atmosphere

Type of the Paper (Review)

Review of performance of Low-cost Sensors for Air Quality Monitoring

Federico Karagulian[1], Maurizio Barbiere[1], Alexander Kotsev[2], Laurent Spinelle[3,4], Michel Gerboles[1]*, Friedrich Lagler[1], Nathalie Redon[5], Sabine Crunaire[5] and Annette Borowiak[1]

[1] European Commission, Joint Research Centre (JRC), Directorate for Energy, Transport and Climate, Air and Climate Unit, Via Enrico Fermi 2749, 21027 Ispra VA, Italy, federico.karagulian@ec.europa.eu; michel.gerboles@ec.europa.eu; maurizo-giuseppe.barbiere@ec.europa.eu; friedrich.lagler@ec.europa.eu; annette.borowiak@ec.europa.eu

[2] European Commission, Joint Research Centre (JRC) Digital Economy Unit, Growth and Innovation Directorate, Via Enrico Fermi 2749, 21027 Ispra VA, Italy), alexander.kotsev@ec.europa.eu

[3] Institut national de l'Environnement Industriel et des Risques (INERIS), Parc technologique Alata, BP 2, F-60550 Verneuil-en-Halatte, France, laurent.spinelle@ineris.fr

[4] Laboratoire Central de Surveillance de la Qualité de l'Air (LCSQA), Parc technologique Alata, BP 2, F-60550 Verneuil-en-Halatte, France

[5] IMT Lille Douai, Univ. Lille, SAGE-Departement Sciences de l'Atmosphère et Génie de l'Environnement, F-59000 Lille, France, nathalie.redon@imt-lille-douai.fr; sabine.crunaire@imt-lille-douai.fr

* Correspondence: michel.gerboles@ec.europa.eu; Tel.: +39-0332-78-5652

Received: date; Accepted: date; Published: date

Abstract

A growing number of companies started commercialising low-cost sensors (LCS) that are said to be able to monitor air pollution in outdoor air. The benefit about the use of LCS is the increased spatial coverage when monitoring air quality in cities and remote locations. Today, there are hundreds of LCSs commercially available on the market with a cost ranging from a few hundred to a few thousand euro. At the same time, independent evaluation of the performance of LCSs against reference measurements is only available in literature for about 110 LCSs in literature. These studies report that LCS are unstable and often affected by atmospheric condition, cross sensitivities from interfering compounds that may change LCS performance depending on site location. In this work, quantitative data about the performance of tested LCS against reference measurement are presented. This information was gathered from published reports and relevant testing laboratories. Other information was drawn from peer-reviewed journals that tested different types of LCSs in research studies. Relevant metrics about the comparison of LCSs systems against reference systems highlighted the most cost-effective LCSs that could be used to monitor air quality pollutants with a good level of agreement represented by a coefficient of determination of $R^2 > 0.75$ and slope close to 1.0. This review highlights the possibility to have versatile LCSs able to operate with multiple pollutants and preferably with transparent LCS data treatment.

Keywords: electrochemical sensors, metal oxide sensors, optical particle counters, nephelometers, citizen science, performance evaluation, sensor validation, air quality monitoring.

1 Introduction

The widening of the commercial availability of micro-sensors technology is contributing to the rapid adoption of low-cost sensors for air quality monitoring by both citizen science initiatives and public authorities [1]. In general, public authorities want to increase the density of monitoring measurements and often want to rely on low-cost sensors because they cannot afford any reference Air Quality Monitoring Station (AQMS) [2]. Low-cost sensors can provide real time measurements at lower cost allowing higher spatial coverage than the current reference methods for air pollutants measurements. Additionally, the monitoring of air pollution with reference measurement methods requires skilled operators for the maintenance and calibration of measuring devices that are described in detailed Standard Operational Procedures [3–7]. Conversely, it is expected that low-cost sensors can be operated without human intervention making it possible for unskilled users to monitor air pollution without the need of additional technical knowledge.

Plenty of institutes in charge of air quality monitoring for regulatory purposes, as well as local authorities, are considering to include low-cost sensors within their routine method of measurements to supplement monitoring with reference measurements. However, the lack of exhaustive and accessible information in order to compare the performance of low-cost sensors and the wide commercial offer make it difficult to select the most appropriate low-cost sensors for monitoring purposes.

For classification and understanding of sensor deployment, one should distinguish between the sole sensor detector produced by Original Equipment Manufacturer (hereafter such sensors are called OEM, or OEM sensors) and sensor systems (SSys), which include OEM sensors together with a protective box, sampling system, power system, electronic hardware, and software for data acquisition, analogue to digital conversion, data treatment and data transfer [8]. Hereafter, OEM and SSys are referred to as low-cost sensors (LCS). From a user point of view, SSys are ready to use out of the box systems, while OEMs needs users to add hardware/software components for protection from meteorological conditions, data storage, data pushing, interoperability of data and generally the calibration of LCSs. The use of LCSs is of major interest for citizen-science initiatives. Therefore, Small and Medium Enterprises make SSys available which can be deployed by citizens who want to monitor the air quality in a chosen environment.

Although a number of reviews of the suitability of LCS for ambient air quality have been published [1,9–15], quantitative data for comparing and evaluating the agreement between LCS and reference data are mostly missing in the existing reviews. Additionally, there is no commonly accepted protocol for the test of LCS [16] and the metrics reported are generally diverse making it difficult to compare the performance of LCS between evaluation studies.

Among the available tests of LCS, there are clear indications that the accuracy of LCS measurements can be questionable [17,18] when comparing LSC values and reference measurements. LCS data can be of variable quality, and it is therefore of fundamental importance to evaluate LCS in order to choose the most appropriate ones for routine measurements or other case studies [19]. However, only a few independent tests are reported in academic publications.

Hereafter, the results of the exhaustive review of existing literature on LCS evaluation that is not available elsewhere are presented. The main purpose of this review was to estimate the agreement between LCS data against reference measurements both with field and tests under controlled conditions carried out by laboratories and research institutes independent from sensor manufacturers and commercial interest. It can provide all stakeholders with exhaustive information for selecting the most appropriate LCS. Quantitative information was gathered from the existing literature about the performance of LCS according to the following criteria:

1. Agreement between LCS and reference measurements
2. Availability of raw data, transparency of data treatment making a-posteriori calibration possible
3. Capability to measure multiple pollutants
4. Affordability of LCSs taking into consideration the number of provided OEMs

2 Sources of available information, method of classification and evaluation

2.1 Origin of data

The research was focused on LCS for Particulate Matter (PM), ozone (O_3), nitric dioxide (NO_2) and carbon monoxide (CO), the pollutants that are included into the European Union Air Quality Directive [2]. References were also included for nitrogen monoxide LCSs.

About 1423 independent laboratory or field tests of LCS versus reference measurements (called 'Records' in the rest of the manuscript) were gathered from peer-reviewed studies of LCS available in the Scopus database, the World Wide Web, the AirMontech website (http://db-airmontech.jrc.ec.europa.eu/search.aspx), ResearchGate, Google search, and reports from research laboratories. Sensor validation studies provided by LCS manufacturers or other sources with concern of a possible conflict of interest were not taken into consideration. Overall, 64 independent studies were found from different sources including reports and peer-reviewed papers.

Additionally, a significant number of test results came from reports published by research institutes. In fact, the rapid technological progress on LCS, the difficulty to publish LCS data that do not agree with reference measurements and the time needed to publish studies in academic journals makes the publication of articles not the preferred route. Consequently, a great part of the available information is found in grey literature, mainly in the form of reports. A substantial quantity of presented results come from research institutes having a LCS testing program in place, e.g. the Air Quality Sensor Performance Evaluation Center (AQ-SPEC) [17], the European Union Joint Research Centre (EU JRC) [18,20–26], and the United States Environmental Protection Agency (US EPA) [14,27–30].

A significant portion of the data comes from the first French field intercomparison exercise (Crunaire [31]) for gas and particle LCS carried out in January/February of 2018. This exercise was carried out by two members of the French Reference Laboratory for Air Quality Monitoring (LCSQA). The objective of the study was to test LCS under field conditions at Air Quality Monitoring Station of urban type sited at the IMT Lille Douai research facilities in Dorignies. A large number of different SSys and OEM were installed in order to evaluate their ability to monitor the main pollutants of interest in the ambient air: NO_2, O_3 and $PM_{2.5}$/PM_{10}. This exercise involved nearly 5 French laboratories in charge of air pollution monitoring and 10 companies (manufacturers or distributors/salers), 23 SSys and OEM of different design and origin (France, Netherlands, United Kingdom, Spain, Italy, Poland, United States), for a total of more than sixty devices, when taking into account replicates.

Within another project, called AirLab (http://www.airlab.solutions/), many LCSs were tested through field and indoor tests. Results are reported based on the Integrated Performance Index (IPI) developed by Fishbain et al. [32] which is an integrated indicator of correlation, bias, failure, source apportionment with LCS, accuracy and time series variability of LCSs and reference measurements. Since the IPI is not available in other studies and cannot be compared with the metrics used in the current review, it was decided not to include the AirLab results in the current work.

A shared database of laboratory and field test results and its associated scripts for summary statistics were created using the collected information. It will be possible to update the database with future results of LCS tests. The purpose of this development was to setup a structured repository to be used for comparing the performances of LCSs.

Each database 'Record' describing laboratory or field LCS test results was included into the database only if comparison against a reference measurement (hereinafter defined as "comparison") was provided. The comparison data allowed to evaluate the correlation between LCS data and reference measurements. Most of the reviewed studies reported only regression parameters obtained from the comparison between LCS and reference measurements, generally without more sophisticated metrics like Root Mean Square Error and measurement uncertainty (see section 3).

2.2 Classification of low-cost sensors

For each model of SSys, the OEM manufacturer was identified and the manufacturer of the SSys as well. Overall, we found 112 models of LCS including both OEMs (31) and SSys (81) manufactured by 77 manufacturers (16 OEM and 61 SSys).

In addition, 19 projects about the evaluation of OEMs and/or SSys reporting quantitative comparison of LCS data and reference measurements were identified. They include the Air Quality Egg, Air Quality Station, AirCasting [17, 27–29], Carnegie Mellon [34,36], CitiSense [28], Cairsense [37], Developer Kit [17], HKEPD/14-02771 [38], making-sense.eu [39], communitysensing.org [30], MacPoll.eu [18], OpenSense II [40,41], Proof of Concept AirSensEUR [21], and SNAQ Heathrow [42,43]). Out of the 1423 Records collected from literature, we identified 1188 Records (197 OEM and 991 SSys) from 89 alive LCS (24 OEM and 65 SSys) and 235 Records (123 OEM and 112 SSys) from 23 "non active" (or discontinued) LCSs (7 OEM and 16 SSys).

"Low-cost" refers to the price of purchase of LCS [9] compared to the purchase and operating cost of reference analysers [44] that can easily exhibit a 10 fold ratio for the monitoring of regulated inorganic pollutants and particulate matter. More recently, ultra-affordable OEMs are starting to appear on the market for PM monitoring [10,45,46]. Many of them are designed to be integrated in Internet of Things (IoT) networks of interconnected devices. Currently, for PM detection it is possible to purchase optical sensors that cost between a few tens and a few hundreds of euro. Those devices are manufactured in emerging economies such as the Republic of China and the Republic of Korea [47]. Some of these LCS can achieve similar performance to more expensive OEMs [17,27-29,32,44–52].

The data treatment of LCSs can be classified in two distinct categories:

1. Processing of LCS data performed by an "open source" software tuned according to several calibration parameters and environmental conditions. All data treatments from data acquisition until the conversion to pollutant concentration levels is known to the user. There were identified 234 Records made of 108 OEMs and 126 SSys using such an open source software for data management. These 401 Records came from 34 unique LCSs. Usually, outputs from these LCS are already in the same measurement units as the reference measurements. In this category, LCS devices are generally connected to a custom-made data acquisition system to acquire LCS raw data. Generally, users are expected to set a calibration function in order to convert LCS raw data to validate against reference measurements.
2. LCS with calibration algorithms whose data treatment is unknown and without the possibility to change any parameter have been identified as "black boxes". This is due to the impossibility for the user to accurately know the whole chain of data treatment. There were identified 1189 Records made up of 212 OEMs and 977 SSys not using an open source software for data treatment. These 1189 Records came from 83 unique LCSs. In most cases, these SSys are previously calibrated against a reference system or, the calibration parameters can be remotely adjusted by the manufacturer. Finally, we should point out that some LCSs used for the detection of Particulate Matter (such as the OPC-N2; OPC-N3 by Alphasense and the PMS series from Plantower) could be used as open source devices if users compute PM mass concentration using the available counts per bins. However, these PM sensors are mostly used as a "black box" with mass concentration computed by unknown algorithms developed by manufacturers.

Clear definitions and examples of the principle of operations used by the different types of sensor (electrochemical, metal oxides, optical particulate counter, optical sensors) are reported in a recent work by WMO [8]. This work also describes several limitations of each type of sensor such as, interference by meteorological parameters, cross-sensitivities to other pollutants, drifts and aging effect. To date, there is a larger number of active and commercially available LCS (Figure 2). However, while most of the OEMs are open sources, allowing end-users to integrate them into SSys, most of the SSys themselves were found to be "black-box" devices. This represents a limitation when the SSys might need a posteriori calibration other then the one provided by the manufacturer since raw-data are unavailable.

LCS are also classified according to their commercial availability. LCSs were assigned to the "Commercial" category if they could be purchased and operated by any user. LCSs fell under the "Non-commercial" category when it was not possible to find any supplier for purchasing. Typically, this type of LCS are used for research and publication while it is difficult for any user to repeat the same sensor setup.

Figure 1. shows the number of LCSs, either OEM or SSys, that were found still active or discontinued, with open or "black box" type of data treatment and that are commercially available.

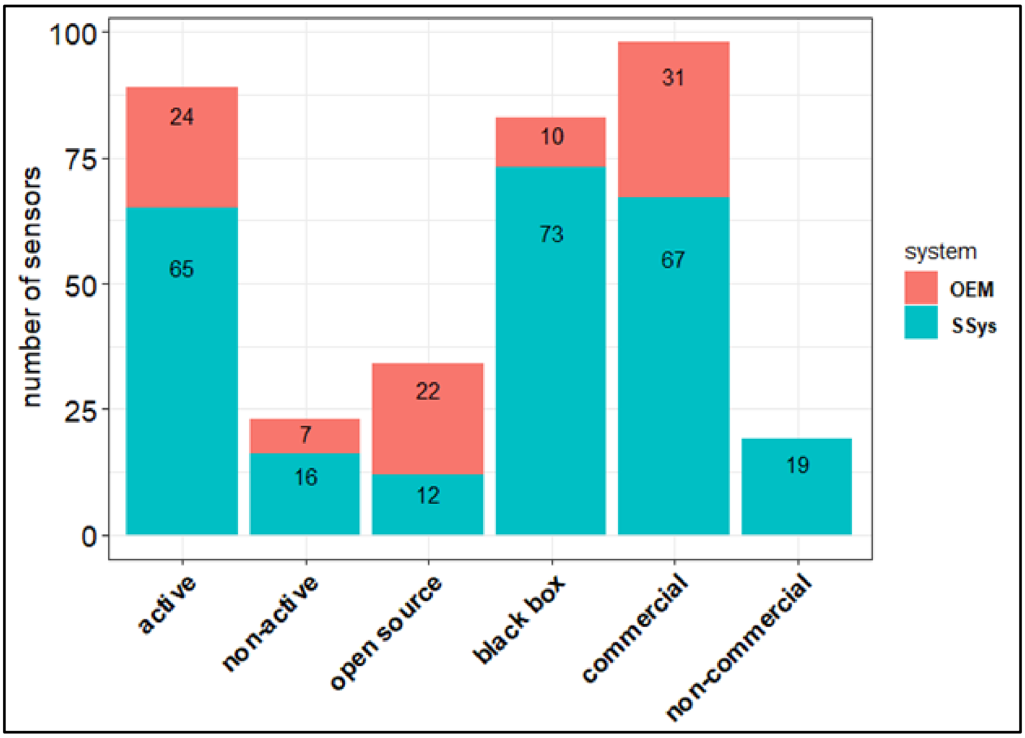

Figure 1. Number of LCS models gathered from the literature review highlighting their open data treatment (open source vs black box) and commercial availability.

2.3 Recent tests per pollutant and per sensor type

Table 1 reports the number of 'Records', by pollutant and sensor technology, gathered in literature about validation and testing of LCSs against a reference system. Records were collected from laboratory (133) and field tests (1290). The majority of records refer to commercially available OEMs and SSys, even though a few references about non-commercial LCS were also picked up.

