## SUPPLEMENTARY MATERIAL

**INHALATION OF SMALL PARTICLES (PM2.5) IN URBAN ROAD TUNNELS AND UNDERGROUND. MADRID (SPAIN). A CITIZEN SCIENCE PROJECT**

**---------------------------------------------------------------------------------**

**PERFORMANCE EVALUATION AND UTILITY OF A LOW-COST PARTICULATE MATTER SENSOR**

**FOR USE IN A CITIZEN SCIENCE CONTEXT**

Ángel López-Encuentra, Esther Gil Cid, Víctor Abraira, Luis Miguel Pozo Coronado.

In the European Union, standards have been established for placing sensors in different urban locations to measure all air pollutants, including PM2.5. One problem is that in "hidden" areas not covered by official stations, there may be elevated and persistent concentrations of PM2.5. Therefore, low-cost sensors (LCS) have become popular to assess the presence of various pollutants in neighborhoods near high-traffic volume roads: population, educational institutions, sports areas, or elderly care facilities, which are distant from fixed stations.

The aim of this study is to determine the real utility of low-cost sensors (LCS) PM2.5 concentration meters and its validation to establish if a threshold value is exceeded. The Airbeam, in all its versions (AB1, AB2, and AB3), uses light scattering sensors to measure PM2.5 concentration and is not without accuracy problems.

Calibrating LCS for atmospheric contamination data can be done either in controlled laboratory environment, or field setup with a sensor from the reference monitoring network; for indoor and outdoor environments; with networks or isolated sensors. The calibration in authorized institutions is usually out of the question in citizen science, where funding is scarce. Besides, different techniques are applied for calibrating the accuracy and precision of the instruments. For field validation by comparison with a sensor from the reference monitoring network, data frames of very different lengths of data can be found, from hours to months. These techniques for ambient validation include, mainly, relative humidity correction factor, regression models and machine learning. Previous validation works in AB sensors show no significant deviations from laboratory or field setup validation form AB sensor. Most of the cited works agree in several aspects: in general, the AB3 has good accuracy and precision, although moderate measure difference between different devices may exist; the AB3 tends to underestimate the PM2.5 concentrations, especially for low concentrations; and the accuracy of AB3 is less reliable for very high concentrations.

There is a need to ensure that "each LCS device" must be validated using accessible methods so that they can be employed by so-called citizen science. This is especially important if the data they produce has significant implications for public health decision-making, backed by scientific evidence. For that purpose, it could be enough to obtain, as a result of the validation process, a threshold measure in the AB3 that ensures that the limit for health hazard PM2.5 concentration has been exceeded. Both goals (the accessible method for validation and the easily usable threshold as a result) constitute the justification for this study.

From December 5, 2023, to January 20, 2024, an Airbeam3 (AB3) sensor was placed in downtown Madrid, at the "Castellana" municipal fixed station of the Air Quality Surveillance Network of the Madrid City Council, which measures PM2.5 using the Tapered Element Oscillating Microbalance method (TEOM) (Table S1). The distance between the AB3 and the air inlet at the fixed station for PM2.5 measurement was one meter. Both are located 3.5 meters above ground. The station is placed near a 14-lane avenue with high traffic intensity, about 77,000 vehicles per day.

To analyze the relationship between AB3 values and government hourly data, an initial analysis was conducted, obtaining the correlation coefficient between the hourly averages obtained from both instruments. All linear model fitting calculations were performed using RStudio.

An empirical model was established between 10-minute averages from AB3 measurements, and 10-minute averages obtained by government fixed station (TEOM). This involved a two-step process: first, studying the correlation and differences between both records, and then determining eight adjustment models through a training model, which were contrasted with the test set.

Based on this study, two models were established for subsequent observations. The first model (Model A) considers temperature and relative humidity, as well as an intercept, allowing for precise adjustment. The second model (Model

B) is based solely on PM2.5 observations, with an intercept, establishing confidence intervals. This second model allows for establishing minimum PM2.5 levels from future observations.

Using a different approach, to study the agreement between both datasets the Bland and Altman method was used, which involves plotting the difference between AB3 and TEOM hourly measurements against their mean and calculating the mean differences and their confidence intervals. To assess the validity of AB3, its ROC curve was constructed for the TEOM 10 μg/m³ threshold (political limit of European Union), and predictive values for AB3 at that cutoff point were calculated using Stata v16.1 statistical software.

