engineering proceedings

Proceedings

Air quality monitoring in a near-city industrial zone by low-cost sensor technologies: a case study

Domenico Suriano [1,*], Mario Prato [1] and Michele Penza [1]

[1] ENEA-Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Laboratory Functional Materials and Technologies for Sustainable Applications, Brindisi Research Centre, Strada Statale 7, Appia, Km. 706, I-72100 Brindisi, Italy; mario.prato@enea.it, michele.penza@enea.it

* Correspondence: domenico.suriano@enea.it; Tel.: +39 0831 201490

Abstract: Urban industrial areas are often a matter of concern due to the emissions of air pollutants that may affect the air quality of the adjacent cities. The aerosol pollutants are monitored by governmental agencies that employ regulatory monitoring stations which are very accurate, but also very expensive, bulky, and maintenance demanding. For this reason, it often happens that the monitoring of the air quality in large areas are covered by few stations. This situation can lead to the building of air pollutant maps having a low spatio-temporal resolution. An appealing way to address this issue is represented by the Low-Cost miniaturized gas Sensors (LCSs) employed in the Low-Cost air quality Monitors (LCMs). Despite the various and unquestionable points of strength characterizing these devices, the scientific community has raised several warnings about the accuracy of their measurements and many caveats in their use. In this study, a new LCM model designed and implemented in our laboratories has been used to perform the measurements of the NO_2 and PM concentrations in the industrial area of Brindisi (Italy). Data gathered by the LCM have been compared with reference instrumentations for a rigorous analysis of the performance achievable through these low-cost technologies in this particular case.

Keywords: air quality monitoring; low-cost sensors; air pollutants; NO_2 low-cost sensors; particulate matter devices; case-study

Citation: To be added by editorial staff during production.

Academic Editor: First name Last name

Published: date

1. Introduction

Several studies have demonstrated the existence of a direct link between exposure to air pollutants and issues concerning the public health [1,2]. Air pollution levels are monitored by the equipment of governmental agencies, but they are characterized by high costs due to their purchasing, maintenance, and logistical issues [3]. As a consequence, often times, few monitoring points are deployed on the territory, causing the impossibility to build air pollutant maps having an adequate spatio-temporal resolution [3,4]. This issue has been addressed in recent years by developing low-cost technologies which have introduced the Low-Cost miniaturized Sensors (LCSs) and the Low-Cost air quality Monitors (LCMs) in the worldwide market [5,6]. The strong points of such devices are directly linked to their cheapness compared to the regulatory-grade equipment, their high portability grade, and their low power consumption [7]. The appealing characteristics concerning these technologies have also induced the scientific community to investigate their use in application areas very similar to the classic "air pollution monitoring", such as, for example, the malodor detection near landfill sites [8,9]. However, the flip side of these technologies is represented by a lower accuracy of measurements compared to the regulatory-grade equipment due to the interfering effects of the environmental variables, such as, temperature and humidity, and also due to their insufficient sensitivity and selectivity [3,10,11]. To improve the performance of LCSs and LCMs, different strategies have been explored, but the most promising appear to be the application of Artificial Neural Networks (ANNs), Multivariate Linear Regression (MLR) algorithms, or other machine learning techniques to the data representing the measurements performed through such devices [3,7,12,13,14,15,16,17,18].

In this work, a LCM called SentinAir was used for monitoring the NO_2, $PM_{2.5}$ and PM_{10} concentrations in the industrial area of Brindisi (Italy). This site is located near the city center and, therefore, the concentrations of the aforementioned pollutants in this area can significantly affect the air quality in the adjacent locations. The LCM used in this experiment was equipped with several LCSs: a couple of NO2B43F sensors by Alphasense [19] for NO_2 detection, and three samples of PMS5003 sensors by Plantower [20] for measuring $PM_{2.5}$ and PM_{10} concentrations.