Table 1. Number of analysed 'Records' for LCSs by pollutant and by type of technology.

pollutant	type	n. Records Field	n. Records Laboratory	references
CO	electrochemical	51	9	AQ-SPEC[17], Jiao[37], Sun[38], Marjovi[56], Karagulian[21], Mead[42], Popoola[43], Borrego[50], Castell[57], Cross[58], Gerboles[22], Wei[59], Gillooly[60], Zimmerman[61], Spinelle[23,26]
CO	MOs	27	2	AQ-SPEC[17], Piedrahita[62], Spinelle[23]
NO	electrochemical	44	6	Jiao[37], Bigi[63], Karagulian[21], Mead[42], Popoola[43], AQ-SPEC[17], Castell[57], Borrego[50], Cross[58], Gillooly[60], Spinelle[24], Gerboles[22], Wei[59], Crunaire[31]

NO	MOs	1	-	Crunaire[31]
NO_2	electrochemical	137	21	AQ-SPEC[17], Jiao[37], Williams[28], Sun[38], Mijling[39], Vaughn[30], Spinelle[18], Mueller[41], Bigi[63], Marjovi[56], Cordero[64], Karagulian[21], Mead[42], Popoola[43], Borrego[50], Castell[57], Cross[58], Spinelle[25], Duvall[65], Gillooly[60], Gerboles[22], Wei[59], Sun[66], Zimmerman[61], Lin[67], Crunaire[31]
NO_2	MOs	28	10	AQ-SPEC[17], Vaugh[68], Williams[28], US-EPA[69], Borrego[50], Piedrahita[62], Spinelle[18], Crunaire[31]
O_3	electrochemical	65	10	AQ-SPEC[17], Jiao[37], Spinelle[18], Mueller[41], Marjovi[56], Karagulian[21], Borrego[50], Castell[57], Cross[58], Duvall[65], Feinberg[34], Gerboles[22], Wei[59], Crunaire[31]
O_3	MOs	54	3	AQ-SPEC[17],Jiao[37], Spinelle[18], Borrego[50], Feinberg[34]
O_3	UV	9	1	Sun[38], AQ-SPEC[17]
$PM_{2.5}$	Electrical	6	-	AQ-SPEC[17]
$PM_{2.5}$	nephelometer	129	24	AQ-SPEC[17], Borghi[33], Jiao[37], Feinberg[34], US-EPA[69], Williams[29], Manikonda[51], Zikova[52], Wang[70], Alvarado[71], Chakrabarti[72], Sousan[53],Borrego[50], Olivares[73],Sun[38], Pillarisetti[74], Holstius[48], Austin[75], Gao[76], Kelly[77], Karagulian[78], Badura[46], Crunaire[31]
$PM_{2.5}$	OPC	428	27	AQ-SPEC[17], Mukherjee[79], Feinberg[34], Jiao[37], Cavaliere[80], Borrego[50], Viana[81], Williams[29], Manikonda[51], Northcross[51], Holstius[48], Steinle[82], Han[83], Jovasevic[84], Dacunto[54], Gillooly[60], Sousan[85], Crilley[86], Badura[46], Kelly[77], Zheng[87], Laquai[45], Budde[10], Liu[88], Crunaire[31]
PM_1	Electrical	6	-	AQ-SPEC[17]
PM_1	nephelometer	1	-	Crunaire[31]
PM_1	OPC	102	8	AQ-SPEC[17], Williams[29], Sousan[85], Crilley[86], Crunaire[31]
PM_{10}	nephelometer	26	1	AQ-SPEC[17], Borrego[50], Alvarado[71], Crunaire[31]
PM_{10}	OPC	176	11	AQ-SPEC[17], Cavaliere[80], Borrego[50], Feinberg[34], Manikonda[51], Sousan[53], Han[83], Jovasevic[84], Williams[29], Sousan[85], Crilley[86], Budde[10], Crunaire[31]

For the detection of Particulate Matter, the largest number of LCS tests were carried out for Optical Particle Counters (OPC) with 752 Records followed by Nephelometers with 181 Records (see Table 1). Both systems detect particulate matter by measuring the light scattered by particles, with the OPC being able to directly count particles according to their size. On the other hand, nephelometers estimate particle density that is subsequently converted into particle mass. For the detection of gaseous pollutants such as CO, NO, NO_2 and O_3, the largest number of tests were performed using electrochemical sensors with 343 Records, followed by Metal Oxides sensors (MOs) with 125 Records (see Table 1). Electrochemical sensors are based on a chemical reaction between gases in the air and the working electrode of an electrochemical cell that is dipped into an electrolyte. In a MOs, also named resistive sensor, semiconductor, gases in the air react on the surface of a semiconductor and exchange electrons modifying its conductance.

Table A2 reports the models of OEMs currently used to monitor Particulate Matter and gaseous pollutants (NO_2, O_3, NO and CO) according to their type of technology. On the other hand, models of SSys measuring concentration of particulate matter and gaseous pollutants are reported in Table A3. We want to point out that several SSys can use the same set of OEMs. In very few cases, the same model of SSys was tested using different types of OEMs when performing validation tests [21,22] .

"Living" LCS are devices that are currently available for commercial or research purposes. Considering only the "living" LCSs of table A2 and Table A3, one may observe that there are less OEMs (24) compared to SSys (65), and therefore different SSys are often based on the same set of OEMs. Additionally, there is a lack of laboratory tests for the OEMs compared to SSys. Among the reviewed 'Records' only ~ 11% were attributed to laboratory tests. Therefore, most LCS (~ 90%) were tested in the field where it is not possible to isolate the effect of single pollutants and/or meteorological parameters, since in the ambient air many of these parameters are correlated with each other. Establishing calibration models relying only on field results might lead to cases where parameters that have no effect on the sensor data but that are correlated with other variables having an effect, are taken into account. The performance of such calibration models can be poor when LCSs are used at other sites than the ones used for calibration where the relationship between the parameter used for calibration and the ones having an effect on the response of LCSs may change [41,89,90].

The research covered the period between 2010 and 2019 (year of publication). As shown in Figure 2, only a few preliminary studies about the evaluation of performance of LCSs were published from 2010 to 2014. In 2015, we recorded the highest number of references with 27 different works publishing results about performances of LCS for air quality monitoring. For the test studies carried out by AQ-SPEC [17], Records were evaluated per model of LCS.

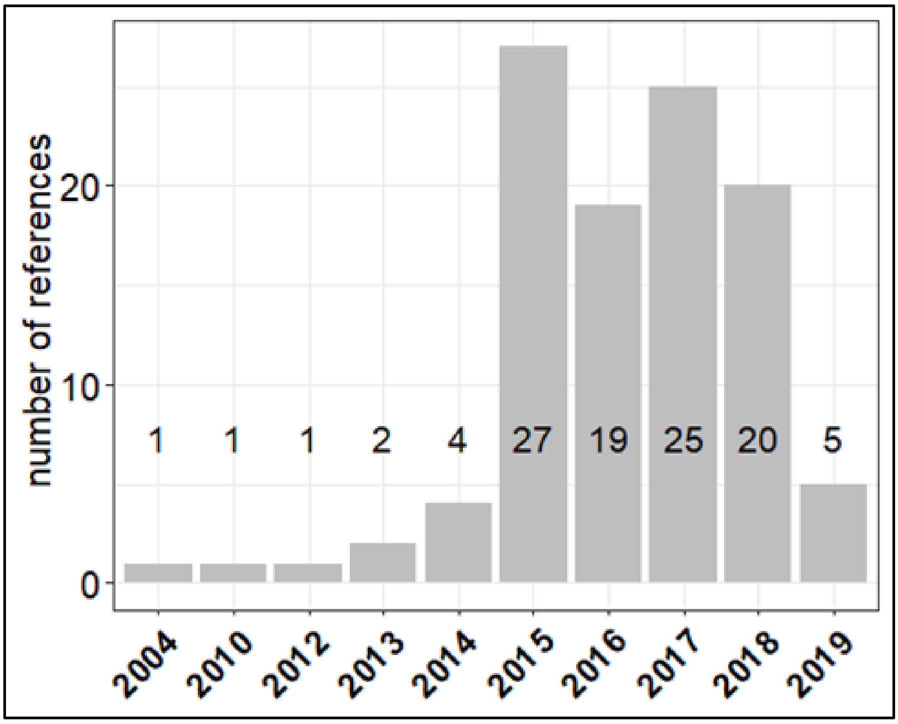

Figure 2. Number of references per year of publication that includes quantitative comparison of sensor data with reference measurements. For 2019, the number publications include the publication between January and April.

Overall, 34 references reporting field tests with LCSs co-located at urban sites were found, as well as 8 references for rural sites, and 10 references for traffic sites. Most of the laboratory and field tests reported hourly data (610 'Records' for 86 models of LCSs). We also found 253 'Records for 40 LCSs using daily data and 248 'Records for 42 LCSs using 5- minute averages (Table A1). Therefore, hourly data were considered statistically more significant.

3 Method of evaluation

The European Union Air Quality Directive [2] indicates that measurement uncertainty [91] shall be the main indicator used for the evaluation of the data quality objective of air pollution measurement methods [2]. However, the evaluation of this metric is cumbersome [92,93] and it is not included in the majority of sensor studies (see Table 2). For the performance criteria used to evaluate air quality modeling applications [94], the set of statistical indicators includes the Root Mean Square Error (RMSE), the bias, the Standard Deviation (SD) and the correlation coefficient (*R*) among which RMSE is thought to be the most explicative one. The statistical indicators can be better visualised in a target diagram [18]. Unfortunately, Table 2 also shows that RMSE is also mainly unavailable in literature. As already mentioned above, integrated indicators like the IPI [32] would breach our objective to use merely quantitative and comparable indicators. Additionally, it is impossible to compute IPIs a posteriori since time series are mainly not available in literature.

Therefore, we had to rely on the most common metrics, i. e., the coefficient of determination R^2, the slope and intercept of linear regression line between LCS data and reference measurement. R^2 can be viewed as a measure of goodness of fit (how close evaluation data is to the reference measurements) and the slope of the regression as level of accuracy. R^2 measures the strength of the association between two variables but it is insensitive to bias between LCS and reference data, either relative bias (slope different from 1) or absolute bias (intercept different from 0). R^2 is a partial measure of how much LCS data agree with reference measurements according to a regression model [95]. A larger R^2 reflects an increase in the predictive precision of the regression model.

An increase of R^2 may not be the result of an improvement of LCS data quality since R^2 may increase when the range of reference measurements increases [96] or according to the seasonality of sampling reported in different studies. Moreover, since LCS are affected by long time drift and ageing, longer field studies are more likely to report lower R^2 than shorter one.

Nearly all published studies report the coefficient of determination (R^2) between reference and LCS data (see Table 2). Fortunately, the majority of these studies also reports the slope and intercept of the regression line between LCS data and reference measurements that describe the possible bias of LCS data. A few studies also report the Root Mean Square of Error, RMSE [18,21,34,39–41,48,57,58,60,62,64,87,88] which clearly indicates the magnitude of the error in LCS data unit and that is sensitive to extreme values and outliers. Only a few studies report the measurement uncertainty [21,24,28,45,49,57,59,61]. Therefore, for the purpose of this work, we only focused on the analysis of the comparison of laboratory and field tests of LCSs.

Table 2. Number of Records gathered by metrics available in literature.

metrics	n. Field Tests	n. Laboratory Tests
Total tests	1290	133
R^2, calibrations	218	60
R^2, comparisons	1160	72
slope of reg. line	1063	55
intercept	1027	54
RMSE	285	5
Measurement uncertainty (U)	153	29

Table 2 also gives the number of R^2 of calibration that was found in literature. Generally, these studies also present the model equations used for calibration. The number of studies reporting the R^2 of calibration represent about 10 % of the studies reporting R^2 of comparison of calibrated LCSs and reference data using linear regressions.

Although the data set of R^2 for calibration is limited in size, we have investigated if the type of calibration has an influence on the agreement between calibrated LCSs data and reference measurements.

In order to estimate the efficiency of calibration models, it was reported the coefficient of determination R^2 as an indicator of the amount of total variability explained by the model (see Calibration of LCSs). This can be used as an indication of performance of the calibration model chosen to validate the LCS against a reference system.

Using the highest R^2 of comparison together with the slope of comparison line near to 1.0, a shorter set of best performing LCS will be drawn together with their sensor technology. It was decided to drop the analysis of intercepts different from 0, accepting that LCS may produce unscaled data with bias provided that LCS data would vary in the same range as reference measurements due to the slope close to 1. Anyhow, Table 4 and Table A4 show that the extent of deviation from 0 of the intercepts did not contribute significantly to the bias of LCS data for the best performing LCSs.

4 Evaluation of sensor data quality

4.1 Calibration of sensors

The method used for the calibration of LCS is generally considered confidential information by the majority of LCS manufacturers. In fact, little information can be found about the calibration of LCS that fall under the category "black box" compared to the ones that fall under the category "Open source". In fact, several studies can be found about the calibration of "Open source" LCSs, both with laboratory and field tests. Calibration consists of setting a mathematical model describing the relationship between LCS data and reference measurements. However, most of the calibrations were carried out during field tests, while only a limited number of laboratory experiments were found available.

Out of a total of 1423 records in the database, 352 Records (25%) included information about LCS calibration giving details of used statistical or deterministic models (see Table 3). However, among these 352 records with details on calibration method, about 20 % do not report R^2, that is the principal metrics used for LCS performance evaluation. This is typically the case for Artificial Neural Networks, Random Forest and support vector regression calibration methods (see below) and it explains why the number of R^2 found for calibration in Table 2 is lower than 352.

The linear model and the multi-linear regression model (MLR) which includes the use of covariates to improve the quality of the calibration are the most widely used techniques to calibrate the LCS data against a reference measurement. Other calibration approaches used the exponential, logarithmic, quadratic, Kohler theory of particles growing factor and few types of supervised learning techniques including Artificial Neural Networks (ANN), Random Forest (RF:), Support Vector Machine (SVM:) and support vector regression (SVR). Most of MLR models used covariates such as meteorological parameters (temperature and relative humidity) and cross-sensitivities from gaseous interferent such as nitric dioxide (NO_2), nitric monoxide (NO) and ozone (O_3) in order to improve LCS calibration. Rarely, LCS data time-drift was included into the list of calibration covariates [37,97].

Table 3. Types of calibration models used for the calibration of LCS. (ANN: artificial neural network, exp: exponential; log: logarithmic; MLR: multilinear regression; quad: quadratic; RF: random forest; SVM: support vector machine; SVR: support vector regression; Kohler: particle size distribution–based correction algorithm accounting for the influence of RH).

Pollutant	Calibration model	n. Records	References	Median R^2 calibration	Median R^2 comparison
CO	ANN	2	Wastine[98], Spinelle[23]	-	0.58
CO	linear	12	Sun[38], Wastine[98], Castell[57], Cross[58], Gerboles[22], Spinelle[23], Zimmerman[61]	0.85	0.15
CO	MLR	21	Jiao[37], Karagulian[21], Wastine[98], Wei[59], Piedrahita[62], Spinelle[23], Zimmerman[61]	0.89	0.83
CO	quad	12	AQ-SPEC[17]	0.63	-

CO	RF	1	Zimmerman[61]	0.91	-
NO	ANN	2	Wastine[98], Spinelle[23]	-	0.57
NO	linear	8	Wastine[98], Castell[57], Cross[58], Spinelle[23], Karagulian[21], Crunaire[31]	0.96	0.032
NO	MLR	20	Jiao[37], Bigi[63], Karagulian[21], Wastine[98], Spinelle[23], Wei[59]	0.92	0.91
NO	RF	2	Bigi[63]	-	0.9
NO	SVR	2	Bigi[63]	-	0.90
NO_2	ANN	7	Cordero[64], Spinelle[18], Wastine[98], Wastine[99]	0.87	0.94
NO_2	linear	25	Sun[38], Spinelle[18], Wastine[98], Wastine[99], Castell[57], Cross[58], Karagulian[21], Zimmerman[61], Lin[67], Crunaire[31]	0.25	0.17
NO_2	log	1	Vaughn[30]	0.89	-
NO_2	MLR	48	Jiao[37], Sun[66], Mijling[39], Spinelle[18], Mueller[41], Bigi[63], Cordero[64], Karagulian[21], Wastine[98], Wastine[99], Piedrahita[62], Wei[59], Zimmerman[61]	0.81	0.81
NO_2	quad	6	AQ-SPEC[17]	0.61	-
NO_2	RF	7	Bigi[63], Cordero[64], Zimmerman[61]	0.86	0.91
NO_2	SVM	4	Cordero[64]	0.85	0.94
NO_2	SVR	2	Bigi[63]	-	0.78
O_3	ANN	2	Spinelle[18], Wastine[98]	-	0.89
O_3	linear	13	Sun[38], Spinelle[18], Wastine[98], Castell[57], Cross[58], Karagulian[21], AQ-SPEC[17], Crunaire[31]	0.84	0.53
O_3	log	1	Vaughn[30]	0.88	-
O_3	MLR	20	Jiao[37], Spinelle[18], Karagulian[21], Wastine[98], Spinelle[24], Wei[59]	0.91	0.88
O_3	quad	9	AQ-SPEC[17]	0.72	-
PM1	Kholer	2	Di-Antonio[100]	-	0.74
PM1	log	6	AQ-SPEC[17]	0.76	-
PM10	exp	6	AQ-SPEC[17]	0.59	-

PM10	linear	3	Cavaliere[80], Jovanovic[84], AQ-SPEC[17]	0.77	0.73
PM10	log	7	AQ-SPEC[17]	0.58	-
PM10	quad	1	Alvarado[68]	0.65	-
PM10-2.5	linear	4	Sousan[54], Han[81], Jovasevic[84]	0.63	0.98
$PM_{2.5}$	exp	3	Dacunto[54], Kelly[74], Austin[72]	0.91	0.97
$PM_{2.5}$	Kholer	2	Crilley[86], Di-Antoni[100]	-	0.78
$PM_{2.5}$	linear	37	Mukherjee[79], Wang[67], Alvarado[68], Cavaliere[80], Jovasevic[84], Olivares[53], Kelly[74], Zheng[17], Holstius[18]	0.84	0.64
$PM_{2.5}$	log	7	AQ-SPEC[17], Laquai[40]	0.73	-
$PM_{2.5}$	MLR	17	Jiao[37], Sun[66], Zheng[83], Holstius[18], Liu[19]	0.81	0.65
$PM_{2.5}$	quad	8	Chakrabarti[69], Alvarado[68], Zheng[17], Gao[76]	0.87	0.88
$PM_{2.5}$	RF	3	Liu[19]	-	0.79
$PM_{2.5\text{-}0.5}$	linear	9	Northcross[87], Steinle[80], Han[81], Jovasevic[84]	0.84	0.98
$PM_{2.5\text{-}0.5}$	MLR	1	Jiao[37]	0.6	0.45
$PM_{2.5\text{-}0.5}$	quad	6	AQ-SPEC[17], Manikonda[51]	0.82	-

When R^2 is both available for calibration and comparison, the median of R^2 is higher for calibration (mean of R^2 = 0.70) than for comparison (median of R^2 = 0.58). This is an expected results since it is easier to fit a model on a short calibration dataset than correctly forecast LCSs data using the calibration model at later dates. For gaseous LCSs, calibration using a linear model gives the worth R^2 of field comparison. Linear calibration should be avoided for gas LCSs.