Several linear models were fitted using PM2.5 values from AB3 and official sources, along with temperature and relative humidity values. In total, eight adjustment models were obtained: only with PM2.5 values; with temperature and RH; only with temperature; only with relative humidity; all in versions with and without an intercept (linear term containing the expected outcome when AB3 PM2.5 is zero).

Using the training set, adjustment coefficients, their standard errors, model residual standard error, multiple R-squared, adjusted R-squared, F-statistic, and p-value were obtained for each model. Additionally, Pearson product-moment correlation between official data and adjusted data was calculated using adjustment coefficients, resulting in a correlation coefficient and a 95% confidence interval. After obtaining the model with the training set, the test set was used to verify the quality of the fit. Residuals from the test set, determined as official PM2.5 minus adjusted PM2.5, were also obtained. Additional tests were conducted on test set residuals to determine their normality (Kolmogorov-Smirnov), and the mean, median, 1st quartile, 3rd quartile, minimum, and maximum from the residuals were obtained.

The optimal model selection was based on all the aforementioned parameters and statistics (Table S2; Figure S1). Models obtained with the training test show small differences between coefficients and their standard errors and within Pearson correlations. However, F-statistic values were significantly different and show that the optimal model was obtained with PM2.5, temperature, RH, and intercept (Model A). Model B only considers PM2.5 values and intercept and was also selected because it showed a good performance and was much simpler (Table S3; Table S4).

In the Bland-Altman plot, the mean difference (average error margin) is -0.94. The plot shows the mean difference (-0.94) as a solid line and the 95% CI limits (-9.84 to 7.96) as dashed lines. However, the plot indicates that it is not constant: at higher pollution values, the difference changes direction (tending to overestimate) and increases its variation. Choosing a value of 20 as the mean of the two methods, the mean difference is -1.75 (95% CI -7.97 to 4.47) for values below 20 (Figure S2).

The use of neuronal networks was discarded, as these results as accurate enough for detecting harmful levels of PM2.5, easy enough to be replicated in citizen science works and completely explicable without hidden layers of a multilayer perceptron neuronal network.

The strength of this study lies in its potential value for citizen science in the field of pollution control, since it provides a reliable threshold value for measures with the AB3·whose surpassing guarantees a health hazard, but also a simple method (based on a regression model with just one independent variable) to compute that threshold for each AB3 device.

We have pointed out some deficiencies of LCS, which justify that there is a need to ensure that "each LCS device" must be validated using accessible methods so that they can be employed by so-called "citizen science". This is especially important if the obtained data can have significant implications for public health decision-making, backed by scientific evidence.

**CONCLUSION**

Some low-cost sensors are useful for measuring small-sized particles (PM2.5), but their validity must be evaluated against a gold standard. In this work, the validation of an AIRBEAM3 sensor, placed in a fixed government installation, is carried out using regression models and Bland-Altman method. These methods are simple enough to be used in a citizen science context. For the linear regression eight linear models were developed from a randomly selected training set. These models were tested on the remaining data, used as test set. From them, two were identified as optimal: the first one, multivariate, predicting PM2.5 from PM2.5, temperature, and relative humidity as measured by the AIRBEAM3; the second model, univariate, predicting PM2.5 only from PM2.5 measurements. Both models showed similar Pearson correlation coefficients between observed and predicted values (0.9301 and 0.9245, respectively), with similar 95% confidence intervals. Regarding the Bland-Altman method, the difference and variability between the hourly data series were greater at higher pollution levels. Therefore, the analysis was repeated for average values above and below the mean of both methods determinations (20 μg/m³), resulting in a difference of 2.56 (95%CI -11.49 to 16.62) and -1.75 (95%CI -7.97 to 4.47), respectively. This study thus stablishes that the probability that the AIRBEAM3 correctly identifies PM2.5 concentrations above 10 μg/m³ (2024-European Union limit) is 96% (Table S5; Figure S3); the probability of the sensor correctly identifying PM2.5 concentrations below 10 μg/m³ is 86% (Table S6; Table S7). Adjustment with the linear model strengthens the conclusions regarding the validity of AIRBEAM3 measurements as predictors of excessive PM2.5 levels.

