Comparative Evaluation of New Low-Cost Particulate Matter Sensors

by

Ishan Patwardhan, Spanddhana Sara, Sachin Chaudhari

Report No: IIIT/TR/2021/-1

Centre for Communications
International Institute of Information Technology
Hyderabad - 500 032, INDIA
August 2021

Comparative Evaluation of New Low-Cost Particulate Matter Sensors

Ishan Patwardhan, Spanddhana Sara, Sachin Chaudhari
Signal Processing and Communication Research Centre,
International Institute of Information Technology Hyderabad, India
ishan.patwardhan@research.iiit.ac.in, spanddhana.sara@research.iiit.ac.in, sachin.chaudhari@iiit.ac.in

***Abstract*—In recent times, a few new low-cost sensors have been introduced to the global market for monitoring particulate matter (PM). In this paper, the performance of three such low-cost PM sensors, namely SDS011, Prana Air, and SPS30, for measuring $PM_{2.5}$ and PM_{10} levels is evaluated against a standard reference Aeroqual Series-500. The test setup was exposed to PM concentrations ranging from 30 μg/cm^3 to 600 μg/cm^3. The results were based on 1 min, 15 min, 30 min, and 1 hr average readings. The experiments were carried out in indoor as well as outdoor environments. The comparative evaluation was performed before and after calibration. The performance of these sensors is evaluated in terms of coefficient of determination (R^2), coefficient of variation (C_v) and root mean square error (RMSE). Evaluation results show that these low-cost sensors have good performance after calibration with a reference sensor.**

***Index Terms*—Coefficient of determination, Coefficient of variability, Low-cost PM sensor, performance evaluation**

I. Introduction

Particulate matter (PM) refers to the small particle impurity suspended in the air. It is one of the major contributors to air pollution. The high-grade PM monitoring devices are usually bulky and expensive, such as instruments based on Tapered Element Oscillating Microbalance (TEOM) technology and Beta Attenuation Monitor (BAM) [1]. They also require frequent servicing to get the best performance. For the above-mentioned reasons, only a few PM-level monitoring systems are deployed in a city.

Many low-cost sensors are now available in the global market for monitoring $PM_{2.5}$ and PM_{10} levels and are mainly used in smart city projects [2]. These sensors are compact, easy to handle, and require very low maintenance. The deployment of these sensors can be in huge numbers in a small area, increasing the spatial and temporal resolution of the PM data. In [3, 4], a few approaches have been discussed to improve the spatio-temporal resolution of PM data using distributed networks of such low-cost sensors. Due to the mass availability and production of low-cost sensors, it is vital to examine their credibility.

Some experiments have already been performed in the past for the evaluation of low-cost PM sensors using the data from various weather stations in and around a particular city as a reference [5, 6]. A study [7] shows that the response of these sensors depend not only upon the particle dimension but also on the type of particle. This means that the sensors manufactured by the different companies can have variations in their measurements of PM levels. In [5], SDS011 (Nova Fitness), ZH03A (Winsen), PMS7003 (Plantower), and OPC-N2 (Alphasense) were used as the test sensors, and a TEOM based reference instrument was utilized. This experiment was carried out by using a stationary test setup near the weather station. In [6], Plantower PMS 1003/3003 sensor was examined in a wind tunnel and outdoor environment. Unlike the existing literature, which mostly used a stationary setup, we have used a portable setup to test the PM sensors in few different parts of the city. This way sensors are exposed to different types of particles from sources like factories, vehicles, restaurants by taking the test setup and the reference instrument near such places. In addition, this paper evaluates two new sensors (Sensirion and Prana) along with one well-studied sensor (SDS011).

The main contributions of this paper are

- The performance of three low-cost sensors, namely SDS011 (Nova Fitness) [8], Prana Air (Prana) [9], SPS30 (Sensirion) [10], is evaluated in this study.
- A mobile test setup is used to study the test sensors' behavior when exposed to different kinds of particles such as dust from the construction sites, traffic .
- The measurements and data analysis are performed separately for indoor and outdoor environments.
- The performance evaluation of these sensors is carried out in terms of coefficient of determination (R^2), coefficient of variation (C_v) and root mean square error (RMSE).

The structure of this paper is as follows. Section II briefly describes the hardware and software setup. Section III describes the indoor and outdoor measurement campaigns. Section IV presents the evaluation parameters used to compare the test sensors using the data collected during the experiments. Results and corresponding discussions are presented in Section V while Section VI concludes this paper.