For CO and NO, the calibration method giving the highest R^2 (about 0.90) is the MLR method using temperature or relative humidity as covariates. The use of supervised learning techniques (ANN, RF or SVR) did not improve performance for CO or gave similar results than MLR for NO. This is in slight contradiction with other studies about the performance of supervised techniques [101,102]. In the majority of cases, these tested LCSs consisted of electrochemical sensors.

For NO_2, supervised learning techniques (ANN, RF, SVM or SVR) performed slightly better than MLRs looking at the R^2 of comparison tests in field, except for SVR which is in slight contradiction with other studies [102]. However, the number of records is much higher MLR than for supervised learning techniques. MLR was applied to both MOs sensor and electrochemical sensors which resulted in scattered R^2 when looking at individual studies. Additionally, supervised learning techniques may be more sensitive to re-location than MLR [89,90].

For O_3, ANN and MLR calibration gave similar R^2 of comparison (median value about 0.90). As for NO_2, the higher number of studies makes the R^2 of the MLR method more significant than the one of ANN.

For PM, the R^2 for comparison tests are very scattered over the calibration methods. Some high values (R^2 higher than 0.95) were reported for studies using a linear calibration while MLR did not perform well ($R^2 < 0.5$). This results are misleading, since the good results with linear calibration are generally obtained by discarding LSCs data obtained with relative humidity exceeding a threshold between 70 and 80% for which humidity is responsible for particle growing [86,100]. This effect is more important for PM_{10} than for PM_1 and $PM_{2.5}$. Other studies did not discard high relative humidity. They took into consideration the particle growing factor either on mass concentration with an exponential calibration model ([54,75,77]) with a median R^2 of 98 or using the Kölher theory on PM mass concentration [83] or directly for the particles beans of each OPC bin [100] leading to R^2 about 0.80.

Figure 3 shows a summary of all mean R^2 obtained from the calibration of SSys against reference measurements. Results were grouped by model of SSys and averaged per reference work. For the same SSys we can observe R^2 ranging between 0.40 and 1.00. This shows the variability of the performance of SSys depending on the type of calibration, type of testing sites and seasonality and making it difficult to compare the results of different studies.

Calibration of LSC against a reference analyser was found to be carried out using different averaging time. Test results with hourly data are presented in Figure A1 and test results with minute data time are given in Figure A2.

The best performance, according to the time average availability in literature and tests in laboratory and/or in the field, were found for:

- For the measurement of $PM_{2.5}$, values of R^2 close to 1 were found for hourly data of PMS1003 by Plantower [77] and for the the PMS3003 , Dylos DC1100 PRO and DC1700 by Dylos for minute data [14,17,82]. Strangely, higher R^2 were reported for the Plantower and Dylos when calibrated with minute data than for hourly data. The OPC-N2 by AlphaSense [17] reported values of R^2 falling within the range of 0.7 - 1.0. The same OPC-N2, reported values of R^2 just above 0.7 when measuring PM_1 while it did not show a good performance when measuring PM_{10} [17] (R^2 less than 0.5). We need to stress out that optical sensors, such as OPCs and nephelometers, are somewhat limited to fight gravity when detecting coarse particulate matter because of the low-efficiency of the sampling system. Most of the regression models used for the calibration of LCSs used hourly data.
- For the calibration of O_3 LCS the highest values of R^2 for hourly data was reported for FIS SP-61 by FIS and O3-3E1F by CityTechnology (Figure A1) [18]. On the other hand, for minute data, values of R^2 close to 1 were found for AirSensEUR (V.2) by LiberaIntentio [21] as well as for the S-500 by Aeroqual [17] (Figure A2). AirSensEUR used a built-in AlphaSense OX-A431 OEM. We want to point out that, most of the MLR models used for calibrating O_3 LCSs needs NO_2 to correct for the strong NO_2 cross-sensitivity.
- For the calibration of NO_2 LCSs, we found values of R^2 for hourly data within the range 0.7 - 1.0 for the NO2-B42F (by Alphasense [59]), for the AirSensEUR (v.2) by LiberaIntentio [21] and for the minute values of MAS [38] (see Figure 3). The NO_2 measurement of the AirSensEUR (v.2) are carried out using the NO2-B43F OEM by AlphaSense.
- Most of the Records about the calibration of CO LCSs showed high values of R^2. As shown in Figure A1, the OEMs CO 3E300 by City Technology [22] and CO-B4 by Alphasense [59] reported $R^2 \sim 1$ for hourly data. High values of R^2 were also reported for the SSys AirSensEUR (v.2) when calibrating CO minute data [21] (Figure A2). Other LCSs reporting values of R^2 within the range 0.7 - 1.0 for hourly data consisted of the MICS-4515 by SGX Sensortech [62], the Smart Citizen Kit by Acrobotic [17] and the RAMP [61].

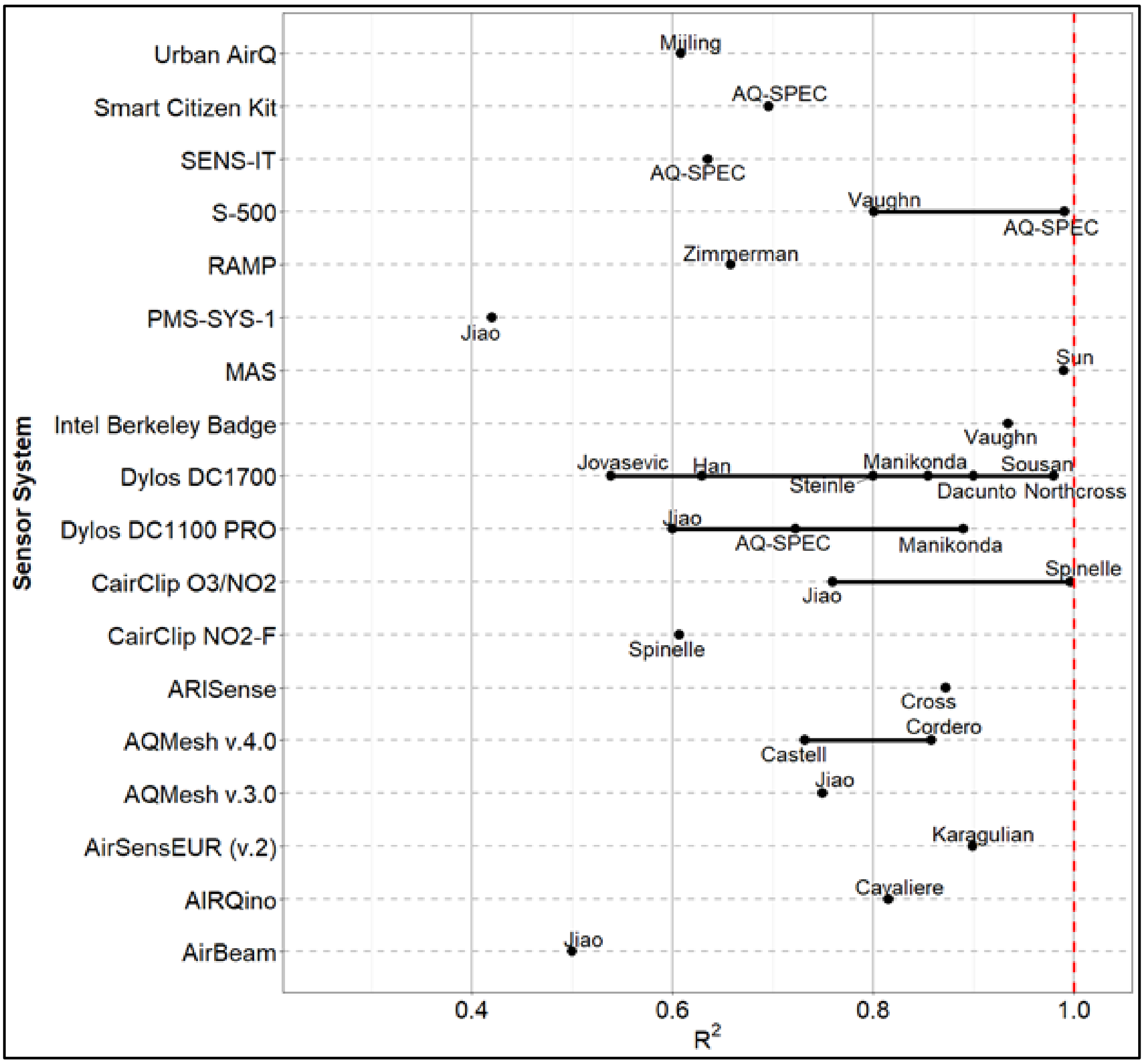

Figure 3. Mean R^2 found for LCSs reporting calibration against reference measurements for all used calibration method. The author name below each bullet gives the 1st author of the publication from which results were drawn.

4.2 Comparison of calibrated low-cost sensors with reference measurements

In this review, Records describing the comparison of LCS data with reference measurements came from “open source” and “black box” LCSs. As for the Records collected from the calibration of LCS, comparison with reference system was carried out at different time-resolutions. Here we only report comparisons of hourly data with 565 and 151 Records from SSys and OEMs, respectively. In Figure 4 we have reported the R^2 values for SSys per reference averaged for all pollutants measured by each SSys. One can observe scattered of R^2 for a few SSys that are tested in several references in different locations, seasons and durations.

Figure A3 and Figure A4 show the distribution of R^2 of LCSs hourly and minute values measuring PM_{10}, $PM_{2.5}$, PM_1, O_3, NO, NO_2, and CO against reference measurements:

- For the SSys, the PA-II by PurpleAir [17] and PATS+ by Berkley Air [74] showed the highest R^2 with values between 0.8 and 1.0. Other LCSs with R^2 values ranging between 0.7-1.0 included the PMS-SYS-1 by Shinyei, the Dylos 1100 PRO by Dylos, the MicroPEM by RTI, the AirNUT by Moji China, the Egg (2018) by Air Quality Egg, the AQT410 v.1.15 by Vaisala, the AirVeraCity by AirVeraCity, the NPM2 by MetOne [31] and the Air Quality Station by AS LUNG [17]. We need to point out that the performance of LCSs measuring PM_{10}, on average, was very poor.

- For the hourly PM measurements of OEMs (Figure A5), the OPC-N2, OPC-N3 [23,40,42,76,83] and the SDS011 by Nova Fitness [76] showed R^2 values in the range 0.7 - 1.0. For the 24-hour PM measurements of OEMs (Figure A6), we found R^2 within the range 0.7 - 1.0 for the OPC-N2 and the OPC-N3 [17].
- For the 24-hour PM measurements of SSys (Figure A7), the PA-II, [17] AirQUINO by CNR [77] showed R^2 values close to 1 for $PM_{2.5}$
- For gaseous pollutants, high R^2 ranging between 0.7 and 1.0 were found for the following multipollutant LCSs: AirSensEUR by LiberaIntentio [21], the AirVeraCity, the AQY and S-500 by Aeroqual and the SNAQ of the University of Cambridge (Figure A3).
- For the hourly gaseous measurements (Figure A5), we found very few OEMs with R^2 in the range 0.7 - 1.0. These included the CairClip O3/NO2 by CairPol [19,24,31,62], the Aeroqual Series 500 (and SM50) [31], the O3-3E1F by CityTechnology [23,24,31] and the NO2-B43F by Alphasense [58,63]. On the other hand, we found very few Records for SSys using daily data. Additionally, one can notice, comparing Figure A4 and Figure A5, that the performance of OEMs is generally enhanced when they are integrated inside SSys, except for PM_{10}.

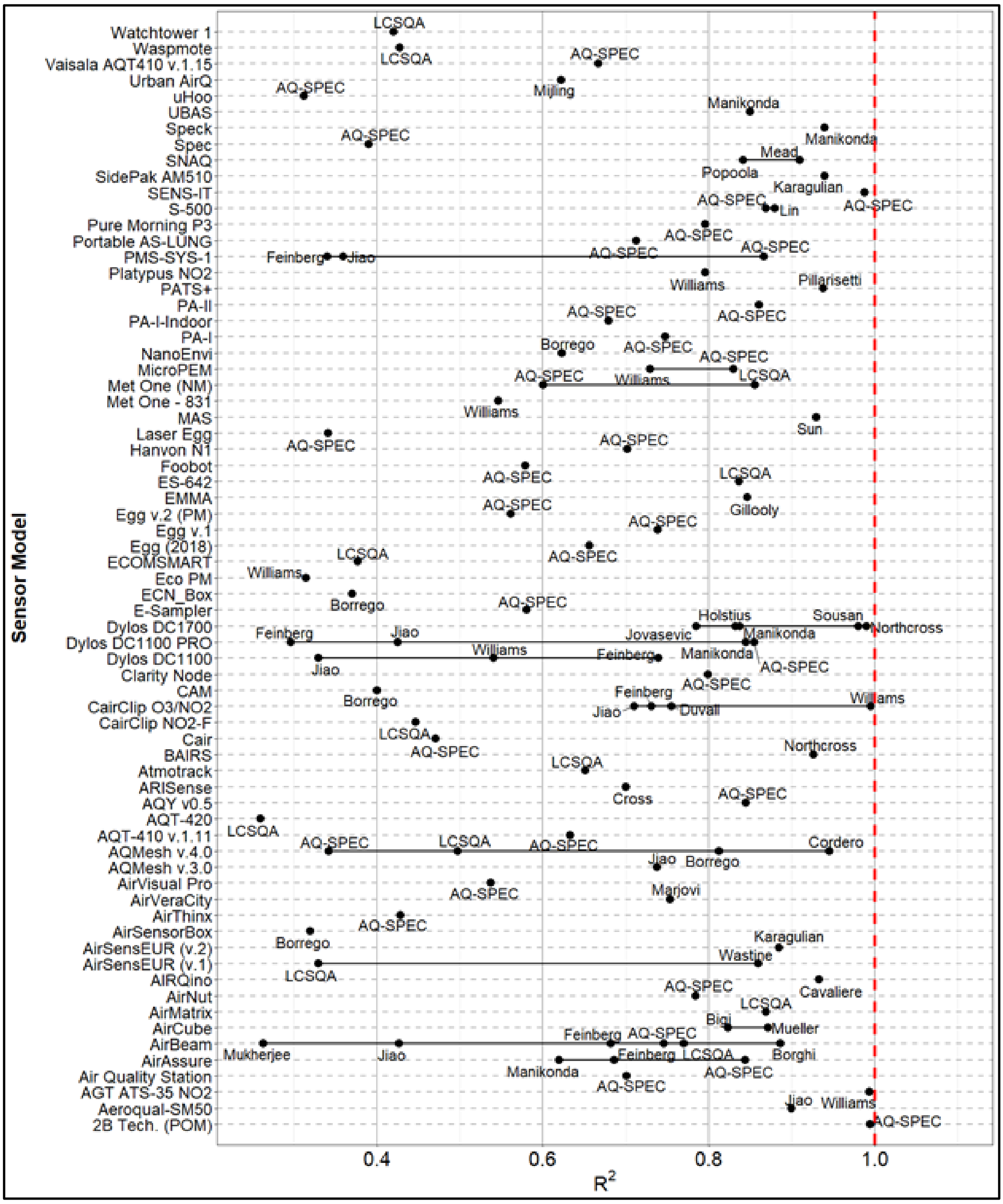

Figure 4. Mean R^2 for obtained from the comparison of SSys against reference measurements at all averaging times (1 min, 5 min, 1 hour and 24h). The author name below each bullet gives the 1st author of the publication from which results were drawn.

Figure A8 and Figure A10 show selected SSys that gave a slope of linear regression line of hourly LCS data versus reference measurement from 0.5 to 1.5 and R^2 higher than 0.7. This selection includes the AirSensEUR, the AirVeracity, and the S-500 for gaseous pollutants and the AirNut, AQY v0.5, Egg v.2 (PM), the NPM2 for hourly data and AIRQuino, AQY v0.5, Egg v.2 (PM) and the PA-I for daily data.

Figure A9 and Figure A11 show the same selection as Figure A8 but for OEMs. This list includes the SM50, the CairClip O3/NO2, the S-500 (NO_2, O_3), the NO2-B4F (NO_2) for gaseous measurements and the Nova Fitness SDS011 for $PM_{2.5}$ measurements for hourly data and the OPC-N2 by Alphasense and the DataRAM for daily data.

5 Cost of purchase

For the evaluation of the price of LCSs, we considered all SSys manufactured by commercial companies. Operating costs such as calibration, maintenance, deployment and data treatment are not included in the estimated price of SSys.

Figure 5 shows the commercial price of LCSs by model and number of measured pollutants and Figure A13 shows the prices for OEMs. There is a large number of SSys measuring one pollutant and only a few ones measuring multiple pollutants. Most OEMs are open source devices (Figure A13). On the other hand, most of the SSys are "black boxes" (Figure 5). Therefore, most of the SSys cannot be easily re-calibrated by users. In fact, most SSys are intended to be ready-to-use air quality monitors.