| | Mean | SD | Min. | 1st Qu | Median | 3rd Qu. | Max. |
|---|---|---|---|---|---|---|---|
| AB3 PM2.5 1min | 10.55 | 12.5118 | 0 | 1 | 6 | 16 | 103 |
| TEOM PM2.5 10 min | 11.47 | 10.3256 | 1 | 4 | 8 | 16 | 58 |
| TEOM PM2.5 hourly mean | 11.44 | 10.1694 | 1 | 4 | 8 | 16 | 55 |
| AB temperatura | 56.91 | 5.8764 | 42 | 52 | 57 | 61 | 96 |
| AB RH | 57.15 | 11.8439 | 24 | 49 | 58 | 66 | 93 |

**Table S1.** A summary of raw data of AB one-minute PM2.5, temperature and relative humidity (RH) (%) and ten-minutes mean and hourly mean from TEOM official data.

| | MIN | 1Q | Median | Mean | 3Q | MAX | Kolmogorov Smirnov | SD |
|---|---|---|---|---|---|---|---|---|
| **Model A** | -38.9024 | -2.2281 | -0.1363 | 0.0621 | 1.9235 | 17.7847 | 0.072675<br>p-value: < 2.2e-16 | 4.012651 |
| **Model B** | -39.4722 | -2.2734 | -0.2410 | 0.1993 | 2.2249 | 19.2143 | 0.078143<br>p-value: < 2.2e-16 | 4.134476 |

**Table S2.** Comparison between residuals of the test set for models A and B: minimum value (Min), 1Q, Median, Mean, 3Q, maximum value (Max), Kolmogorov-Smirnov normality test and standard deviation.

| | Residual standard error | Multiple R-squared | Adjusted R-squared | F. statistic | Correl Pearson | p-valor | 95 % Pearson confidence interval |
|---|---|---|---|---|---|---|---|
| **Model A** | 3.751 on 5075 DF | 0.8651 | 0.865 | 1.085e+04 on 3 and 5075 DF | 0.9300983 | < 2.2e-16 | (0.9262903,0.9337164) |
| **Model B** | 3.892 on 5077 DF | 0.8547 | 0.8546 | 2.986e+04 on 1 and 5077 DF | 0.9244817 | < 2.2e-16 | (0.9203804, 0.9283797) |

**Table S3**. Results of different statistics for the training set

| | Coef. PM2.5 | Estimated Std Error | Coef. Temp | Estimated Std. Error | Coef. RH | Estimated Std Error | Intercept | Estimated Std Error |
|---|---|---|---|---|---|---|---|---|
| **Model A** | 0.766299 | 0.004299 | 0.035027 | 0.009140 | -0.084560 | 0.004498 | 6.201582 | 0.61625 |
| **Model B** | 0.767020 | 0.004439 | ---- | ---- | ---- | ---- | 3.326358 | 0.071751 |

**Table S4**. Values of the adjustment coefficients of PM2.5, temperature (temp), relative humidity (RH) and intercept of Models A and B, as well as their typical error.

| | AB3 PM2.5 >10 µg/m³ | AB3 PM2.5 <10 µg/m³ | TOTAL |
|---|---|---|---|
| TEOM. PM2.5 >10 µg/m³ | 364 | 97 | 461 |
| TEOM. PM2.5 <10 µg/m³ | 15 | 597 | 612 |
| TOTAL | 379 | 694 | 1.073 |

**Table S5.** Distribution of data according to their upper or lower value of PM2.5 10 ng/m3 for AB3 and the reference station (TEOM).