II. Test Setup

This section contains information about the equipment used for the experiment and the technical parameters used to collect and analyze the data points.

A. Reference Instrument

The reference instrument that was used for this experiment is Series-500 manufactured by Aeroqual [11]. It is a portable air pollution monitoring device that can measure $PM_{2.5}$ and PM_{10} levels simultaneously with a minimum time resolution of 1 min. It works on laser particle counter (LPC) technology and is factory calibrated. A case study [12] reports a very high correlation between this portable monitor and higher specification environmental monitors.

B. Nodes

Three identical test nodes were created for the experiment. Figs. 1 and 2 show the schematic view and actual view, respectively, for each such node, which consists of one unit of SDS011, Prana Air, and SPS30 each. ESP8266 based Wi-Fi enabled NodeMCU v1.0 microcontroller module was used to interface these sensors.

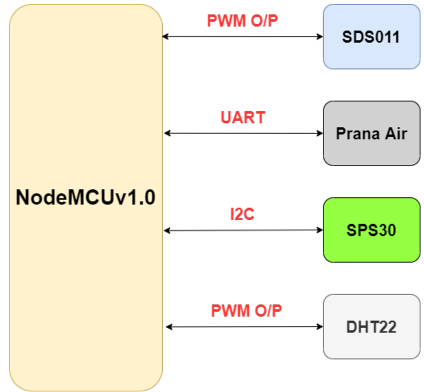


Fig. 1: Schematic view of a node.

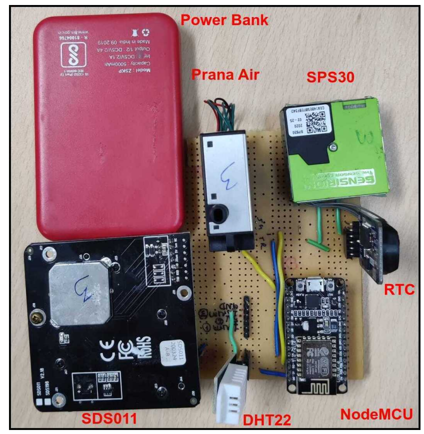


Fig. 2: Actual view of a node.

C. Data Collection

Samples were collected at 2 s intervals, and all data was pushed to Thingspeak, an MQTT-based IoT platform. The Wi-Fi access point was created using a smartphone, and a 4G cellular network was employed to access the internet. All information was downloaded in .CSV format from the Thingspeak platform for all three nodes. We also downloaded the data from our standard reference Series-500. The data was processed and analyzed using the python programming language for 1 min, 15 min, 30 min, and 1 hr averaged readings.

III. Measurement Campaigns

In this paper, the performance evaluation was done in indoor and outdoor environments. The following two subsections briefly describe the process of both experiments.

A. Indoor Experiment

In this method, the test-setup was stationary and placed near an open window to record the PM concentration for one week continuously. The windows of the room remained open for the whole duration of this experiment to keep good ventilation.

B. Outdoor Experiment

In the outdoor experiment, all three nodes and the reference instrument Series-500 were placed inside a vehicle with open windows. All devices were powered using a 5000 mAh power bank. The vehicle was taken to several parts of the city to cover all kinds of areas, industrial, residential and commercial, and places with high and low traffic movement. A few halts for short intervals of 30-45 minutes were made to collect the readings of that particular area. The main objective was to evaluate the ability of sensors to record the change in PM levels. Fig. 3 displays the $PM_{2.5}$ levels at the locations chosen for this experiment.

IV. Data Pre-processing and Performance Evaluation Parameters

The data obtained from the test setup was filtered for outliers before performing any analysis. All sensors work in a specific humidity range, so all the readings that were obtained beyond the operating humidity range were removed.