In Figure 6 we have shortlisted the best SSys according to their level of agreement with reference systems. Figure 6 includes SSys with hourly and daily data showing R^2 higher than 0.85 and slopes ranging between 0.8 and 1.2. The Figure shows the price, the number of pollutants being measured, the averaging time and the data openness of the selected SSys. Table 4 reports the SSys shortlisted in Figure 6 with the R^2 and slope mean values, the list of pollutants being measured, the openness of data, their commercial availability and price.

Among "open source" SSys, we could identify the AirSensEUR by LiberaIntentio and the AIRQuino by the CNR. The remaining shortlisted SSys were identified as "black box". The AirSensEUR (v.2) resulted in a mean R^2 value of 0.90 and a slope of 0.94 while the AIRQuino resulted in a mean R^2 value of 0.91 and a slope of 0.97. We need to point out that, to date, the AIRQuino can be used for the detection of up to five pollutants (NO_2, CO, O_3, NO, and PM). However, only data for PM were available at the time of this review.

Table 4. Shortlist of SSys showing good agreement with reference systems ($R^2 > 0.85$; $0.8 < slope < 1.2$) for 1 hour time averaged data.

model	pollutant	mean R^2	mean slope	mean absolute intercept	open/close	living	commercial	price (EUR)
AirNut	$PM_{2.5}$	0.86	0.88	8.6	black box	Y	commercial	132
PA-I	PM_1	0.95	0.92	0.52	black box	N	commercial	132
PA-II	PM_1	0.99	0.82	1.8	black box	Y	commercial	176
Egg (2018)	PM_1	0.87	0.85	0.095	black box	Y	commercial	219
PATS+	$PM_{2.5}$	0.96	0.92	0.05	black box	Y	commercial	440
S-500	NO_2, O_3	0.88	0.97	0.27	black box	Y	commercial	440
CairClip O3/NO2	O_3	0.88	0.88	12	black box	Y	commercial	600
Portable AS-LUNG	PM_1	0.89	0.87	1.0	black box	Y	non commercial	880
AirSensEUR (v.1)	NO_2, O_3, CO, NO	0.95	0.98	-	open source	Y	commercial	1600
AirSensEUR (v.2)	NO_2, O_3, CO, NO	0.89	1.1	5.7	open source	Y	commercial	1600
Met One (NM)	$PM_{2.5}$	0.86	1.1	2.8	black box	Y	commercial	1672
Air Quality Station	PM_1	0.88	0.90	0.85	black box	Y	non commercial	1760
AQY v0.5	$PM_{2.5}$	0.87	0.97	4.0	black box	updated	commercial	2640
Vaisala AQT410 v.1.15	CO	0.87	0.97	0.23	black box	Y	commercial	3256
2B Tech. (POM)*	O_3	1.00	1.00	0.74	black box	Y	commercial	3960
AQMesh v.3.0	NO	0.87	0.88	0.76	black box	N	commercial	8800

** The 2B Tech. (POM) is a miniaturized reference method, UV photometry, which explains the perfect R^2 and slope of 1*

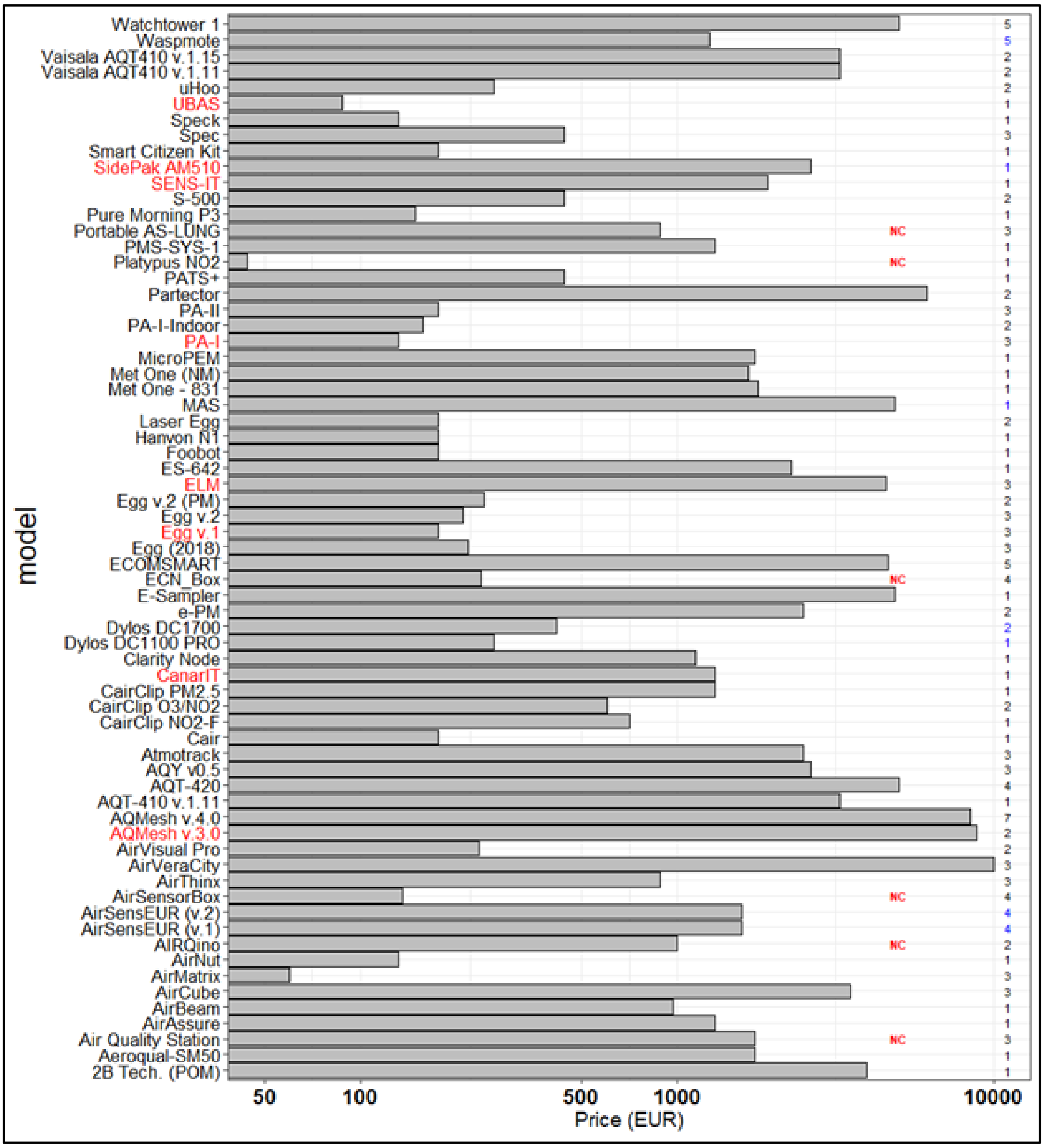

Figure 5. Prices of SSys grouped by model. Numbers at right indicates the number of pollutants measured by each SSys, with open source in blue and black box in black. x-axis uses logarithmic scale. Names of ‘living’ and ‘non-living’ SSys are indicated in black and red color on the labels of the y-axis, respectively. NC indicates commercially unavailable sensors.

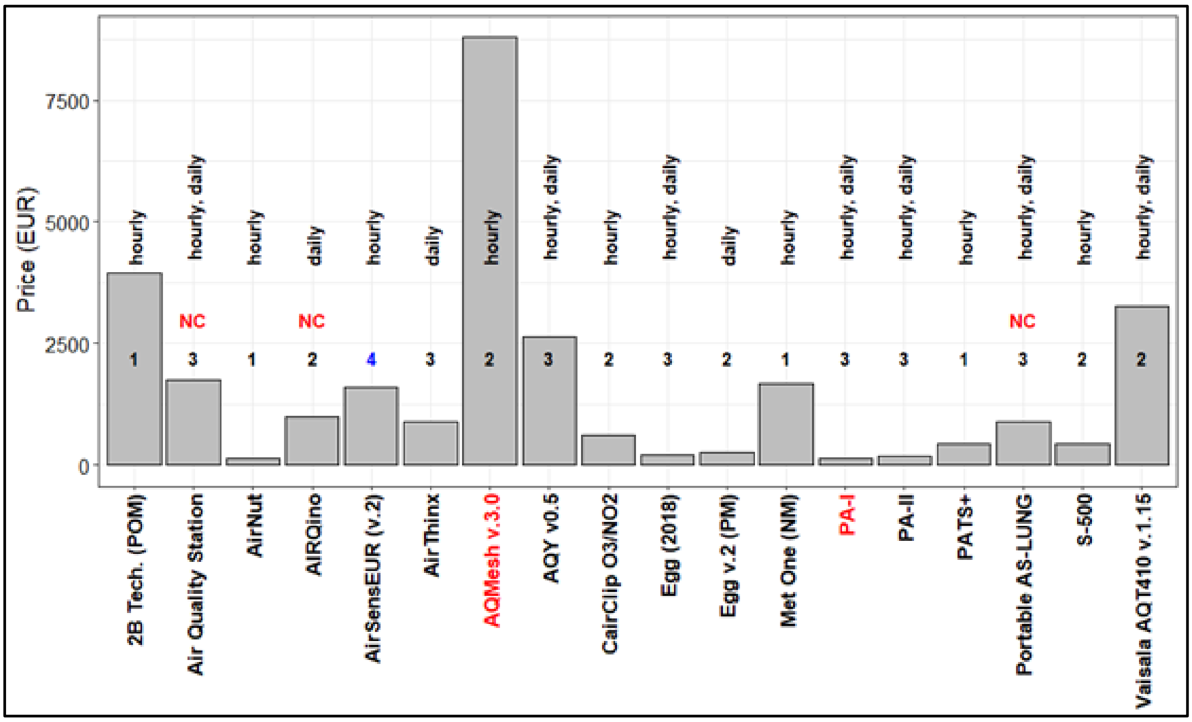

Figure 6. Price of low-cost SSys. Numbers in bold indicate the number of pollutant measured by open source (blue) and black box (black) sensors. Only Records with $R^2 > 0.85$ and $0.8 < \text{slope} < 1.2$ are shown. Names of 'living' & 'updated' and 'non-living' sensors are indicated in black and red on the labels of x-axis, respectively. NC indicates commercially unavailable sensor. Labels reporting hourly/daily indicate the averaging time of reviewed records.

Figure 7 shows the relationship between the mean R^2 of SSys and the decimal logarithm of the price of LCSs. In Figure 7 only the "living" LCSs are compared. It shows that for OEMs there is not a significant linear relationship between the price of OEMs and the value of R^2. Conversely, there is a significant light increase of R^2 with the logarithm of the price of SSys. The regression equations indicated in Figure 7 shows that R^2 can increase of 14 ± 6% for a 10-fold increase of the prices of SSys which is a limited increase at high cost. Figure 7 also shows a higher scattering of R^2 at the low end of the price scale with SSys price lower than 500 euro with more fluctuation of the SSys performance.

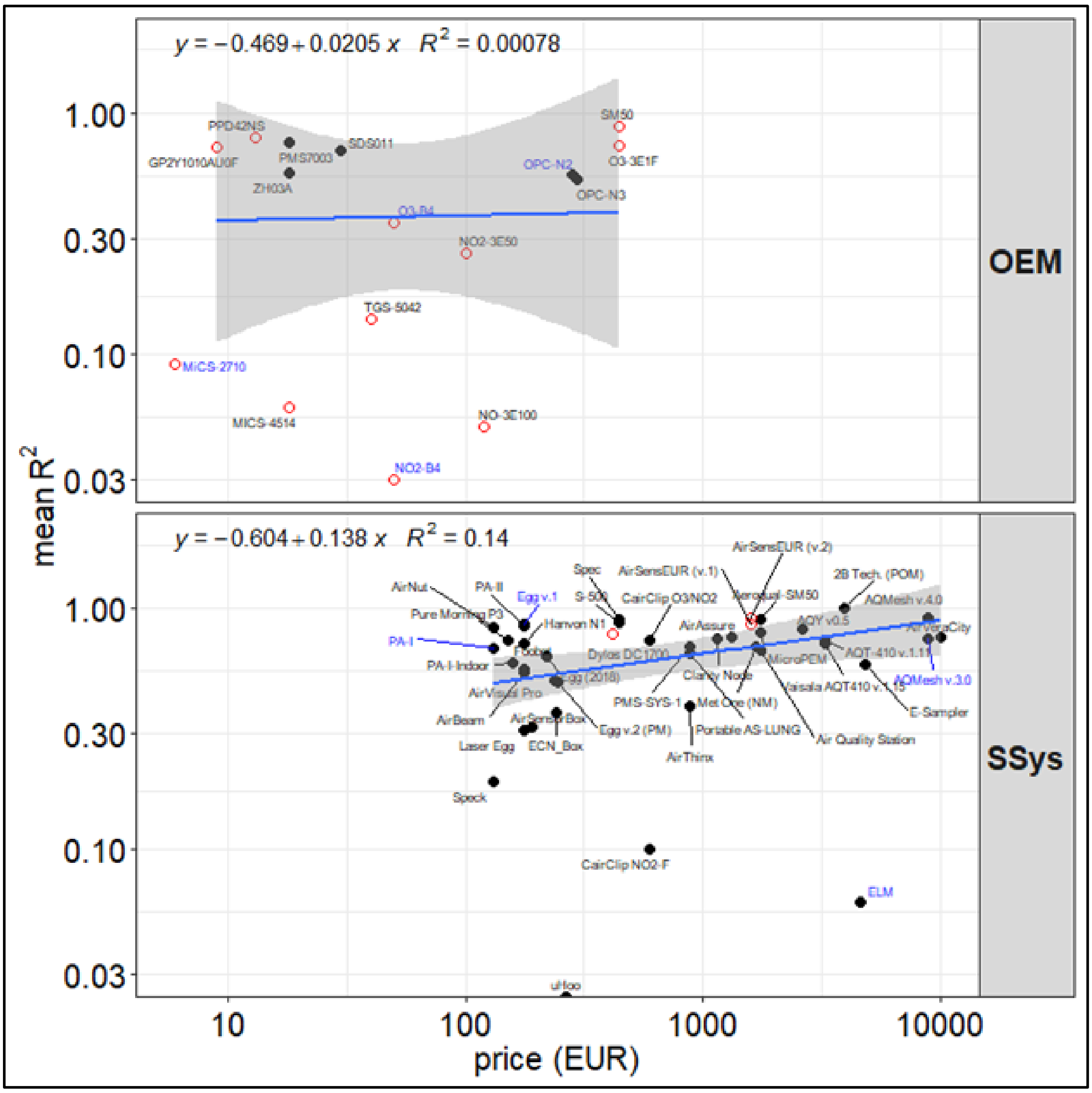

Figure 7. Relation between prices of LCSs and R^2 for field test only. Logarithmic scale has been set for both axis. Open source and black box models are indicated with red open dots and black solid dots, respectively. Names of 'living' and 'non-living' sensors are indicated in black and blue color, respectively. R^2 refers to data averaged over 1 hour. Grey shade in the fit plots indicate a pointwise 95% confidence interval on the fitted values.

6 Conclusions

There is little information available in the literature regarding calibration of LCSs. Nevertheless, it was anyhow possible to list the calibration methods giving the highest R^2 when applied to the results of field tests. For CO and NO reviewed works showed that the MLR models were the most suitable for calibration. ANN gave the same level of performance than MLR only for NO. For NO_2 and O_3, supervised learning models such as, SVR, SVM, (not for O_3), ANN, and RF followed by MLR models showed to be the most suitable method of calibration. Regarding Particulate Matter, the best results were obtained with linear models when calibrating $PM_{2.5}$. However, these models were applied only to $PM_{2.5}$ with high relative humidity data (> 75-80%) that were discarded. For higher relative humidity, models accounting for the growing of the particulates must be further developed. So far, the calibration using the Khöler theory seems to be the promising method.

A list of SSys with R^2 and slope close to 1.0 were drawn from the whole database of Records of comparison tests of LCSs data versus reference measurements that indicates the best performance of SSys as shown in Figure 8. In fact, Figure 8 evidences a best selection region for SSys with blue background. The best SSys would be the one which reaches the point with coordinates R^2 = 1 and slope = 1. Within the blue background region, the following SSys can be found: the 2B Tech. (POM), the PA-II, the AirSensEUR (v.1), the PA-I, the S-500, the AirSensEUR (v.1), the SNAQ, the Vaisala AQT410 v.15, the MetOne (NM), the Egg (v.2), the AQY v0.5, the CairClip O3/NO2, the AQMesh v3.0, the AQT410 v.11 and the AirVeraCity. Additionally, Figure 8 shows that there are more SSys underestimating reference measurements with slopes lower than 1 than SSys overestimating reference measurements.

Analysing the price of SSys and their price, it was found that R^2 increases of 14 % for a 10-fold increase of the prices of SSys, a limited improvement for a large price increase.