2. Materials and methods

The LCM used for this experiment is a device designed and implemented in the laboratory of the ENEA Research Center of Brindisi called SentinAir. A complete description of the hardware and the software of SentinAir can be found in earlier published articles [21,22]. The LCM was used to measure NO_2, $PM_{2,5}$, and PM_{10} concentrations in a location which coordinates are 40°38'03.6"N, 17°58'39.0"E. Data quality provided by this device was evaluated through the use of reference instruments. NO_2 sensor measurements were compared with the 405 nm NOx monitor by 2Btech [23], while the APM-2 by Comde-Derenda [24] was employed for assessing PM measurements.

As concerning the performance indicators, the coefficient of determination (R^2), the Mean Absolute Error (MAE), and the Root Mean Squared Error (RMSE) were adopted to understand the grade of reliability of data provided by the LCSs. R^2, MAE, and RMSE are defined as follows:

$$R^2 = \frac{(\sum_{i=1}^{N}(s_i-\bar{s})(r_i-\bar{r}))^2}{\sum_{i=1}^{N}(s_i-\bar{s})^2\,(r_i-\bar{r})^2} \quad (1),$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|s_i - r_i| \quad (2),$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(s_i - r_i)^2} \quad (3),$$

where r_i represents the i_{th} measurement of the reference, s_i is the i_{th} concentration value given by the sensor, N is the number of observations, $\bar{s}$ represents the average of the sensor concentration measurements, and $\bar{r}$ is the average of the reference measurements. Data elaboration has been carried out by using the Scikit-learn libraries written in Python language [25-28]. Python platform is an open-source software freely downloadable from the on-purpose website [26].

The sampling rate of both LCM and reference instruments was set to 5 minutes, and therefore hourly averages were considered for each pollutant.

The analog output signals provided by the electrochemical sensors for NO_2 measurements were converted into gas concentrations by using a Multi-Layer Perceptron (MLP) [29], which is an ANN already successfully used in previous works [30,31]. The MLP used in this work has three hidden layers composed of respectively 150, 50, and 150 neurons, which have a "logistic" activation function (for further details, see [26,32]).

The PMS5003 sensors provide PM_{10} and $PM_{2.5}$ concentrations as output; therefore, we compared their measurements with the reference ones, and subsequently, a MLR algorithm was applied to improve their performance. It is already known that the ambient humidity is a source of error for optical PM sensors [33,34]. For this reason, the relative humidity and the raw PM concentrations given by the sensor outputs were selected as predictors for the MLR model [34].

3. Results

The NO_2 measurements started on the 28th of April 2023 and ended on the 2nd of May 2023 due to an unexpected breakdown of the reference instrument. The final dataset composed of hourly averages was split in two parts containing roughly the same number of records. The first one (the calibration period) was used to train the ANN, while the second one (the validation period) was used to validate the measurements. The predictors used as inputs for the ANN were the signals of the "working" and "auxiliary" electrodes of the NO2B43F sensors [19], and the relative humidity. In table 1, some statistics of the environmental variables monitored are reported.

Table 1. Statistics of the environmental variables monitored during the NO_2 measurements.

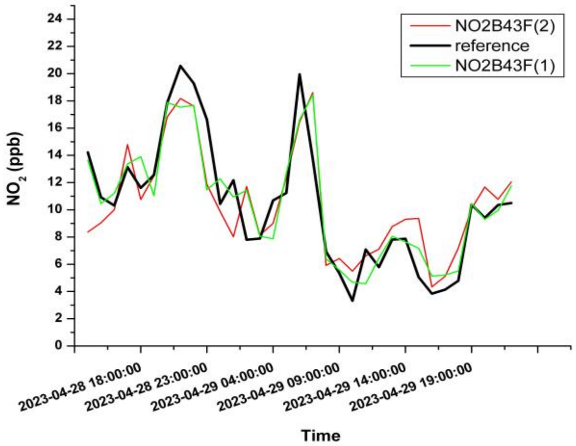
	Calibration period			Validation period		
	Min	**Max**	**Median**	**Min**	**Max**	**Median**
NO_2 [ppb]	3.3	20.6	10.4	2.8	20.9	6.3
T [°C]	12	26	20	13	26	17
RH[%]	32	78	49	43	88	73

The performance indicators related to the two NO_2 sensors installed inside the LCM are reported in table 2.