Prevalence TEOM PM2.5 ≥10 µg/m³: 461/1073: 43%

----------------------------------------------------------------------------------------------------

| **Concept** | **Frecuency** | **IC95%** |
|---|---|---|
| Prevalence | 43% | 40% - 46% |
| Sensitivity | 79% | 74.9% - 82.6% |
| Specificity | 97.5% | 96% - 98.6% |
| Likelihood ratio (+) | 32.2 | 19.5 - 53.2 |
| Likelihood ratio (-) | .216 | .181 - .258 |
| Odds ratio | 149 | 85.8 - 260 |
| Positive predictive value | 96% | 93.6% - 97.8% |
| Negative predictive value | 86% | 83.2% - 88.5% |

**Table S6.** Operational characteristics of the test under study (AB3)

| **Prevalence** | **20%** | **30%** | **43%** | **50%** | **60%** |
|---|---|---|---|---|---|
| Positive predictive value | 89% | 93% | 96% | 97% | 98% |
| Negative predictive value | 95% | 92% | 86% | 82% | 76% |

**Table S7.** Predictive values with different prevalences (Bayes theorem)

Values and confidence intervals are based on likelihood ratios, assuming that the prevalence is known exactly.

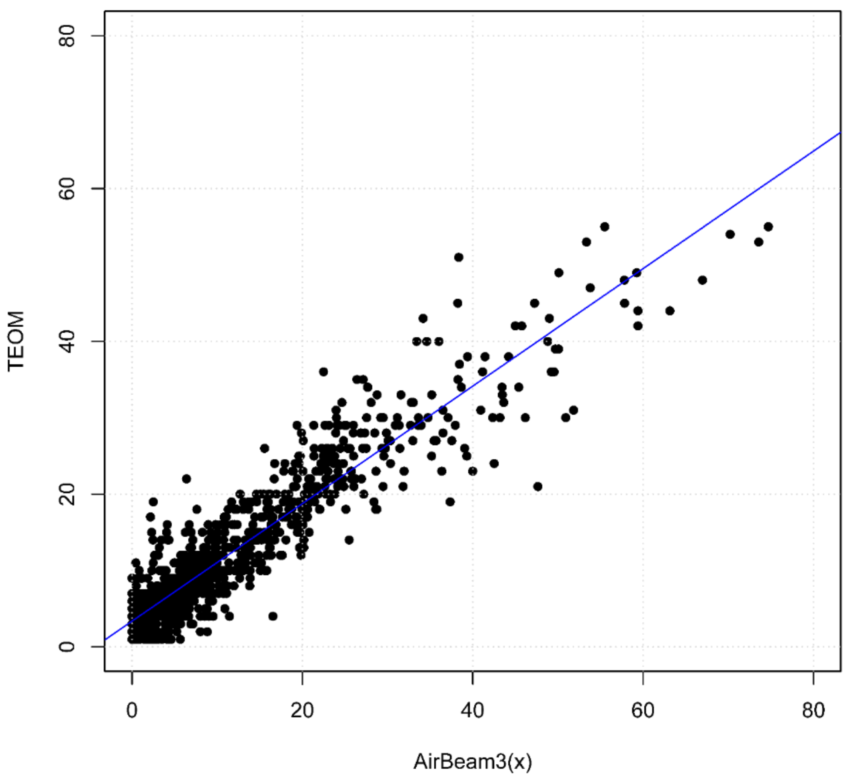


**FIGURE S1**

DISPERSION OF CONCENTRATION DATA OF PM2.5
PM2.5: Small-sized particles with a diameter of 2.5 micrometers
TEOM: Tapered Element Oscillating Microbalance method.