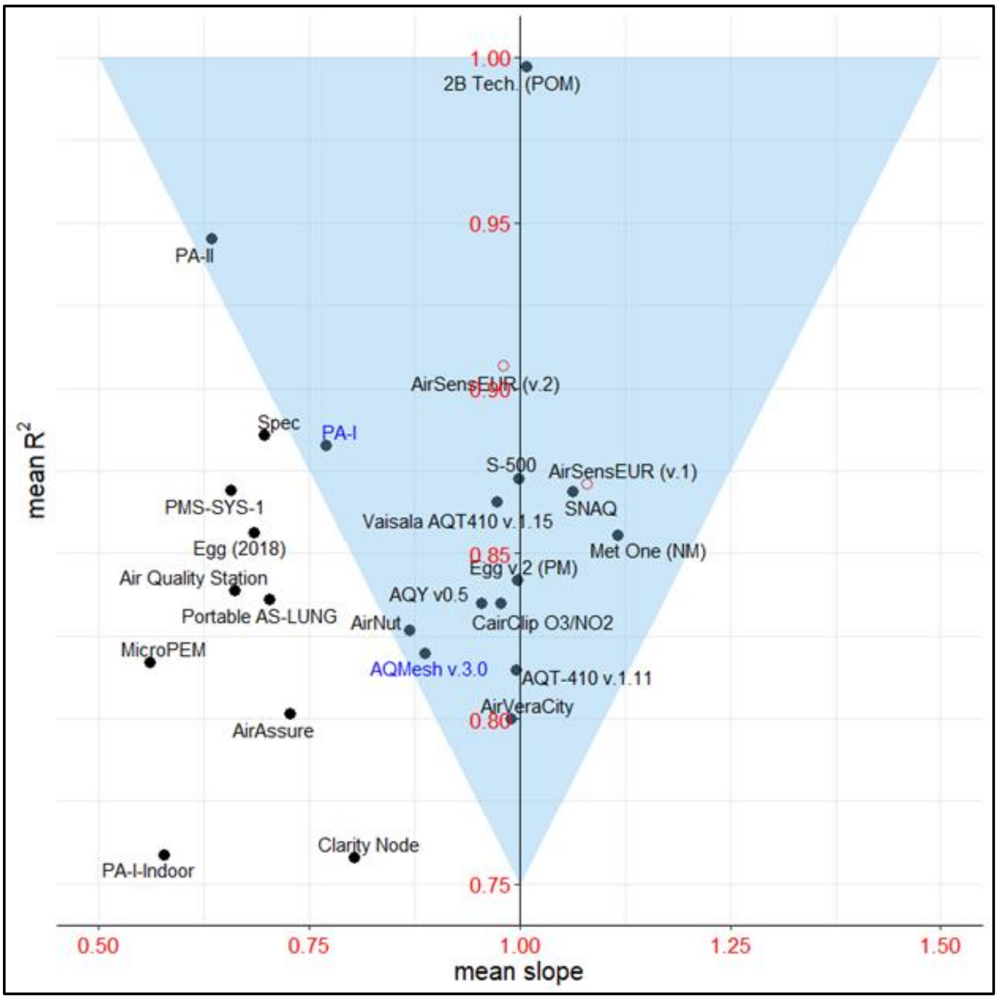

Figure 8. Correspondence between R^2 and *slope* for SSys. Only SSys models with $R^2 > 0.75$ and $0.5 < slope < 1.2$ are shown. Names of 'living' and 'non-living' sensors are indicated in black and blue color, respectively.

Although this paper gives an exhaustive survey of the independent LCS evaluations, the concept of comparing LCS field tests from different studies can be difficult or results in misleading conclusions. It is difficult because of the lack of uniformity of the metrics representing LCS data quality that are different between studies and difficult to compare. Comparing field tests of LCS may also be misleading since in order to take into consideration the highest number of studies it was necessary to mainly rely on the coefficient of determination R^2. However, R^2 is too much dependent on the range of reference measurements, on the duration of test field and on the season and location of the tests making change of R^2 not completely dependent on LCS data quality or of calibration methods. This shortcoming makes the standardisation of a protocol of evaluation of LCSs at international level of great importance and the results of intercomparison exercise where gathering LCSs are gathered at the same test sites and at the same time greatly desired.

Author Contributions: conceptualization, Federico Karagulian and Michel Gerboles; methodology, Federico Karagulian and Michel Gerboles; software, Federico Karagulian; validation, Federico Karagulian and Michel

Gerboles; formal analysis, Federico Karagulian and Michel Gerboles; investigation, Federico Karagulian and Michel Gerboles; data curation, Federico Karagulian, Michel Gerboles, Sabine Crunaire, Nathalie Redon, Laurent Spinelle, Caroline Marchand and Benoît Herbin; original draft preparation, Federico Karagulian; review and editing, Federico Karagulian, Michel Gerboles, Alexander Kotsev, Laurent Spinelle, Maurizio Barbiere and Annette Borowiak; visualization, Federico Karagulian; supervision, Michel Gerboles; project administration, Annette Borowiak.; funding acquisition, Annette Borowiak.

Funding: This research was funded by the European Commission's Joint Research Centre and the Directorate General for Environment. The French intercomparison exercise was funded by the French Ministry of Environment through the national reference laboratory for air quality monitoring (LCSQA).

Acknowledgments: In this section you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments). Insert

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the collection, analysis, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Table A1. Number of analyzed Records and sensor models by averaging time.

Averaging time	n. Records	n. OEMs & SSys
hourly	610	86
5 min	253	40
daily	248	42
1 min	214	33

Table A2. Model of OEMs by pollutant, type, openness and price.

model	pollutant	type	reference	open/close	living	price
CO-B4	CO	electrochemical	Wei[59]	open source	N	50
CO 3E300	CO	electrochemical	Gerboles[22]	open source	Y	100
DataRAM pDR-1200	$PM_{2.5}$	nephelometer	Chakrabarti[72]	black box	N	-
DiscMini	PM	OPC	Viana[81]	open source	Y	11000
DN7C3CA006	$PM_{2.5}$	nephelometer	Sousan[85]	open source	Y	10
DSM501A	$PM_{2.5}$	nephelometer	Wang[70], Alvarado[71]	open source	Y	15
FIS SP-61	O_3	MOs	Spinelle[25]	open source	Y	50
GP2Y1010AU0F	$PM_{2.5}$, PM_{10}	nephelometer	Olivares[73], Manikonda[51], Sousan[54], Alvarado[71], Wang[70]	open source	Y	10

MiCS-2710	NO_2	MOs	Spinelle[18], Williams[28]	open source	N	7
MICS-4514	CO, NO_2	MOs	Spinelle[18,23]	open source	Y	20
NO-3E100	NO	electrochemical	Spinelle[23], Gerboles[22]	open source	Y	120
NO-B4	NO	electrochemical	Wei[59]	open source	Y	50
NO2-3E50	NO_2	electrochemical	Spinelle[18], Gerboles[22]	open source	Y	100
NO2-A1	NO_2	electrochemical	Williams[28]	black box	Y	50
NO2-B4	NO_2	electrochemical	Spinelle[18,25]	open source	N	50
NO2-B42F	NO_2	electrochemical	Wei[59]	open source	N	50
NO2-B43F	NO_2	electrochemical	Sun[66]	open source	Y	50
O3-B4	O_3	electrochemical	Spinelle[18,25], Wei[59]	open source	N	50
O3-3E1F	O_3	electrochemical	Spinelle[25], Gerboles[22]	open source	Y	500
OPC-N2	PM_1, $PM_{2.5}$	OPC	AQ-SPEC[17], Mukherjee[79], Sousan[85], Feinberg[34], Crilley[86], Badura[46], Crunaire[31]	open source, black box	N	362
OPC-N3	PM_1, $PM_{2.5}$	OPC	AQ-SPEC[17]	open source	Y	338
PMS1003	$PM_{2.5}$	OPC	Kelly[77]	black box	Y	20
PMS3003	$PM_{2.5}$	OPC	Zheng[87], Kelly[77]	open source, black box	Y	30
PMS5003	$PM_{2.5}$	OPC	Laquai[45]	black box	Y	15
PMS7003	$PM_{2.5}$	OPC	Badura[46]	black box	Y	20

PPD42NS	$PM_{2.5}$, PM_3, PM_2	nephelometer	Wang[70], Holstius[48], Austin[75], Gao[76],Kelly[77]	open source	Y	15
SDS011	$PM_{2.5}$,	OPC	Budde[10], Laquai[45], Badura[46], Liu[88]	open source	Y	30
SM50	O_3	MOs	Feinberg[34]	open source	Y	500
TGS-5042	CO	MOs	Spinelle[23]	open source	Y	40
TZOA-PM Research Sensors	PM	nephelometer	Feinberg[34]	open source	Y	90
ZH03A	$PM_{2.5}$	nephelometer	Badura[46]	black box	Y	20

Table A3. Models of Sensor Systems by pollutant, type, openness and price.

model	pollutant	type	reference	open/close	living	price
2B Tech. (POM)	O_3	UV	AQ-SPEC[17]	black box	Y	4500
Aeroqual-SM50	O_3	MOs	Jiao[37]	black box	Y	2000
AGT ATS-35 NO2	NO_2	MOs	Williams[28]	black box	N	-
Air Quality Station	PM_1, $PM_{2.5}$,	OPC	AQ-SPEC[17]	black box	Y	2000
AirAssure	$PM_{2.5}$	nephelometer	AQ-SPEC[17], Feinberg[34], Manikonda[51]	black box	Y	1500
AirBeam	$PM_{2.5}$	OPC, nephelometer	AQ-SPEC[17], Mukherjee[79], Feinberg[34], Borghi[33], Jiao[37], Crunaire[31]	black box	Y	200
AirCube	NO_2, O_3, NO	electrochemical	Mueller[41], Bigi[63]	black box	Y	3538
AirMatrix	PM1, , $PM_{2.5}$	nephelometer	Crunaire[31]	black box	Y	60

AirNut	$PM_{2.5}$	nephelometer	AQ-SPEC[17]	black box	Y	150
AIRQino	$PM_{2.5}$	OPC	Cavaliere[80]	open source	Y	1000
AirSensEUR (v.1)	NO, NO_2, O_3	electrochemical	Crunaire[31]	black box	Y	1600
AirSensEUR (v.2)	CO, NO, NO_2, O_3	electrochemical	Karagulian[21]	open source	Y	1600
AirSensorBox	NO_2, CO, O_3,	electrochemical, MOs, nephelometer	Borrego[50]	black box	Y	280
AirThinx	PM_1, $PM_{2.5}$	OPC	AQ-SPEC[17]	black box	Y	1000
AirVeraCity	CO, NO_2, O_3	electrochemical, MOs	Marjovi[56]	black box	Y	10000
AirVisual Pro	$PM_{2.5}$	nephelometer	AQ-SPEC[17]	black box	Y	270
AQMesh v.3.0	CO, NO	electrochemical	Jiao[37]	black box	N	10000
AQMesh v.4.0	NO_2, CO, NO, O_3	electrochemical	Cordero[64], AQ-SPEC[17], Castell[57], Borrego[50], Crunaire[31]	black box	updated	10000
AQT410 v.1.11	O_3	electrochemical	AQ-SPEC[17]	black box	Y	3700
AQT-420	NO_2,O_3, , $PM_{2.5}$	electrochemical, OPC	Crunaire[31]	black box	Y	3256
AQY v0.5	$PM_{2.5}$, NO_2, O_3	OPC, electrochemical, MOs	AQ-SPEC[17]	black box	updated	3000
ARISense	NO_2, CO, NO, O_3	electrochemical	Cross[58]	black box	Y	-
Atmotrack	PM_1, , $PM_{2.5}$	nephelometer	Crunaire[31]	black box	Y	2500
BAIRS	$PM_{2.5-0.5}$	OPC	Northcross[103]	open source	N	475
Cair	$PM_{2.5}$,$PM_{10-2.5}$	OPC	AQ-SPEC[17]	black box	Y	200
CairClip O3/NO2	O_3, NO_2	electrochemical	Jiao[33], Spinelle[24],	black box	Y	600

			Williams[28], Duvall[65], Feinberg[34]			
CairClip NO2-F	NO_2	electrochemical	Spinelle[18], Duvall[65], Crunaire[31]	black box	Y	600
CairClip PM2.5	$PM_{2.5}$	nephelometer	Williams[29]	black box	Y	1500
CAM	PM_{10}, $PM_{2.5}$, NO_2, CO, NO	OPC, electrochemical	Borrego[50]	black box	Y	-
CanarIT	PM	nephelometer	Williams[29]	black box	N	1500
Clarity Node	$PM_{2.5}$	nephelometer	AQ-SPEC[17]	black box	Y	1300
Dylos DC1100	$PM_{2.5-0.5}$	OPC	Jiao[37], Williams[29], Feinberg[34]	black box, open source	Y	300
Dylos DC1100 PRO	$PM_{2.5-0}$, $PM_{10-2.5}$, PM_{10}	OPC	Jiao[37], AQ-SPEC[17], Feinberg[34], Manikonda[51]	black box, open source	Y	300
Dylos DC1700	$PM_{2.5-0.5}$, PM_{10},$PM_{10-2.5}$, PM_3, PM_2, $PM_{2.5}$	OPC	Manikonda[51], Sousan[85], Northcross[103], Holstius[48], Steinle[82], Han[83], Jovasevic[84], Dacunto[54]	open source	Y	475
e-PM	PM_{10}, $PM_{2.5}$	nephelometer	Crunaire[31]	black box	Y	2500
E-Sampler	$PM_{2.5}$	OPC	AQ-SPEC[17]	black box	Y	5500
ECN_Box	PM_{10}, $PM_{2.5}$, NO_2, O_3	nephelometer, electrochemical	Borrego[50]	black box	Y	274
Eco PM	PM1	OPC	Williams[29]	black box	N	
ECOMSMART	NO_2, O_3, PM_1, PM_{10}, $PM_{2.5}$	electrochemical, OPC	Crunaire[31]	black box	Y	4560
Egg (2018)	PM_1, $PM_{2.5}$, PM_{10}	OPC	AQ-SPEC[17]	black box	Y	249

Egg v.1	CO, NO_2, O_3	MOs	AQ-SPEC[17]	black box	N	200
Egg v.2	CO, NO_2, O_3	electrochemical	AQ-SPEC[17]	black box	Y	240
Egg v.2 (PM)	$PM_{2.5}$, PM_{10}	nephelometer	AQ-SPEC[17]	black box	Y	280
ELM	NO_2, PM_{10}, O_3	MOs, nephelometer	AQ-SPEC[17], US-EPA[69]	black box	N	5200
EMMA	$PM_{2.5}$, CO, NO_2, NO	OPC, electrochemical	Gillooly[60]	black box	Y	-
ES-642	$PM_{2.5}$	OPC	Crunaire[31]	black box	Y	2600
Foobot	$PM_{2.5}$	OPC	AQ-SPEC[17]	black box	Y	200
Hanvon N1	$PM_{2.5}$	nephelometer	AQ-SPEC[17]	black box	Y	200
Intel Berkeley Badge	NO_2, O_3	electrochemical, MOs	Vaughn[30]	open source	N	-
ISAG	NO_2, O_3	MOs	Borrego[50]	black box	N	-
Laser Egg	$PM_{2.5}$, PM_{10}	nephelometer	AQ-SPEC[17]	black box	Y	200
M-POD	CO, NO_2	MOs	Piedrahita[62]	black box	N	
MAS	CO, NO_2, O_3, $PM_{2.5}$	electrochemical, UV, nephelometer	Sun[38]	black box, open source	N, Y	5500
Met One - 831	PM_{10}	OPC	Williams[29]	black box	Y	2050
Met One (NM)	$PM_{2.5}$	OPC	AQ-SPEC[17]	black box	Y	1900
MicroPEM	$PM_{2.5}$	nephelometer	AQ-SPEC[17], Williams[29]	black box	Y	2000
NanoEnvi	NO_2, O_3, CO	electrochemical, MOs	Borrego[50]	black box	Y	-
PA-I	PM_1, $PM_{2.5}$, PM_{10}	OPC	AQ-SPEC[17]	black box	N	150
PA-I-Indoor	$PM_{2.5}$, PM_{10}	OPC	AQ-SPEC[17]	black box	Y	180

PA-II	PM_1, $PM_{2.5}$, PM_{10}	OPC	AQ-SPEC[17]	black box	Y	200
Partector	PM_1, $PM_{2.5}$	Electrical	AQ-SPEC[17]	black box	Y	7000
PATS+	$PM_{2.5}$	nephelometer	Pillarisetti[74]	black box	Y	500
Platypus NO2	NO_2	MOs	Williams[28]	black box	Y	50
PMS-SYS-1	$PM_{2.5}$	nephelometer	Jiao[37], AQ-SPEC[17], Williams[29], Feinberg[34]	black box	Y	1000
Portable AS-LUNG	PM_1, $PM_{2.5}$, PM_{10}	OPC	AQ-SPEC[17]	black box	Y	1000
Pure Morning P3	$PM_{2.5}$	OPC	AQ-SPEC[17]	black box	Y	170
RAMP	CO, NO_2	electrochemical	Zimmerman[61]	open source	Y	-
S-500	NO_2, O_3	MOs	Lin[67], AQ-SPEC[17], Vaughn[30]	black box	Y	500
SENS-IT	O_3, CO, NO_2	MOs	AQ-SPEC[17]	black box	N, Y	2200
SidePak AM510	$PM_{2.5}$	nephelometer	Karagulian[78]	open source	Y	3000
Smart Citizen Kit	CO	MOs	AQ-SPEC[17]	black box	Y	200
SNAQ	NO_2, CO, NO	electrochemical	Mead[42], Popoola[43]	black box	Y	-
Spec	CO, NO_2, O_3	electrochemical	AQ-SPEC[17]	black box	Y	500
Speck	$PM_{2.5}$	nephelometer	Feinberg[34], US-EPA[69], Williams[29], AQ-SPEC[17], Manikonda[51], Zikova[52]	black box	Y	150
UBAS	$PM_{2.5}$	nephelometer	Manikonda[51]	black box	N	100
uHoo	$PM_{2.5}$, O_3	nephelometer, MOs	AQ-SPEC[17]	black box	Y	300
Urban AirQ	NO_2	electrochemical	Mijling[39]	open source	N	-