The performance parameters used in this paper are R^2, C_v, and RMSE. The coefficient of determination R^2 analyzes the ability to report the changes in PM levels in

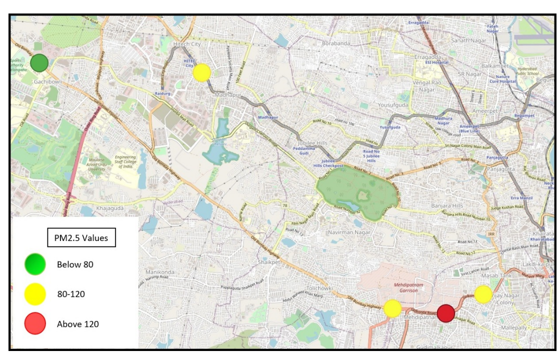


Fig. 3: Measurement of outdoor $PM_{2.5}$ levels.

comparison to the reference instrument. It is given in [13] by

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}, \quad (1)$$

where y_i denotes the observations, $\bar{y}$ denotes the average of the observations and $\hat{y}_i$ is the prediction of y_i using the fitted model. The R^2 values were calculated with raw sensor output values as well as calibrated values. The calibration of the test sensors with respect to the reference sensor is done using simple linear regression.

Coefficient of variability C_v measures the reproducibility or the variance across multiple units of the same sensor and is given by

$$C_v = \frac{1}{N} \sum_{i=1}^{N} \frac{\hat{\sigma}_i}{\hat{\mu}_i} \times 100\%, \quad (2)$$

where N is the total number of samples, $\hat{\mu}_i$ denotes the sample mean of readings of all three units of a particular sensor at one moment of time and $\hat{\sigma}_i$ is the average standard deviation of all the copies of a sensor [14]. C_v was calculated only on the raw data in this paper.

RMSE denoted by E_{rms} is a metric representing the average of the square root of the sum of squares of errors and is given by.

$$E_{rms} = \frac{1}{N} \sqrt{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}, \quad (3)$$

where $\hat{y}_i$ is the prediction of y_i using the fitted model. RMSE was only calculated for the calibrated data in this paper.

V. Results and Discussion

A. Indoor experiment

Based on approximately 200,000 data points, the performance of the sensors is analyzed. All the sensors and reference instruments data in the indoor setup plotted against time are showed in Figs. 4 and 5 for $PM_{2.5}$ and PM_{10}, respectively. The graphs in Figs. 4 and 5 show that the sensors' data and reference instruments' data follow a similar trend. The sensors underestimate the PM values most of the time when compared to Aeroqual. It can also be observed that the bias between the sensors and the reference instrument is lower at lower PM values and higher at higher PM values.

Tables I and II contain the average R^2 values of all three copies of the same sensor for $PM_{2.5}$ and PM_{10} before performing the calibration. A very high correlation value was observed for all the sensors in this setup. All sensors respond equally well while capturing the changes in $PM_{2.5}$ levels. A minor drop in R^2 values is observed in the estimation of PM_{10} levels for all the sensors.

Table III indicates the C_v values for estimation of $PM_{2.5}$ and PM_{10}. It can be observed that this parameter is not more than 10% for any sensor. SDS011 is found to have the most variation of approximately 9% for both kinds of particles, followed by Prana Air (≈ 5-7%), and SPS30 (≈ 3%).

TABLE I: R^2 values for $PM_{2.5}$ for indoor experiment

Averaging Interval	**Sensor Name**		
	SDS011	***Prana Air***	***SPS30***
1 min	0.99	0.99	0.98
15 min	0.99	0.98	0.98
30 min	0.98	0.98	0.97
1 hr	0.97	0.97	0.96

TABLE II: R^2 values for PM_{10} for indoor experiment

Averaging Interval	**Sensor Name**		
	SDS011	***Prana Air***	***SPS30***
1 min	0.87	0.98	0.98
15 min	0.98	0.90	0.90
30 min	0.98	0.90	0.90
1 hr	0.97	0.89	0.90

TABLE III: C_v values for $PM_{2.5}$ and PM_{10} in % for indoor experiment

Particle Dimension	**Sensor Name**		
	SDS011	***Prana Air***	***SPS30***
$PM_{2.5}$	9.6	5.73	2.75
PM_{10}	9.05	6.64	2.95

B. Outdoor Experiment

The results obtained in this section are based on an analysis of approximately 30,000 data points, collected at 2 s intervals, before calibration. Figs. 6 and 7 show the plots for PM2.5 and PM10, respectively. The graphs show that the bias between the sensors and reference instrument is high, and sensors sometimes do not follow the same trend. This might be due to the sudden increase in the PM concentration and the change in the type of particle.