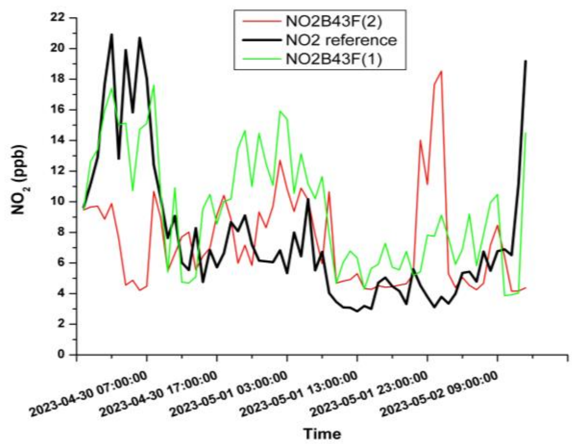
Table 2. Performance indicators of NO_2 sensors.

	Calibration period			Validation period		
	R^2	**MAE [ppb]**	**RMSE [ppb]**	**R^2**	**MAE [ppb]**	**RMSE [ppb]**
NO2B43F(1)	0.818	1.4	2.0	0.439	3.4	3.6
NO2B43F(2)	0.727	1.9	2.4	0.005	3.8	5.9

In figure 1, the plots of the time series referring to the NO_2 measurements are exposed for a better understanding of the LCS performance.



(**a**)



(**b**)

Figure 1. Time series of NO_2 measurements related to: (**a**) the calibration or training period of the ANN; (**b**) the validation of the ANN predictions.

The PM measurements were carried out from the 22th of May 2023 to the 4th of July 2023. Due to an unexpected power failure, data from the 16th of June to the 21th of June were lost. The final dataset consisting of these hourly averages was split in two parts, each of them containing 1053 records: the first one was used for the "calibration" of the MLR model, while the second one for its "validation".

In table 3, some significant statistics about the most relevant environmental variables are summarized.

Table 3. Statistics of the environmental variables monitored during the PM measurements.

	Calibration period			**Validation period**		
	Min	**Max**	**Median**	**Min**	**Max**	**Median**
$PM_{2.5}$ [μg/m³]	4.2	29.3	10.0	0.3	43.8	8.0
PM_{10} [μg/m³]	10.2	54.8	24.8	4.6	94.9	20
T [°C]	17	34	24	20	38	26
RH[%]	28	80	62	26	88	59

In table 4, the performance indicators related to the raw measurements carried out by the LCM (without the application of the MLR algorithm) are reported; while in table 5, the performance indicators representing the improvements achieved through the application of the MLR algorithm are shown.

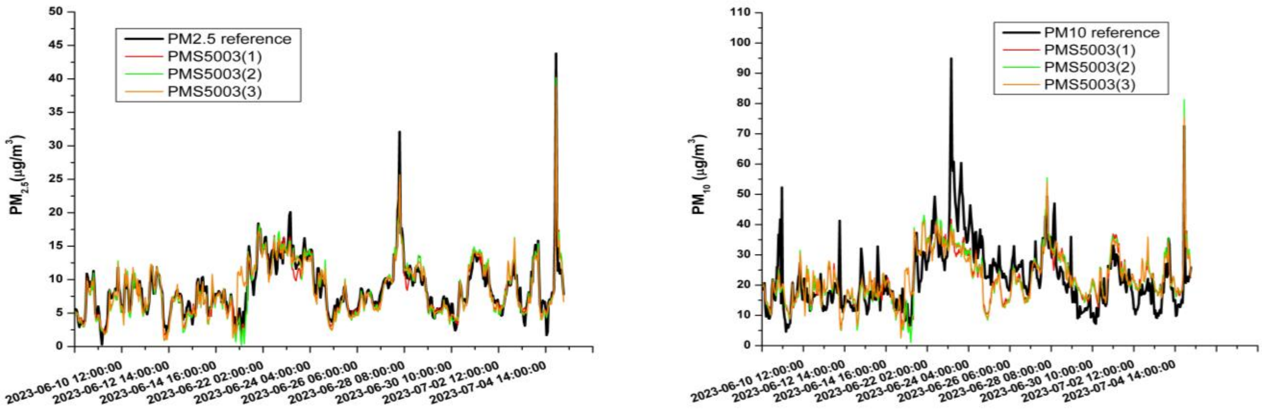
Table 4. Data related to the three sensors installed in the LCM under test and to the dataset composed of raw measurements.