Vaisala AQT410 v.1.11	CO, NO_2	electrochemical	AQ-SPEC[17]	black box	Y	3700
Vaisala AQT410 v.1.15	CO, NO_2	electrochemical	AQ-SPEC[17]	black box	Y	3700
Waspmote	NO, NO_2, PM_1, PM_{10}, $PM_{2.5}$	MOs, OPC	Crunaire[31]	black box	Y	1270
Watchtower 1	NO_2,PM_1, PM_{10}, $PM_{2.5}$, O_3	electrochemical, OPC	Crunaire[31]	black box	Y	5000

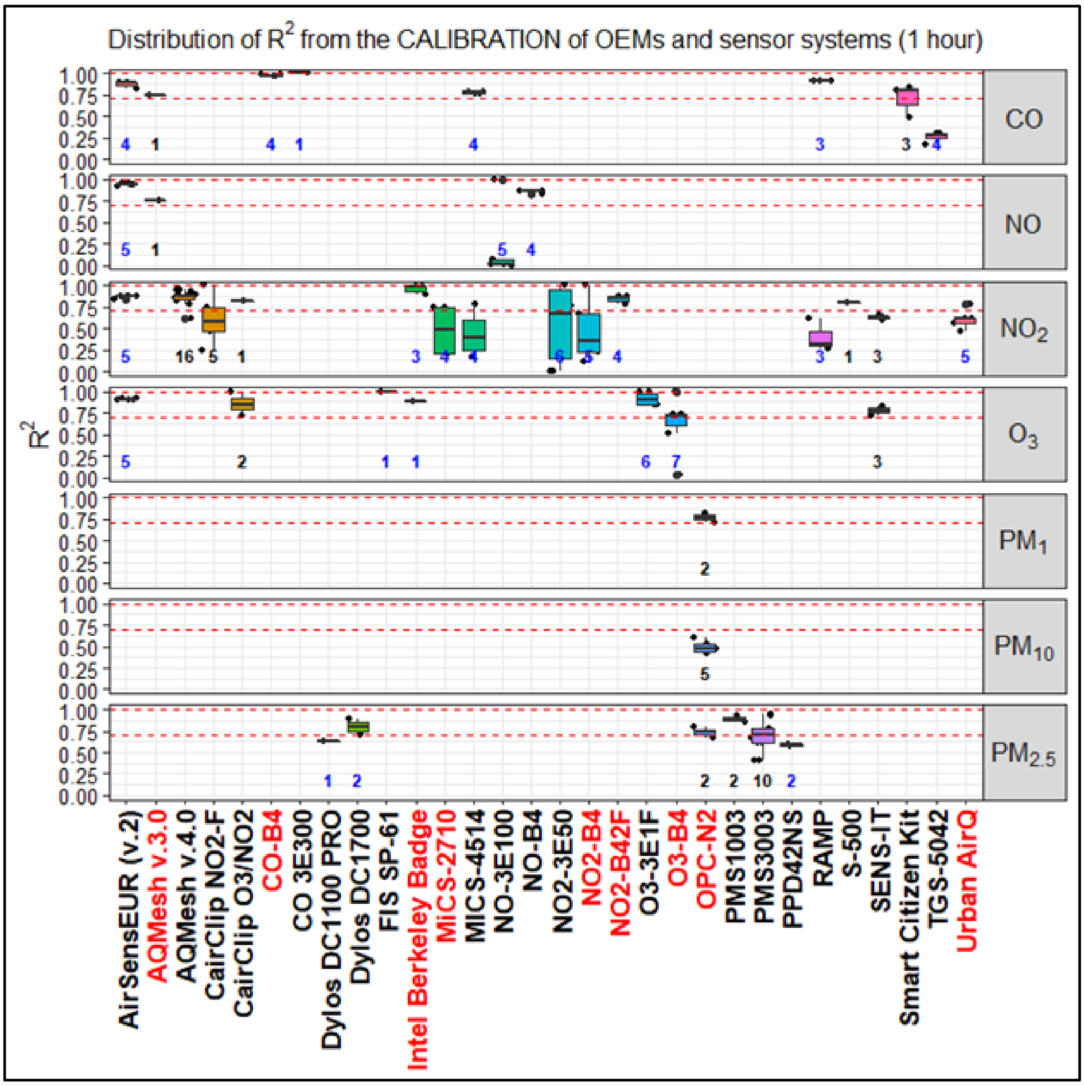

Figure A1. Distribution of R^2 for LCSs hourly data against the reference for different pollutants. Dashed lines indicate the R^2 value of 0.7 and 1.0. Numbers in blue and black indicate the number of open source and black box Records, respectively. Names of ‘living’ and ‘non-living’ sensors are indicated in black and red labels of the x-axis color, respectively.

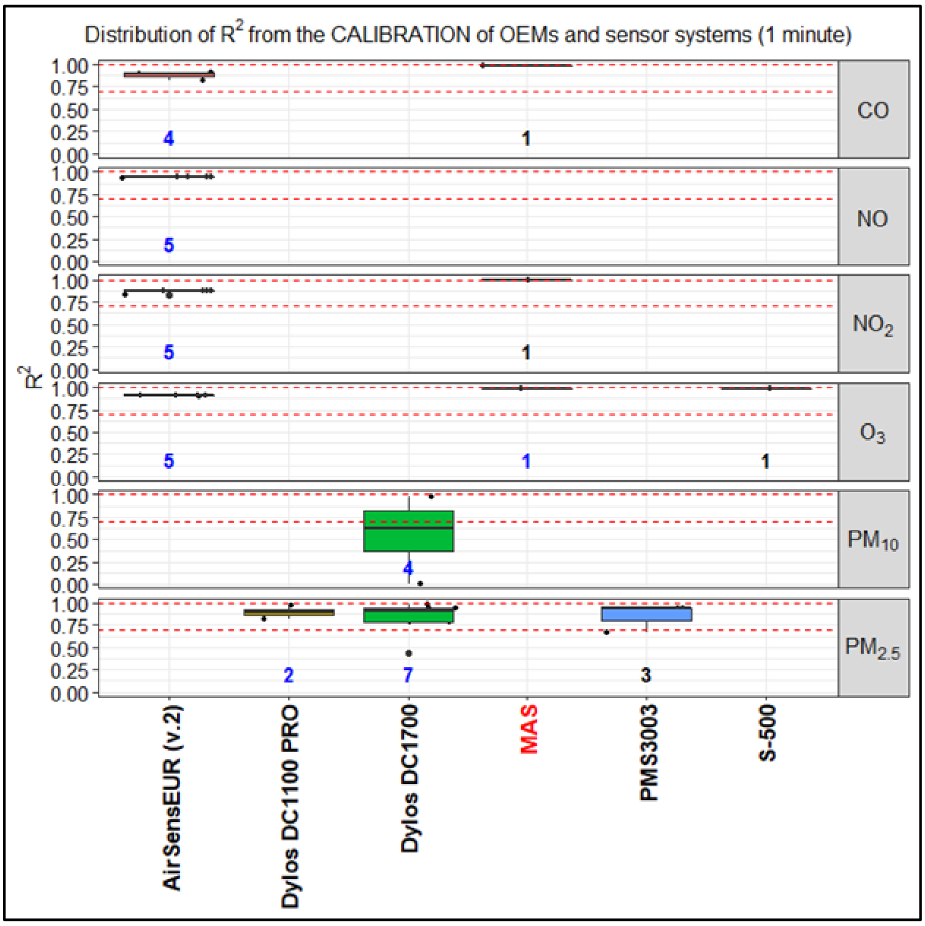

Figure A2. Distribution of R^2 for LCSs minute data against the reference for different pollutants. Dashed lines indicate the R^2 value of 0.7 and 1.0. Numbers in blue and black indicate the number of open source and black box Records, respectively. Names of 'living' and 'non-living' sensors are indicated in black and red labels of the x-axis color, respectively.

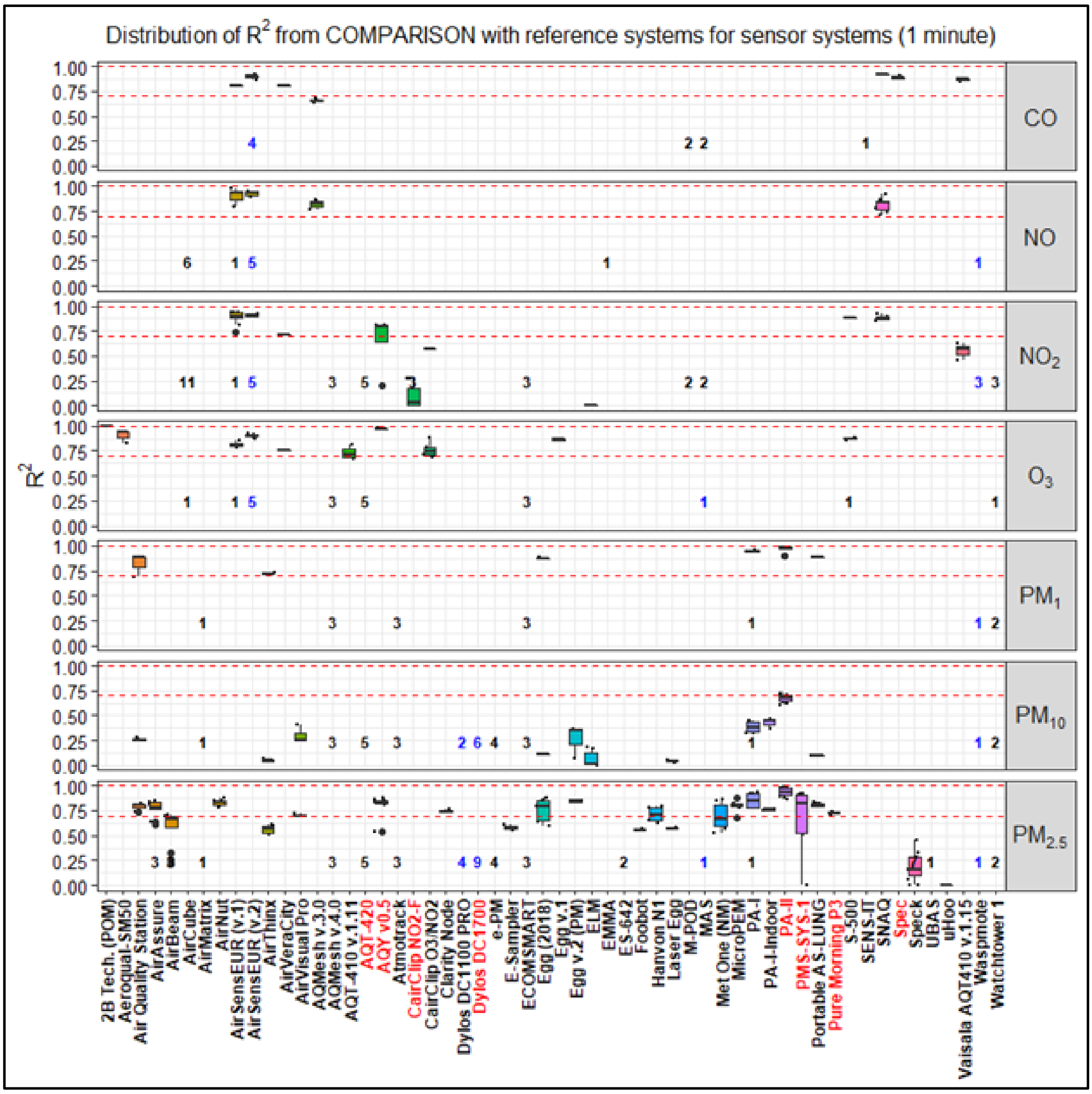

Figure A3. Distribution of R^2 from the comparison of SSys minute data against reference measurements. Numbers in blue and black indicate the number of open source and black box Records, respectively. Names of 'living' and 'non-living' sensors are indicated in black and red labels of the x-axis color, respectively.

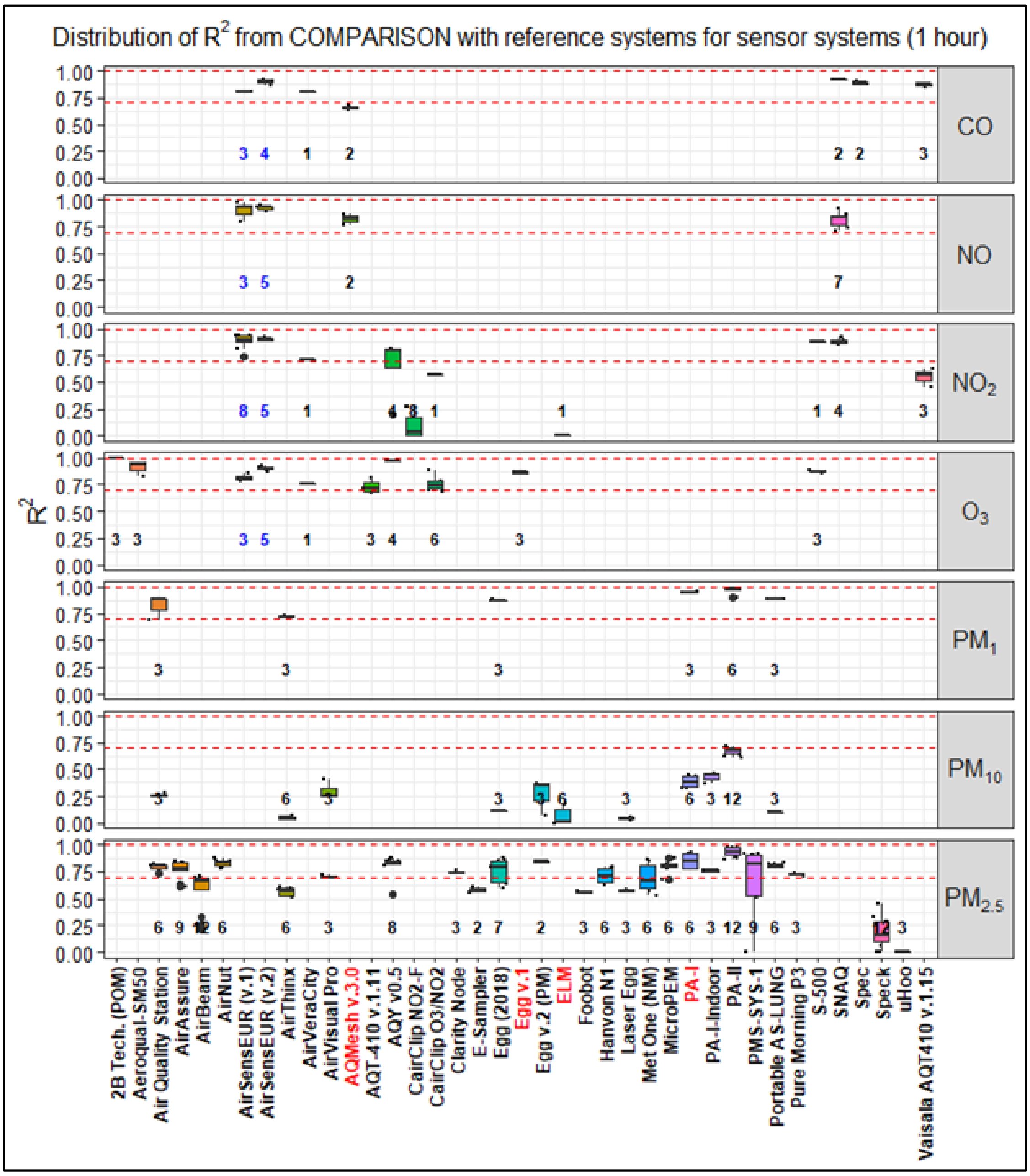

Figure A4. Distribution of R^2 from the comparison of SSys hourly data against reference measurements. Numbers in blue and black indicate the number of open source and black box Records, respectively. Names of 'living' and 'non-living' sensors are indicated in black and red labels of the x-axis color, respectively.