Tables IV and V contain the R^2 values for $PM_{2.5}$ and PM_{10} respectively. The R^2 and C_v values for mobile setup were not calculated for 30 min and 1 hr intervals due to insufficient data points. From these tables, the overall response is observed to be less accurate as compared to the results of the indoor experiment. It is observed that SDS011 has the highest R^2 value for both kinds of particles. Prana Air and SPS30 have an almost similar response for $PM_{2.5}$ and PM_{10}.

Table VI shows the C_v values obtained from the outdoor experiment. There is a minor change in C_v values for all sensors except SDS011 when compared to the indoor

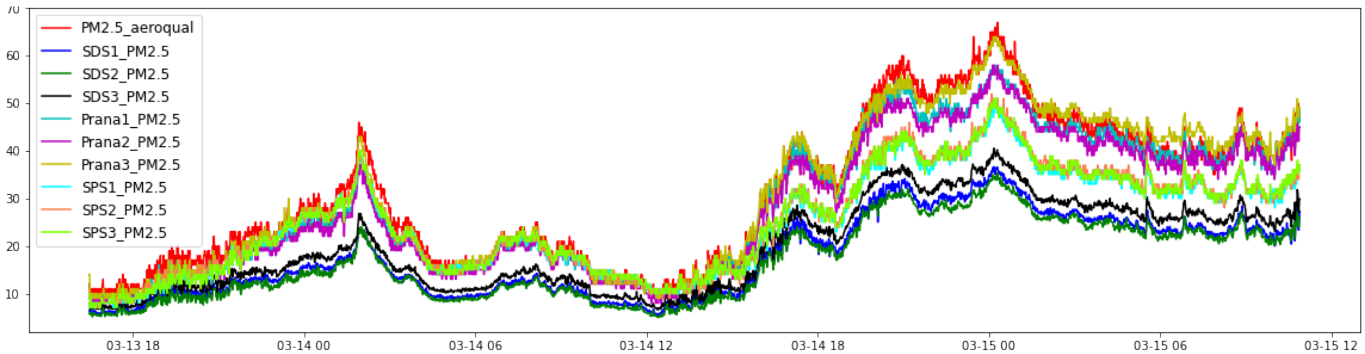


Fig. 4: $PM_{2.5}$ trend for indoor experiment.

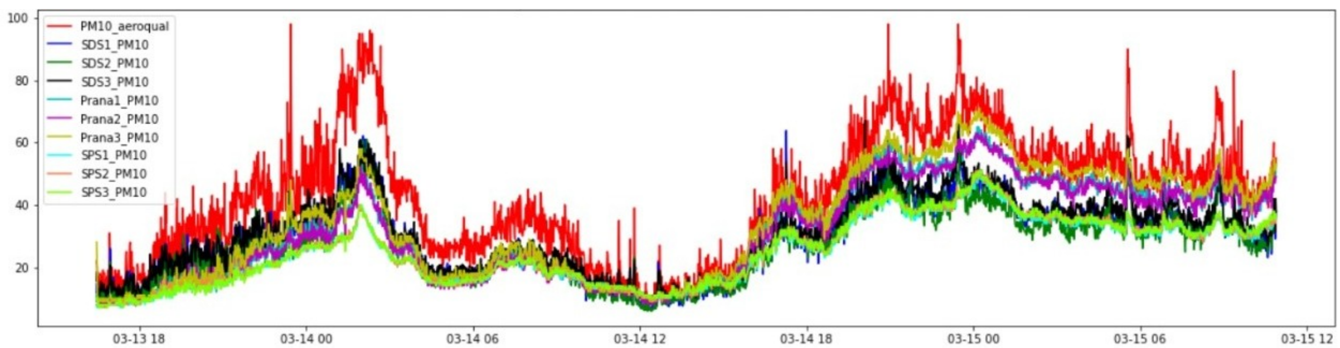


Fig. 5: PM_{10} trend for indoor experiment.