		R^2	**MAE [μg/m³]**	**RMSE [μg/m³]**
PMS5003(1)		0.411	7.9	9.6
PMS5003(2)	PM_{10}	0.391	7.5	10.0
PMS5003(3)		0.359	8.3	10.8
PMS5003(1)		0.859	9.9	11.3
PMS5003(2)	$PM_{2.5}$	0.854	12.5	14.0
PMS5003(3)		0.835	14.2	15.8

To offer a more intuitive view of the improvements achieved by considering the effect of the relative humidity, in figure 2, the plots of the time series related to the PM sensor data after the application of the MLR algorithm are shown along with the reference measurements.

Table 5. Data related to the three sensors installed in the LCM under test. The performance indicators were computed for both the calibration and the validation dataset.

		Calibration period			**Validation period**		
		R^2	**MAE [μg/m³]**	**RMSE [μg/m³]**	**R^2**	**MAE [μg/m³]**	**RMSE [μg/m³]**
PMS5003(1)		0.567	4.9	6.8	0.495	5.8	7.8
PMS5003(2)	PM_{10}	0.541	5.1	6.9	0.481	5.9	7.9
PMS5003(3)		0.533	5.2	7.0	0.448	6.1	8.1
PMS5003(1)		0.843	1.2	1.7	0.906	0.9	1.3
PMS5003(2)	$PM_{2.5}$	0.834	1.2	1.7	0.907	1.0	1.3
PMS5003(3)		0.863	1.0	1.5	0.863	1.0	1.5



(**a**) (**b**)

Figure 2. Time series plots of the PM data in the validation period related to: (**a**) $PM_{2.5}$ measurements; (**b**) PM_{10} measurements.

4. Discussion and conclusion

The NO_2 concentration levels during the experiment were very low (20.9 ppb as maximum value). By considering the validation period, this element has determined a poor capability to follow the trend of the reference for the NO2B43F(2) sensor ($R^2 = 0.005$), and a moderate capability for the NO2B43F(1) ($R^2 = 0.439$). This difference is likely due to the different sensitivity characterizing the two sensors. As matter of fact, the sensor sensitivity can vary from 200 nA/ppm to 650 nA/ppm (see [19]), namely, one sensor can have more than three times the sensitivity of another one. This element can cause a remarkable difference in terms of R^2. The global indication that we can get from these data can be summarized by the fact that electrochemical sensors tested in this experiment have a limited capability to provide reliable measurements in case of environments very low polluted by NO_2.

The PM sensors under test have shown a good performance in the case of $PM_{2.5}$ measurements and a moderate performance in the case of PM_{10} (see table 4). By considering the relative humidity variable, we found that their performance has been further improved. This aspect suggests us that the measurements of $PM_{2.5}$ provided by these sensors can be considered very reliable; while, in the case of PM_{10}, we can conclude that their measurements are moderately reliable, even though we consider the effect of the relative humidity.