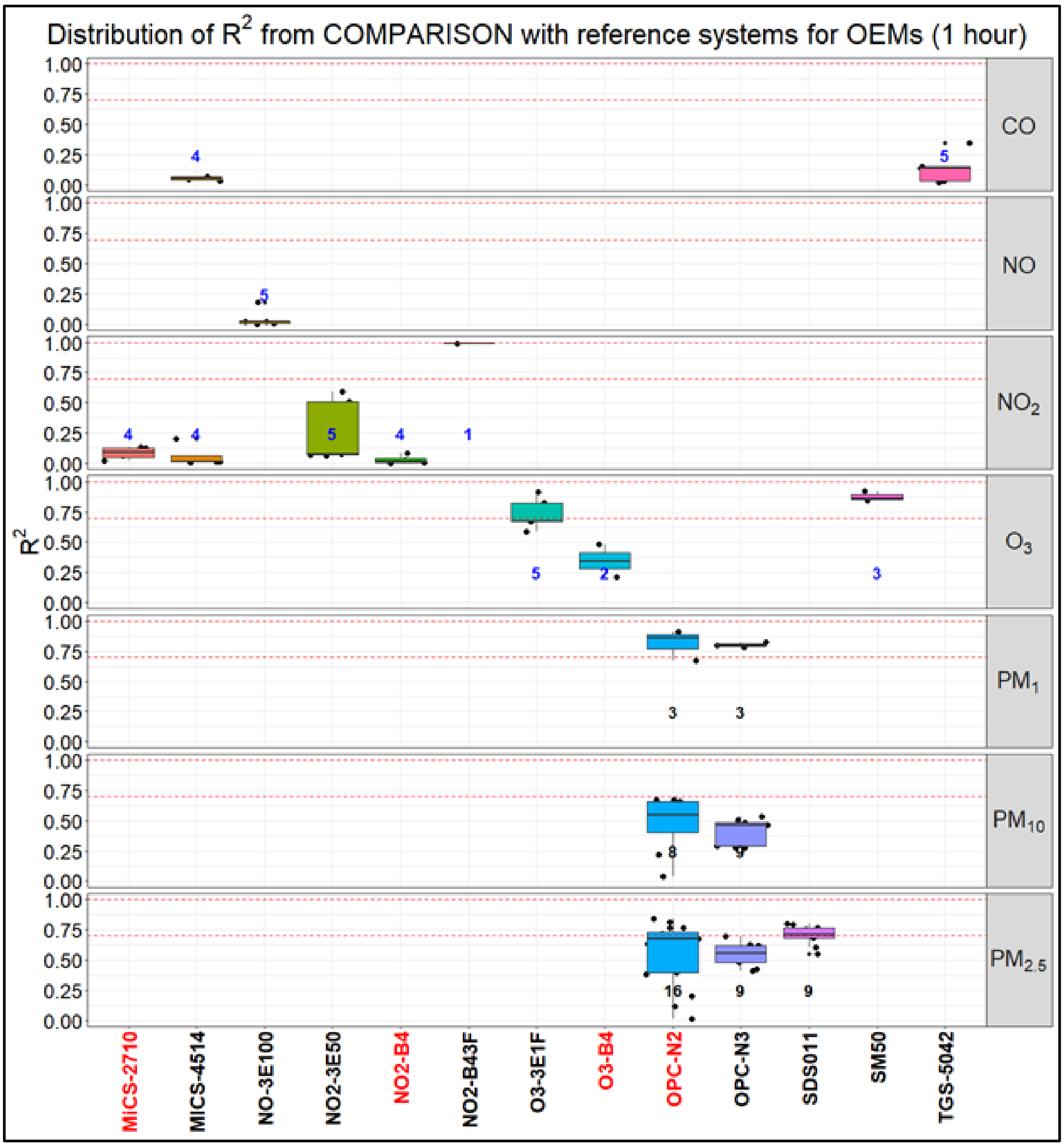

Figure A5. Distribution of R^2 from the comparison of all OEMs against reference systems. Records were averaged over a time-scale of 1 hour. Numbers in blue and black indicate the number of open source and black box Records, respectively. Names of 'living' and 'non-living' sensors are indicated in black and red labels of the x-axis color, respectively..

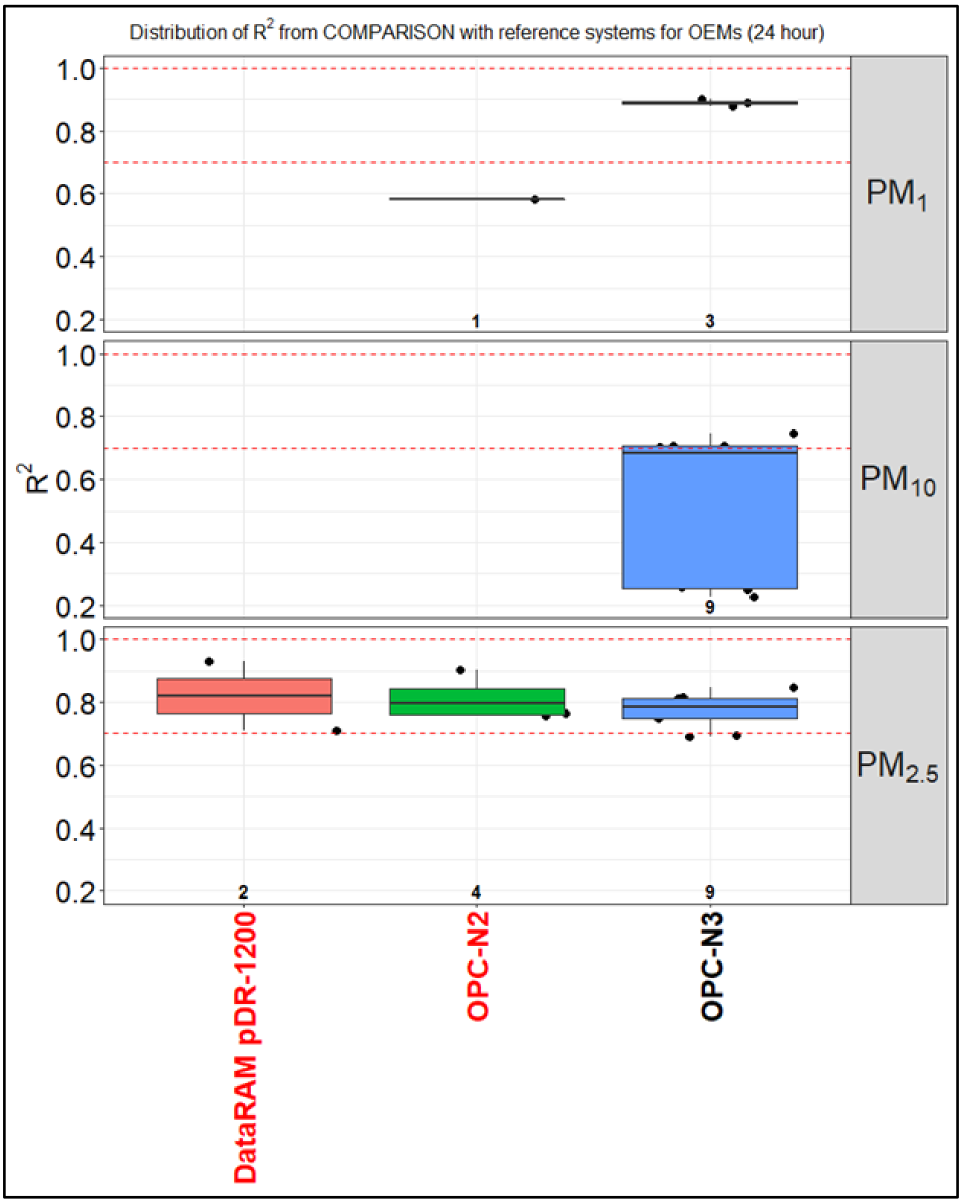

Figure A6. Distribution of R^2 from the comparison of all OEMs against reference systems. Records were averaged over a time-scale of daily data. Numbers in blue and black indicate the number of open source and black box Records, respectively. Names of 'living' and 'non-living' sensors are indicated in black and red labels of the x-axis color, respectively.

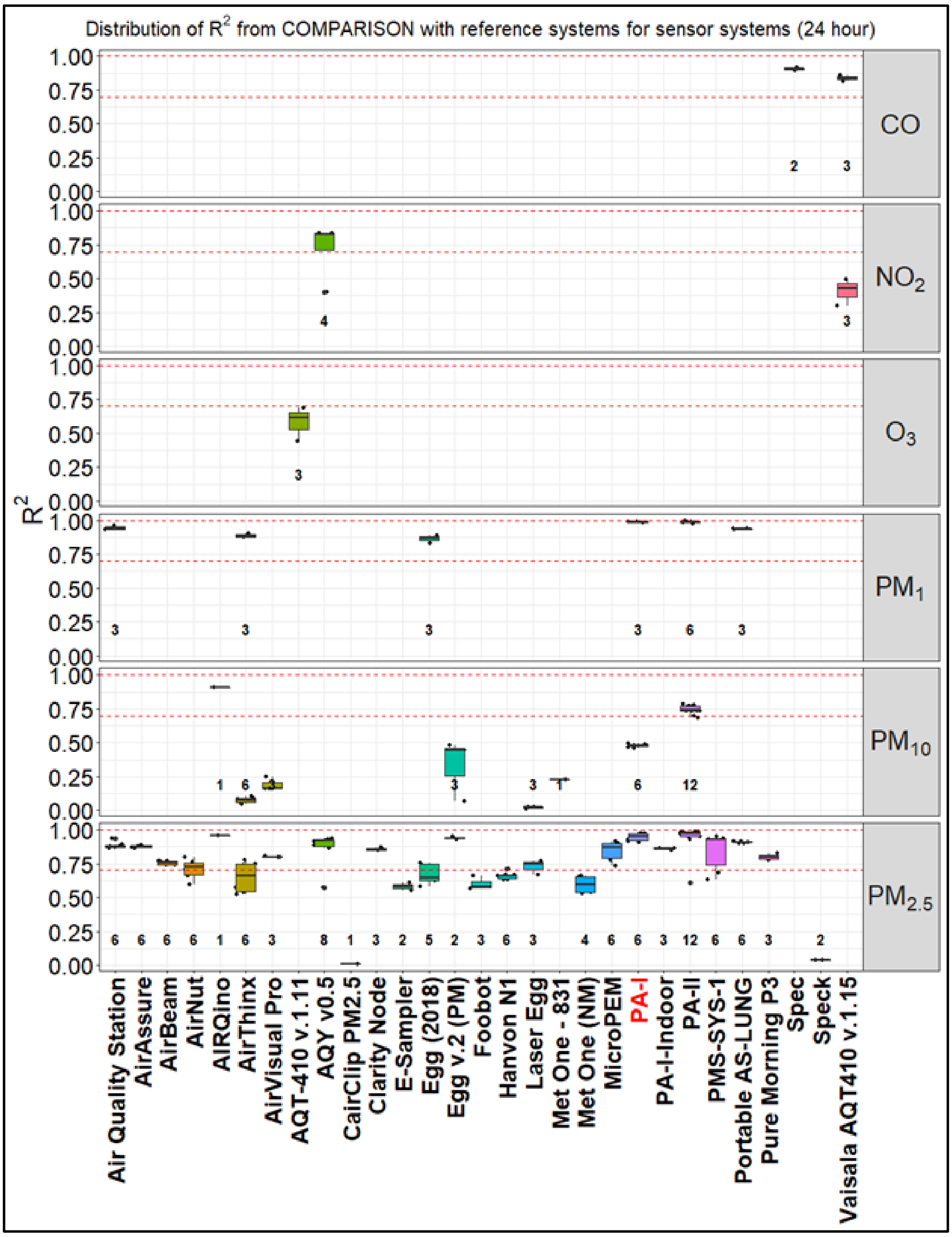

Figure A7. Distribution of R^2 from the comparison of all sensor systems against reference systems. Records were averaged over a time-scale of daily data. Numbers in blue and black indicate the number of open source and black box Records, respectively. Names of 'living' and 'non-living' sensors are indicated in black and red labels of the x-axis color, respectively.

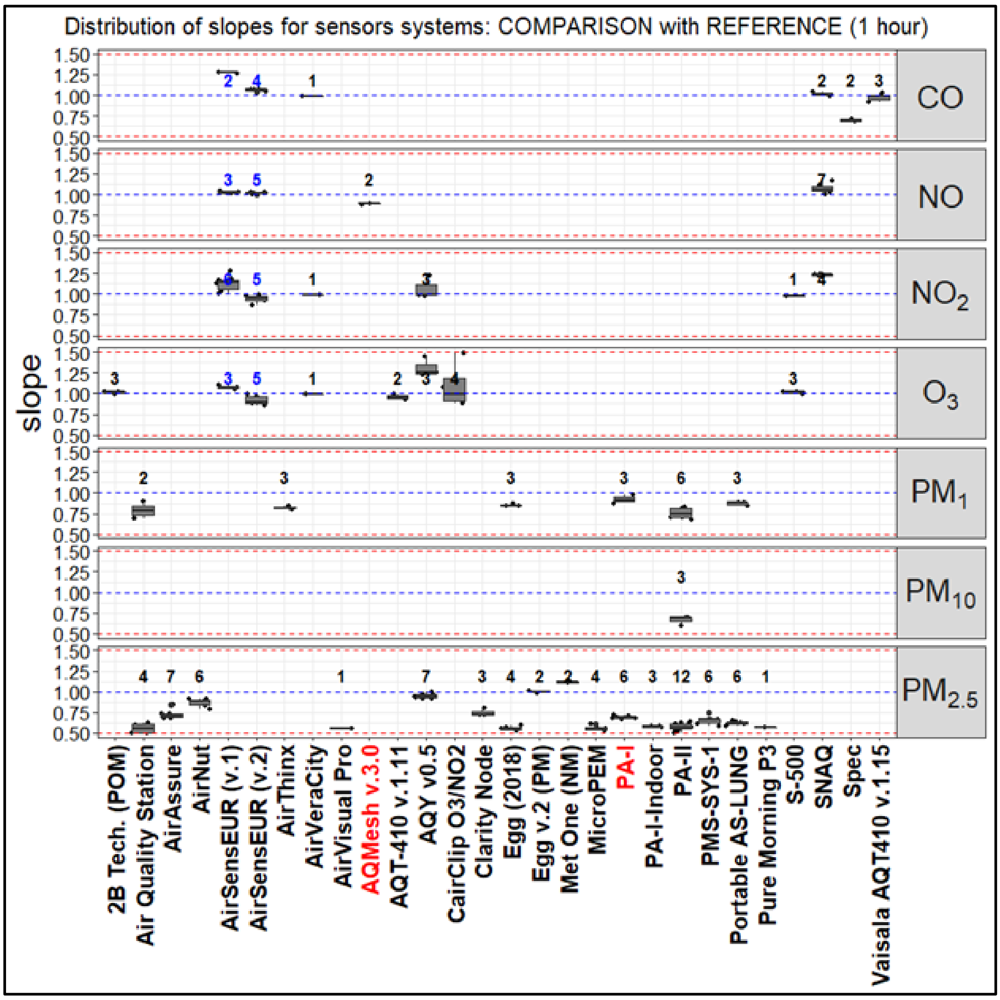

Figure A8. Distribution of slopes from the comparison of SSys against the reference. Only Records with $R^2 > 0.7$ and $0.5 < \text{slope} < 1.5$ are shown. Records were averaged over a time-scale of 1 hour. Numbers in blue and black indicate the number of open source and black box Records, respectively. Names of 'living' and 'non-living' sensors are indicated in black and red labels of the x-axis color, respectively.

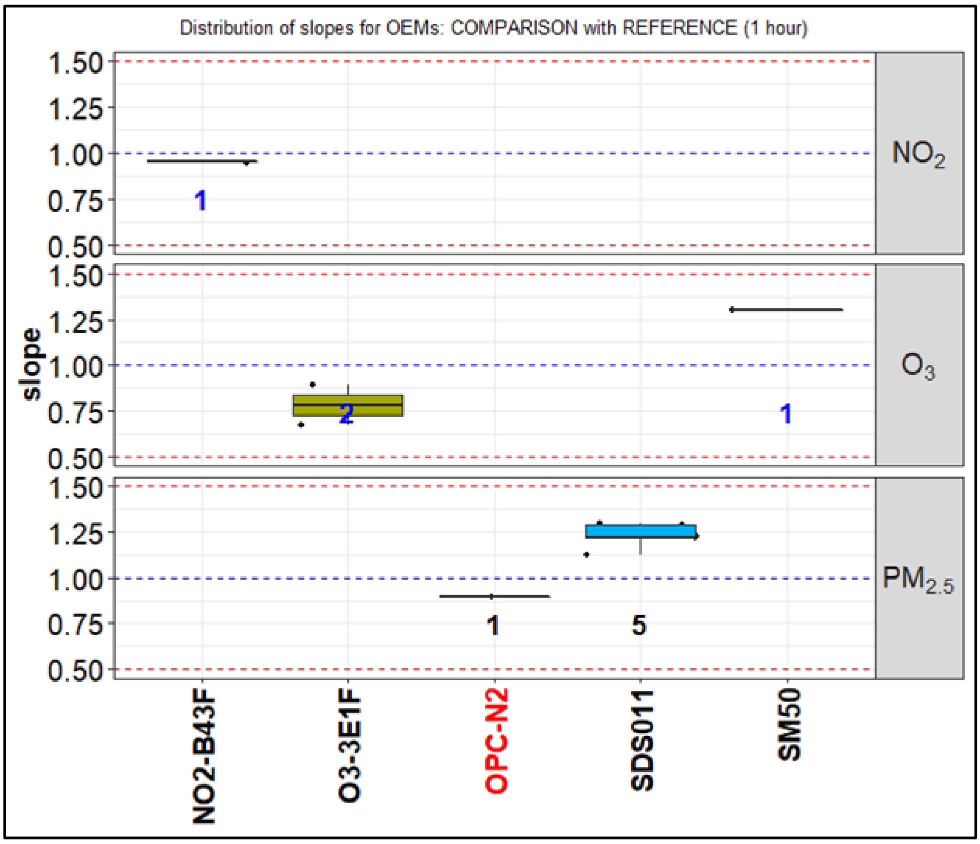

Figure A9. Distribution of slopes from the comparison of OEMs against the reference. Only hourly Records with $R^2 > 0.7$ and $0.5 <$ slope < 1.5 are shown. Numbers in blue and black indicate the number of open source and black box Records, respectively. Names of 'living' and 'non-living' sensors are indicated in black and red labels of the x-axis color, respectively.