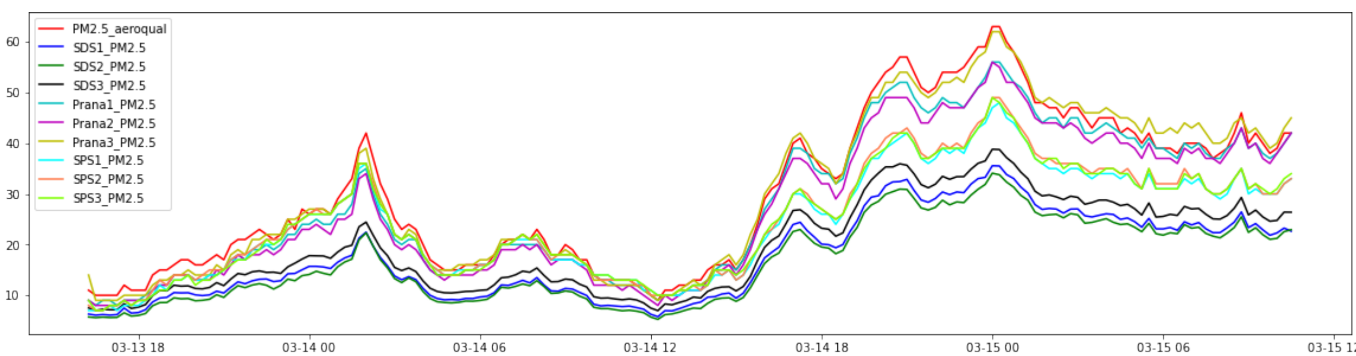


Fig. 6: $PM_{2.5}$ trend for outdoor experiment.

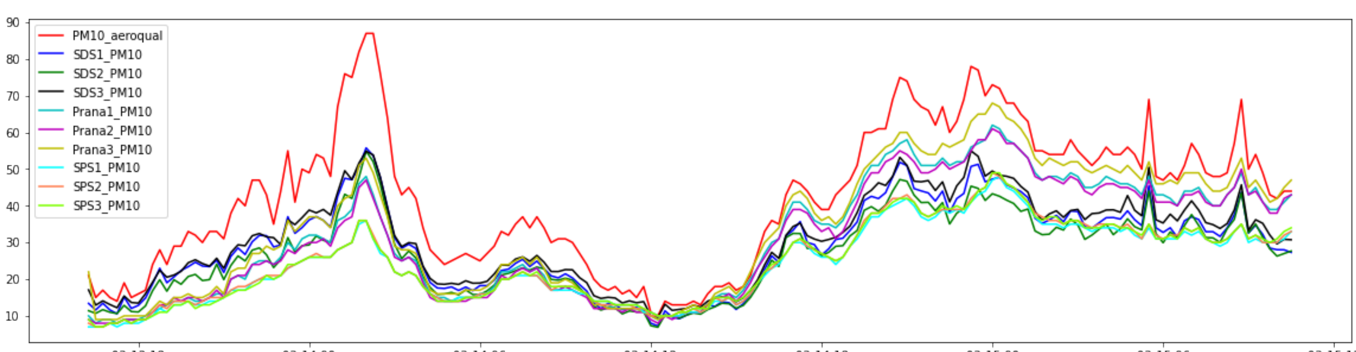


Fig. 7: PM_{10} trend for outdoor experiment.

experiment. The C_v values were observed to be in the range of 3-9% approximately for $PM_{2.5}$ and 3-20% for PM_{10}. It is seen that the SDS011 has the highest C_v value for this experiment.

TABLE IV: R^2 values for $PM_{2.5}$ for outdoor experiment

Averaging Interval	**Sensor Name**		
	SDS011	***Prana Air***	***SPS30***
1 min	0.89	0.60	0.70
15 min	0.89	0.74	0.71

TABLE V: R^2 values for PM_{10} for outdoor experiment

Averaging Interval	**Sensor Name**		
	SDS011	***Prana Air***	***SPS30***
1 min	0.73	0.61	0.70
15 min	0.87	0.68	0.65

TABLE VI: C_v values for $PM_{2.5}$ and PM_{10} in % for outdoor experiment

Particle Dimension	**Sensor Name**		
	SDS011	***Prana Air***	***SPS30***
$PM_{2.5}$	9.29	6.86	3.88
PM_{10}	20.13	7.89	3.38

C. Calibrated values

This section contains results obtained after performing the calibration for the indoor experiment only. Data points obtained from the outdoor experiment are in fewer numbers, so they have been excluded from this section. The R^2 values for 1 hr averaged samples, after performing calibration, were calculated as 0.92, 0.91, 0.86 for SDS011, Prana Air, and SPS30, respectively. Table VII indicates the E_{rms} values for 1 hr averaged samples for all test sensors. It was also observed to be reasonable for the collected data.