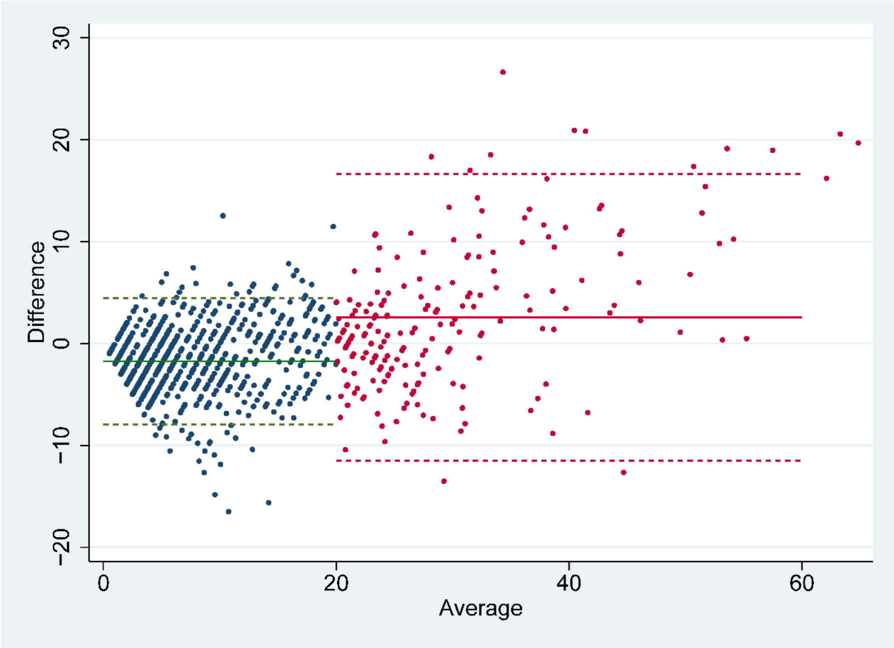


**FIGURE S2**

BLAND–ALTMAN CHART. Difference AB3-TEOM versus average with 95% limits of agreement. The solid lines are the mean and the dotted lines are the limits. Measurements in average lower than 20 µg/$m^3$ in blue and measurements higher in red.

95% CI: dashed lines

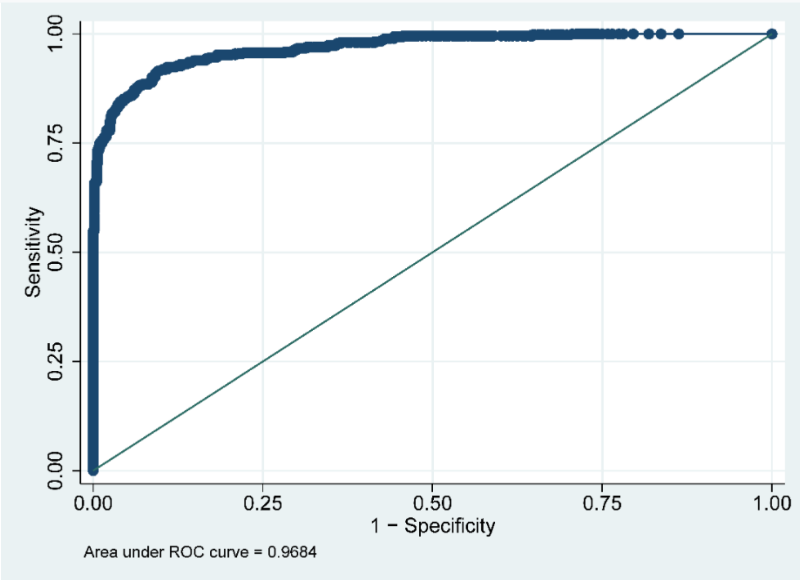


**FIGURE S3**

ROC curve of AB3 considering contamination TEOM values greater than or equal to 10 µg/$m^3$

**Acknowledgements:** Narros GA (Escuela Superior de Ingenieros Industriales. UPM), Madres por el Clima Lourdes, Coordinadora Vecinal del Distrito Latina (COVELA), ENVIRA Sostenible y Servicio de Calidad del Aire del Ayuntamiento de Madrid.

## REFERENCES

**Aircasting: AQ-SPEC; South Coast Air Quality Management District (SCAQMD)** (n.d.) *Air Quality Sensor Performance Evaluation Reports*. Available at: http://www.aqmd.gov/aq-spec/evaluations#&MainContent_C001_Col00=2 (Accessed: 16 July 2025).

**Ali, S., Glass, T., Parr, B., Potgieter, J. and Alam, F.** (2020) 'Low cost sensor with IoT LoRaWAN connectivity and machine learning-based calibration for air pollution monitoring', *IEEE Transactions on Instrumentation and Measurement*, 70, pp. 1–11. doi: 10.1109/TIM.2020.3034109

**Anastasiou, E., Vilcassim, M.J.R., Adragna, J., Gill, E., Tovar, A., Thorpe, L.E. *et al.*** (2022) 'Feasibility of low-cost particle sensor types in long-term indoor air pollution health studies after repeated calibration, 2019–2021', *Scientific Reports*, 12, 14571. doi:10.1038/s41598-022-18200-0

**Ayuntamiento de Madrid** (n.d.) *Portal de datos abiertos: Calidad del aire. Estaciones de control*. Available at: https://datos.madrid.es/sites/v/index.jsp?vgnextoid=f3c0f7d512273410VgnVCM2000000c205a0aRCRD&vgnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD (Accessed: 26 October 2024).

