelä, T.; Aurela, M.; Teinilä, K.; Varjonen, S.; González, Ó.; Timonen, H. Laboratory evaluation of particle-size selectivity of optical low-cost particulate matter sensors. *Atmospheric Measurement Techniques* **2020**, *13*, 2413–2423. Publisher: Copernicus GmbH, doi:10.5194/amt-13-2413-2020.
19. Vogt, M.; Schneider, P.; Castell, N.; Hamer, P. Assessment of Low-Cost Particulate Matter Sensor Systems against Optical and Gravimetric Methods in a Field Co-Location in Norway. *Atmosphere* **2021**, *12*, 961. doi:10.3390/atmos12080961.
20. Molina Rueda, E.; Carter, E.; L'Orange, C.; Quinn, C.; Volckens, J. Size-Resolved Field Performance of Low-Cost Sensors for Particulate Matter Air Pollution. *Environmental Science & Technology Letters* **2023**, *10*, 247–253. Publisher: American Chemical Society, doi:10.1021/acs.estlett.3c00030.
21. Kaur, K.; Kelly, K.E. Performance evaluation of the Alphasense OPC-N3 and Plantower PMS5003 sensor in measuring dust events in the Salt Lake Valley, Utah. *Atmospheric Measurement Techniques* **2023**, *16*, 2455–2470. Publisher: Copernicus GmbH, doi:10.5194/amt-16-2455-2023.
22. Bittner, A.S.; Cross, E.S.; Hagan, D.H.; Malings, C.; Lipsky, E.; Grieshop, A. Performance Characterization of Low-cost Air Quality Sensors for Off-grid Deployment in Rural Malawi **2021**. doi:10.5194/amt-2021-372.
23. Topalović, D.B.; Davidović, M.D.; Jovanović, M.; Bartonova, A.; Ristovski, Z.; Jovašević-Stojanović, M. In search of an optimal in-field calibration method of low-cost gas sensors for ambient air pollutants: Comparison of linear, multilinear and artificial neural network approaches. *Atmospheric Environment* **2019**, *213*, 640–658. doi:10.1016/j.atmosenv.2019.06.028.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.