Author Contributions: Conceptualization, D.S.; methodology, D.S.; software, D.S.; validation, D.S.; formal analysis, D.S.; investigation, D.S.; resources, M.P.; data curation, D.S.; writing-original draft preparation, D.S.; writing-review and editing, D.S.; visualization, D.S.; supervision, M.P.; project administration, M.P.; funding acquisition, M.P.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: please**,** contact the author at domenico.suriano@enea.it.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kheirbek, I.; Wheeler, K.; Walters, S.; Kass, D.; Matte, T. $PM_{2.5}$ and ozone health impacts and disparities in New York City: Sensitivity to spatial and temporal resolution. *Air Qual. Atmos. Health,* **2013**, *6*, 473–486.
2. Makri, A.; Stilianakis, N.I. Vulnerability to air pollution health effects. *Int. J. Hyg. Environ. Health,* **2018**, *211*, 326–336.
3. Karagulian, F.; Barbiere, M.; Kotsev, A.; Spinelle, L.; Gerboles, M.; Lagler, F.; Redon, N.; Crunaire, S.; Borowiak, A. Review of the Performance of Low-Cost Sensors for Air Quality Monitoring. *Atmosphere,* **2019**, *10*, 506

4. Suriano, D.; Cassano, G.; Penza, M. Design and Development of a Flexible, Plug-and-Play, Cost-Effective Tool for on-Field Evaluation of Gas Sensors. *J. Sens.* **2020**, *2020*, 8812025
5. Kumar, P.; Morawska, L.; Martani, C.; Biskos, G.; Neophytou, M.; Di Sabatino, S.; Bell, M.; Norford, L.; Britter, R. The rise of low-cost sensing for managing air pollution in cities. *Environ. Int.* **2015**, *75*, 199–205
6. Snyder, E.G.; Watkins, T.H.; Solomon, P.A.; Thoma, E.D.; Williams, R.W.; Hagler, G.S.; Shelow, D.; Hindin, D.A.; Kilaru, V.J.; Preuss, P.W. The Changing Paradigm of Air Pollution Monitoring. *Environ. Sci. Technol.* **2013**, *47*, 11369–11377
7. Suriano, D.; Penza, M. Assessment of the Performance of a Low-Cost Air Quality Monitor in an Indoor Environment through Different Calibration Models. *Atmosphere* **2022**, *13*, 567.
8. Suriano, D.; Rossi, R.; Alvisi, M.; Cassano, G.; Pfister, V.; Penza, M.; Trizio, L.; Brattoli, M.; Amodio, M.; De Gennaro, G. A Portable Sensor System for Air Pollution Monitoring and Malodours Olfactometric Control. In *Sensors and Microsystems: AISEM 2011 Proceedings*; Lecture Notes in Electrical Engineering; Springer: Boston, MA, USA, 2012; Volume 109
9. Guidi, V.; Carotta, M.C.; Fabbri, B.; Gherardi, S.; Giberti, A.; Malagù, C. Array of sensors for detection of gaseous malodors in organic decomposition products. *Sens. Actuators B Chem.* **2012**, *174*, 349–354.
10. Lewis, A.; Edwards, P. Validate personal air-pollution sensors. *Nature* **2016**, *535*, 29–31.
11. Kang, Y.; Aye, L.; Ngo, T.D.; Zhou, J. Performance evaluation of low-cost air quality sensors: A review. *Sci. Total Environ.* **2022**, *818*, 151769.
12. Trizio, L.; Brattoli, M.; De Gennaro, G.; Suriano, D.; Rossi, R.; Alvisi, M.; Cassano, G.; Pfister, V.; Penza, M. Application of artificial neural networks to a gas sensor-array database for environmental monitoring. In *Sensors and Microsystems*; Springer: Boston, MA, USA, 2012; pp. 139–144
13. Bigi, A.; Mueller, M.; Grange, S.K.; Ghermandi, G.; Hueglin, C. Performance of NO, NO2 low cost sensors and three calibration approaches within a real world application. *Atmos. Meas. Tech.* **2018**, *11*, 3717–3735