**Baldelli, A.** (2021) 'Evaluation of a low-cost multi-channel monitor for indoor air quality through a novel, low-cost, and reproducible platform', *Measurement: Sensors*, 17, 100059. doi: 10.1016/j.measen.2021.100059

**Bland, J.M. and Altman, D.G.** (1986) 'Statistical methods for assessing agreement between two methods of clinical measurement', *The Lancet*, 1(8476), pp. 307–310. PMID: 2868172

**Borghi, F., Spinazzè, A., Campagnolo, D., Rovelli, S., Cattaneo, A. and Cavallo, D.M.** (2018) 'Precision and accuracy of a direct-reading miniaturized monitor in PM2.5 exposure assessment', *Sensors*, 18(9), 3089. doi: 10.3390/s18093089

**Castell, N., Dauge, F.R., Schneider, P., Vogt, M., Lerner, U., Fishbain, B., Broday, D. and Bartonova, A.** (2017) 'Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?', *Environment International*, 99, pp. 293–302. doi: 10.1016/j.envint.2016.12.007

**Cho, H. and Baek, Y.** (2021) 'Practical particulate matter sensing and accurate calibration system using low-cost commercial sensors', *Sensors*, 21(18), 6162. doi: 10.3390/s21186162

**Council of the European Union** (2023) 'Air quality: Council ready to start talks with Parliament on new rules to strengthen standards in the EU', *Press release*, 9 November. Available at: https://www.consilium.europa.eu/es/press/press-releases/2023/11/09/air-quality-council-ready-to-start-talks-with-parliament-on-new-rules-to-strengthen-standards-in-the-eu/ (Accessed: 26 October 2024).

**Core Team, R.** (2024) *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available at: https://www.r-project.org (Accessed: 26 October 2025).

**DeWitt, H.L., Crow, W.L. and Flowers, B.** (2020) 'Performance evaluation of ozone and particulate matter sensors', *Journal of the Air & Waste Management Association*, 70(3), pp. 292–306. doi: 10.1080/10962247.2020.1713921

**Fraisl, D., Hager, G., Bedessem, B. *et al.*** (2022) 'Citizen science in environmental and ecological sciences', *Nature Reviews Methods Primers*, 2, 64. doi: 10.1038/s43586-022-00144-4

**Fritz, S., See, L. and Grey, F.** (2022) 'The grand challenges facing environmental citizen science', *Frontiers in Environmental Science*, 10, 1019628. doi: 10.3389/fenvs.2022.1019628

**González Rivero, R.A., Morera Hernández, L.E., Schalm, O., Hernández Rodríguez, E., Alejo Sánchez, D., Morales Pérez, M.C. *et al.*** (2023) 'A low-cost calibration method for temperature, relative humidity, and carbon dioxide sensors used in air quality monitoring systems', *Atmosphere*, 14(2), 191. doi: 10.3390/atmos14020191

**HabitatMap** (n.d.) *AirBeam: How it works*. Available at: https://www.habitatmap.org/airbeam/how-it-works (Accessed: 26 October 2024).

**Heimbinder, M. and Lim, C.** (n.d.) *AirBeam3 technical specifications, operation & performance*. HabitatMap. Available at: https://www.habitatmap.org/blog/airbeam3-technical-specifications-operation-performance (Accessed: 26 October 2024).

**Huang, J., Kwan, M.P., Cai, J., Song, W., Yu, C., Kan, Z. *et al.*** (2022) 'Field evaluation and calibration of low-cost air pollution sensors for environmental exposure research', *Sensors*, 22(6), 2381. doi: 10.3390/s22062381

**Jayaratne, R., Liu, X., Ahn, K.H., Asumadu-Sakyi, A., Fisher, G., Gao, J. *et al.*** (2020) 'Low-cost PM2.5 sensors: an assessment of their suitability for various applications', *Aerosol and Air Quality Research*, 20(3), pp. 520–532. doi: 10.4209/aaqr.2018.10.0390