Figs. 8-10 present the scatter plots of all units represented by different colors before and after performing the calibration. These figures further explain the E_{rms} values. For the sensors having the least E_{rms} values, Prana Air for $PM_{2.5}$ and SDS011 for PM_{10}, the data points align very well around the average value of the reference instrument.

TABLE VII: E_{rms} values for 1 hr averaged readings

Particle Dimension	**Sensor Name**		
	SDS011	***Prana Air***	***SPS30***
$PM_{2.5}$	3.40	1.80	2.63
PM_{10}	2.42	8.3	8.8

VI. Conclusion

In this study, we found that the low-cost PM sensors were able to follow the trend of the reference instrument for most of the time with reasonably correlated values. The sensors underestimated the PM values, which can be corrected to some extent by performing calibration. We were able to achieve low E_{rms} values after doing the calibration using a simple linear regression for the indoor experiment. Also, different copies of the same sensor output the PM values in a very close range, indicating a low inter-unit variability for the sensors examined in this paper. In general, it can be concluded that the new low-cost PM sensors can be used for measuring the PM levels, but calibration is required to get a better output.

Acknowledgment

Ishan Patwardhan thanks IHub-Data situated in IIITH, India, for a Research Fellowship. This research was supported partly by National Geospatial Programme (NGP), India, under grant no. 2073 (2020), PRIF Social Incubator Program (2019), and the Ministry of Electronics and Information Technology (MEITY) under grant no. 3070665 (2020), with no conflict of interests.

References

[1] J. Schwab et al. "New York State Urban and Rural Measurements of Continuous PM2.5 Mass by FDMS, TEOM, and BAM". In: *J. of Air & Waste Manage. Assoc.* 56.4 (2006), pp. 372–383.

[2] M. Wong et al. "Towards a Smart City: Development and Application of an Improved Integrated Environmental Monitoring System". In: *Sustainability* 10.3 (2018).

[3] M. Budde, L. Zhang, and M. Beigl. "Distributed, low-cost particulate matter sensing: scenarios, challenges, approaches". In: *Proc. Int. Conf. Atmospheric Dust.* 2014, pp. 230–236.

[4] C. Rajashekar et al. "Improving Spatio-Temporal Understanding of Particulate Matter using Low-Cost IoT Sensors". In: *Int. Symp. Personal, Indoor and Mobile Radio Commun.* 2020, pp. 1–7.

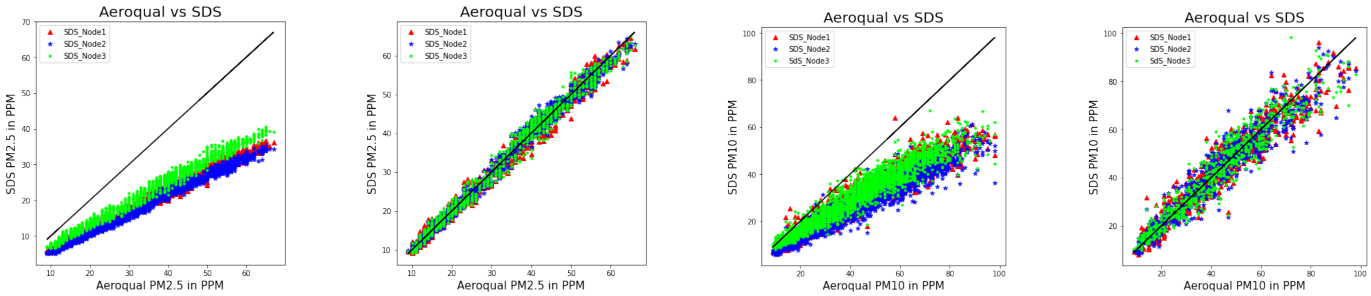
[5] M. Badura et al. "Evaluation of low-cost sensors for ambient PM2. 5 monitoring". In: *J. Sensors* (2018).

[6] K. Kelly et al. "Ambient and laboratory evaluation of a low-cost particulate matter sensor". In: *Environmental Pollution* 221 (2017), pp. 491–500.

[7] A. Manikonda et al. "Laboratory assessment of low-cost PM monitors". In: *J. Aerosol Sci.* 102 (2016), pp. 29–40.