14. Spinelle, L.; Gerboles, M.; Villani, M.G.; Aleixandre, M.; Bonavitacola, F. Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: Ozone and nitrogen dioxide. *Sens. Actuators B Chem.* **2015**, *215*, 249–257
15. Yamamoto, K.; Togami, T.; Yamaguchi, N.; Ninomiya, S. Machine Learning-Based Calibration of Low-Cost Air Temperature Sensors Using Environmental Data. *Sensors* **2017**, *17*, 1290
16. Si, M.; Xiong, Y.; Du, S.; Du, K. Evaluation and calibration of a low-cost particle sensor in ambient conditions using machine-learning methods. *Atmos. Meas. Tech.* **2020**, *13*, 1693–1707
17. Cordero, J.M.; Borge, R.; Narros, A. Using statistical methods to carry out in field calibrations of low cost air quality sensors. *Sens. Actuators B Chem.* **2018**, *267*, 245–254
18. Kosmopoulos, G.; Salamalikis, V.; Pandis, S.N.; Yannopoulos, P.; Bloutsos, A.A.; Kazantzidis, A. Low-cost sensors for measuring airborne particulate matter: Field evaluation and calibration at a South-Eastern European site. *Sci. Total Environ.* **2020**, *748*, 141396
19. NO2-B43F sensor datasheet. Available online:https://www.alphasense.com/wp-content/uploads/2022/09/Alphasense_NO2-B43F_datasheet.pdf(accessed on 26th of June 2023).
20. PMS5003 sensor datasheet. Available online:https://www.plantower.com/en/products_33/74.html(accessed on 26th of June 2023).
21. Suriano, D. A portable air quality monitoring unit and a modular, flexible tool for on-field evaluation and calibration of low-cost gas sensors. *HardwareX* **2021**, *9*, e00198.
22. Suriano, D. SentinAir system software: A flexible tool for data acquisition from heterogeneous sensors and devices. *SoftwareX* **2020**, *12*, 100589.
23. Model 405nm monitor. Available online: https://twobtech.com/model-405-nm-nox-monitor.html(accessed on 26th of June 2023).
24. APM-2 monitor. Available online: https://www.comde-derenda.com/produkte/apm-2-2 (accessed on 26th of June 2023).
25. Python website: https://www.python.org (accessed on 26th of June 2023)
26. Scikit website: https://scikit-learn.org/stable/index.html (accessed on 26th of June 2023)
27. Lars Buitinck, Gilles Louppe, et al., API design for machine learning software: experiences from the scikit-learn project, 2013, https://arxiv.org/abs/1309.0238v1
28. Pedregosa et al., "Scikit-learn: Machine Learning in Python", JMLR 12, pp. 2825-2830, 2011
29. Rumelhart, D.E.; McClelland, J.L.; P.D.P. Research Group (Eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*; MIT Press: Cambridge, MA, USA, 1986.
30. Spinelle, L.; Gerboles, M.; Villani, M.G.; Aleixandre, M.; Bonavitacola, F. Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: Ozone and nitrogen dioxide. *Sens. Actuators B Chem.* **2015**, *215*, 249–257.
31. Spinelle, L.; Gerboles, M.; Villani, M.G.; Aleixandre, M.; Bonavitacola, F. Field calibration of a cluster of low-cost commercially available sensors for air quality monitoring. Part B: NO, CO and CO_2. *Sens. Actuators B Chem.* **2017**, *238*, 706–715.
32. Bishop, C.M. *Neural Networks for Pattern Recognition*; Oxford University Press Inc.: New York, NY, USA, 1995.
33. Suriano, D.; Prato, M. An Investigation on the Possible Application Areas of Low-Cost PM Sensors for Air Quality Monitoring. *Sensors* **2023**, *23*, 3976. https://doi.org/10.3390/s23083976
34. Lee, H.; Kang, J.; Kim, S.; Im, Y.; Yoo, S.; Lee, D. Long-Term Evaluation and Calibration of Low-Cost Particulate Matter (PM) Sensor. *Sensors* **2020**, *20*, 3617. https://doi.org/10.3390/s20133617