**Karagulian, F., Barbiere, M., Kotsev, A., Spinelle, L., Gerboles, M., Lagler, F., Redon, N., Crunaire, S. and Borowiak, A.** (2019) 'Review of the performance of low-cost sensors for air quality monitoring', *Atmosphere*, 10(9), 506. doi: 10.3390/atmos10090506

**Lewis, A.C., Lee, J.D., Edwards, P.M., Shaw, M.D., Evans, M.J., Moller, S.J., Smith, K.R., Buckley, J.W., Ellis, M., Gillot, S.R. and White, A.** (2016) 'Evaluating the performance of low cost chemical sensors for air pollution research', *Faraday Discussions*, 189, pp. 85–103. doi: 10.1039/C5FD00201J

**Liang, L.** (2021) 'Calibrating low-cost sensors for ambient air monitoring: techniques, trends, and challenges', *Environmental Research*, 197, 111163. doi: 10.1016/j.envres.2021.111163
**Maah, B., Zhou, Z. and Thiele, L.** (2018) 'A survey on sensor calibration in air pollution monitoring deployments', *IEEE Internet of Things Journal*, 5(6), pp. 4857–4870. Available at: https://ink.library.smu.edu.sg/sis_research/4533 (Accessed: 26 October 2024).
**Mahajan, S., Chung, M.K., Martinez, J. *et al.*** (2022) 'Translating citizen-generated air quality data into evidence for shaping policy', *Humanities and Social Sciences Communications*, 9, 122. doi: 10.1057/s41599-022-01135-2
**Mei, H., Han, P., Wang, Y., Zeng, N., Liu, D., Cai, Q. *et al.*** (2020) 'Field evaluation of low-cost particulate matter sensors in Beijing', *Sensors*, 20(16), 4381. doi: 10.3390/s20164381
**Mukherjee, A., Stanton, L.G., Graham, A.R. and Roberts, P.T.** (2017) 'Assessing the utility of low-cost particulate matter sensors over a 12-week period in the Cuyama Valley of California', *Sensors*, 17(8), 1805. doi: 10.3390/s17081805
**Orellano, P., Reynoso, J., Quaranta, N., Bardach, A. and Ciapponi, A.** (2020) 'Short-term exposure to particulate matter (PM10 and PM2.5), nitrogen dioxide (NO2), and ozone (O3) and all-cause and cause-specific mortality: systematic review and meta-analysis', *Environment International*, 142, 105876. doi: 10.1016/j.envint.2020.105876
**Pocock, M.J.O., Tweddle, J.C., Savage, J. *et al.*** (2017) 'The diversity and evolution of ecological and environmental citizen science', *PLoS ONE*, 12(4), e0172579. doi: 10.1371/journal.pone.0172579
**Raysoni, A.U., Pinakana, S.D., Mendez, E., Wladyka, D., Sepielak, K. and Temby, O.** (2023) 'A review of literature on the usage of low-cost sensors to measure particulate matter', *Earth*, 4(1), pp. 168–186. doi: 10.3390/earth4010009
**Roddee, S., Changphuek, S., Jirakajohnkool, S., Tochaiyaphum, P., Phairuang, W., Chetiyanukornkul, T. *et al.*** (2024) 'Using a low-cost sensor to estimate fine particulate matter: a case study in Samutprakarn, Thailand', *Atmosphere*, 15(3), 336. doi: 10.3390/atmos15030336
**StataCorp** (n.d.) *Stata statistical software*. Available at: https://www.stata.com/ (Accessed: 26 October 2024).
**Vanbrabant, K., Van Dam, D., Bongaerts, E., Vermeiren, Y., Bové, H., Hellings, N. *et al.*** (2024) 'Accumulation of ambient black carbon particles within key memory-related brain regions', *JAMA Network Open*, 7(2), e245678. doi: 10.1001/jamanetworkopen.2024.5678
**Villanueva, E., Espezua, S., Castelar, G., Diaz, K. and Ingaroca, E.** (2023) 'Smart multi-sensor calibration of low-cost particulate matter monitors', *Sensors*, 23(7), 3776. doi: 10.3390/s23073776
**World Health Organization (WHO)** (2021) *WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide: executive summary*. Geneva: World Health Organization. Available at: https://www.who.int/publications/i/item/9789240034433 (Accessed: 26 October 2025).