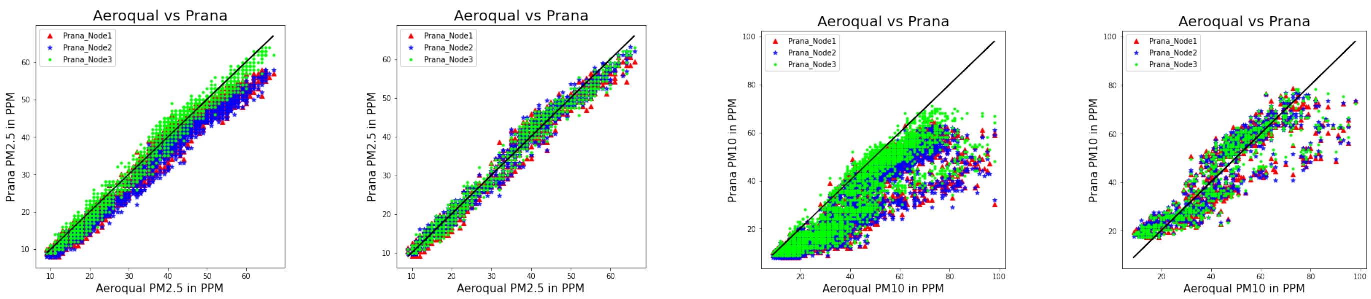
[8] *Nova Fitness SDS011.* https://cdn-reichelt.de/documents/datenblatt/X200/SDS011-DATASHEET.pdf. Oct. 2015.

[9] *Prana Air.* https://www.pranaair.com/air-quality-sensor/outdoor-pm-sensor/.



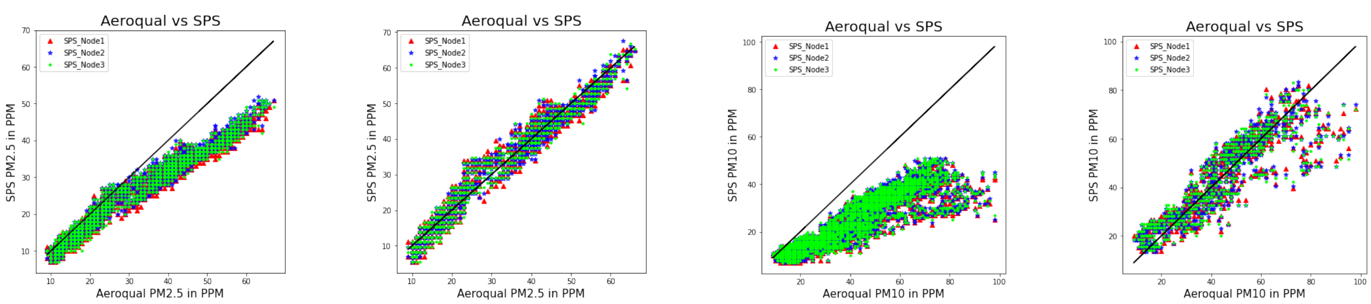
(a) $PM_{2.5}$ before calibration **(b)** $PM_{2.5}$ after calibration **(c)** PM_{10} before calibration **(d)** PM_{10} after calibration

Fig. 8: SDS011 scatter plots



(a) $PM_{2.5}$ before calibration **(b)** $PM_{2.5}$ after calibration **(c)** PM_{10} before calibration **(d)** PM_{10} after calibration

Fig. 9: Prana Air scatter plots



(a) $PM_{2.5}$ before calibration **(b)** $PM_{2.5}$ after calibration **(c)** PM_{10} before calibration **(d)** PM_{10} after calibration

Fig. 10: SPS30 scatter plots

[10] *Sensirion*. https://www.sensirion.com/en/download-center/particulate-matter-sensors-pm/particulate-matter-sensor-sps30/. (Accessed on 03/19/2021). Mar. 2020.

[11] *Series- 500 Brochure*. https://www.aeroqual.com/wp-content/uploads/Aeroqual-Portable-Monitors-Long-Brochure-LR.pdf. (Accessed on 03/19/2021).

[12] *Portable PM Colocation Study*. https://www.aeroqual.com/wp-content/uploads/Portable-PM-Colocation-Study.pdf. (Accessed on 03/19/2021).

[13] A. Bucchianico. "Coefficient of Determination". In: *Encyclopedia of Statistics in Quality and Reliability* (Mar. 2008).

[14] C. Brown. *Applied Multivariate Statistics in Geohydrology and Related Sciences*. Springer, Berlin, Heidelberg, 1998.