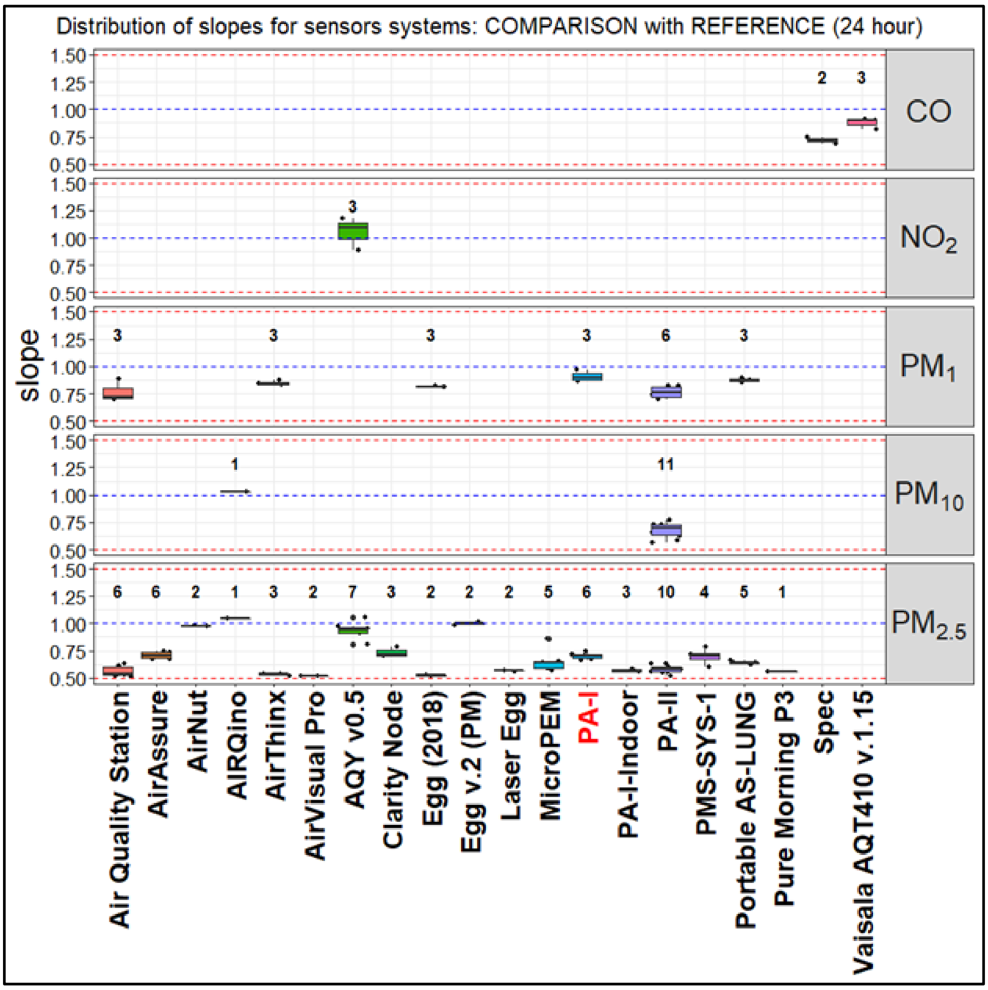

Figure A10. Distribution of slopes from the comparison of SSys against the reference. Only Records with $R^2 > 0.7$ and $0.5 < \text{slope} < 1.5$ are shown. Records were averaged over a time-scale of daily data. Numbers in blue and black indicate the number of open source and black box Records, respectively. Names of 'living' and 'non-living' sensors are indicated in black and red labels of the x-axis color, respectively.

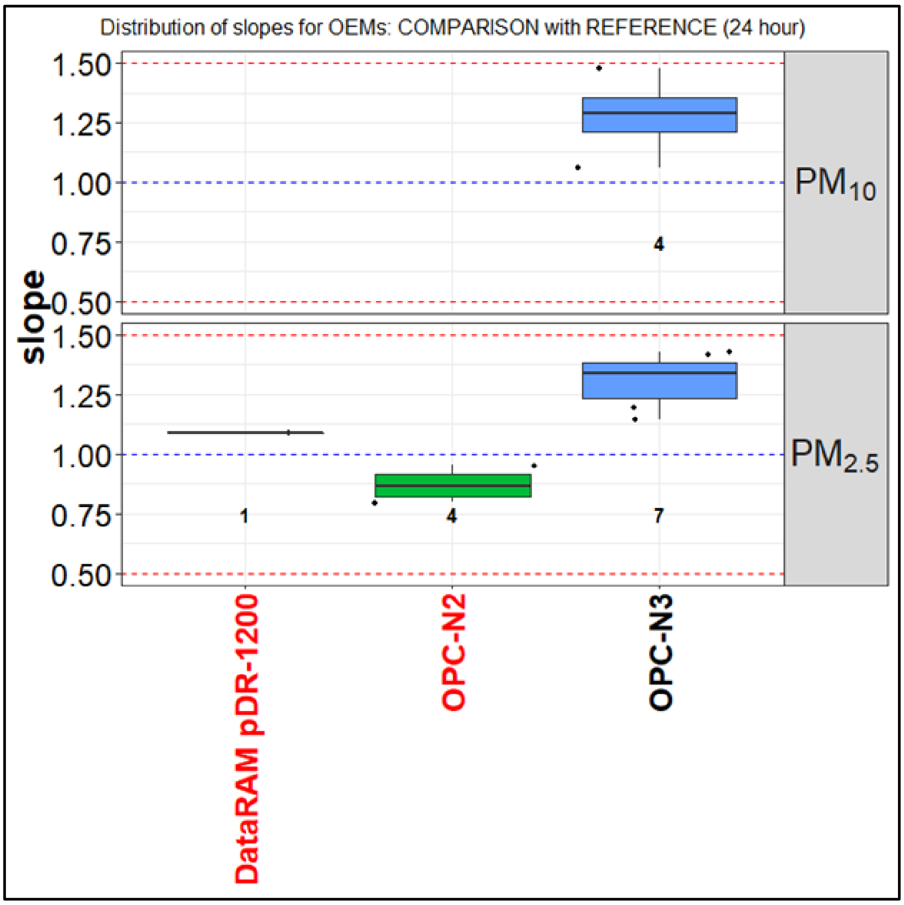

Figure A11. Distribution of slopes from the comparison of OEMs against the reference. Only Records with $R^2 > 0.7$ and $0.5 < \text{slope} < 1.5$ are shown. Records were averaged over a time-scale of daily data. Numbers in blue and black indicate the number of open source and black box Records, respectively. Names of 'living' and 'non-living' sensors are indicated in black and red labels of the x-axis color, respectively.

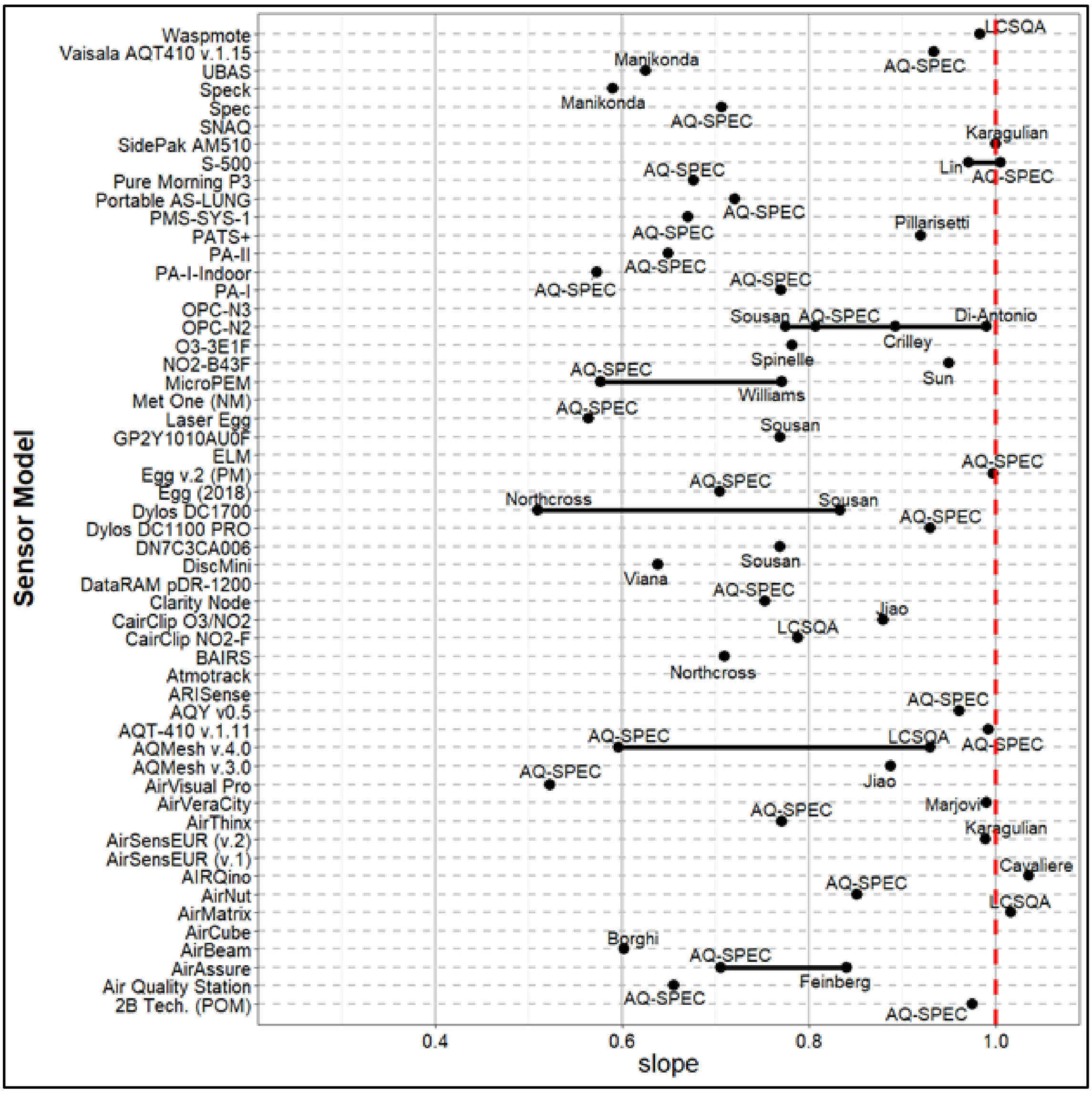

Figure A12. Mean slope obtained from the comparison of LCSs against reference measurements.

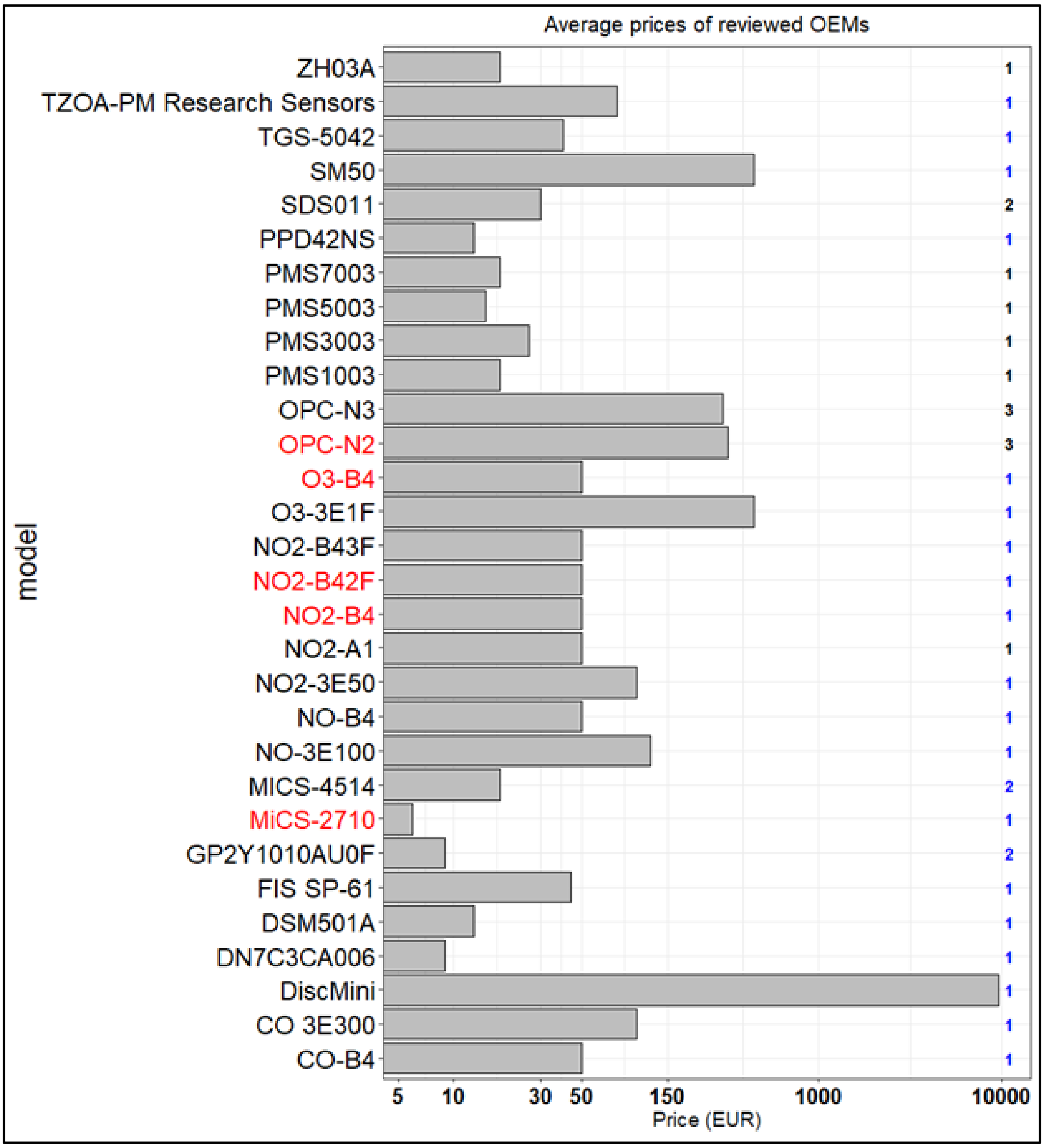

Figure A13. Prices of OEMs grouped by model. Numbers at right indicates the number of pollutants measured by each OEMs, with open source in blue and black box in black. x-axis uses logarithmic scale. Names of 'living' and 'non-living' OEMs are indicated in black and red color on the labels of the y-axis, respectively.

Table A7. Shortlist of SSys showing good agreement with reference systems ($R^2 > 0.85$; $0.8 <$ slope $<$ 1.2) for daily data.

model	pollutant	mean	mean slope	mean absolute intercept	open/close	living	commercial	price (EUR)
PA-I	PM_1	0.99	0.9	0.47	black box	N	commercial	132
PA-II	PM_1	0.99	0.8	1.8	black box	Y	commercial	176
Egg (2018)	PM_1	0.88	0.8	0.33	black box	Y	commercial	219
Egg v.2 (PM)	$PM_{2.5}$	0.94	1	3.3	black box	Y	commercial	246
AirThinx	PM_1	0.89	0.8	1.3	black box	Y	commercial	880
Portable AS-LUNG	PM_1	0.93	0.9	1.5	black box	Y	non commercial	880
AIRQino	$PM_{2.5}$,PM_{10}	0.91	1	1.1	open source	Y	non commercial	1000
Air Quality Station	PM_1	0.94	0.9	1.1	black box	Y	non commercial	1760
AQY v0.5	$PM_{2.5}$	0.91	0.9	4.0	black box	updated	commercial	2640
Vaisala AQT410 v.1.15	CO	0.86	0.9	0.25	black box	Y	commercial	3256

References

1. 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. *Environment International* **2015**, *75*, 199–205.
2. 2008/50/EC: Directive of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe Available online: http://eur-lex.europa.eu/Result.do?RechType=RECH_celex&lang=en&code=32008L0050 (accessed on Feb 15, 2011).
3. CEN *Ambient air - Standard gravimetric measurement method for the determination of the PM10 or PM2,5 mass concentration of suspended particulate matter (EN 12341:2014). European Committee for Standardization.*; 2014;
4. CEN *Ambient air - Standard method for the measurement of the concentration of carbon monoxide by non-dispersive infrared spectroscopy, (EN 14626:2012). European Committee for Standardization.*; 2012;
5. CEN *Ambient air - Standard method for the measurement of the concentration of nitrogen dioxide and nitrogen monoxide by chemiluminescence (EN 14211:2012). European Committee for Standardization.*; 2012;
6. CEN *Ambient air - Standard method for the measurement of the concentration of ozone by ultraviolet photometry, (EN 14625:2012). European Committee for Standardization.*; 2012;
7. CEN *Ambient air - Standard method for the measurement of the concentration of sulphur dioxide by ultraviolet fluorescence, (EN 14212:2012). European Committee for Standardization.*; 2012;
8. Lewis, A.C.; von Schneidemesser, E.; Peltier, R. *Low-cost sensors for the measurement of atmospheric composition: overview of topic and future applications (World Meteorological Organization) WMO-No. 1215.*; 2018;
9. Aleixandre, M.; Gerboles, M. Review of small commercial sensors for indicative monitoring of ambient gas. *Chemical Engineering Transactions* **2012**, *30*, 169–174.
10. Budde, M.; Müller, T.; Laquai, B.; Streibl, N.; Schwarz, A.; Schindler, G.; Riedel, T.; Beigl, M.; Dittler, A. Suitability of the Low-Cost SDS011 Particle Sensor for Urban PM-Monitoring. **2018**.
11. iScape *Summary of Air Quality sensors and recommendations for application. iScape project D1.5, February 2017*; 2017;
12. Snyder, E.G.; Watkins, T.H.; Solomon, P.A.; Thoma, E.D.; Williams, R.W.; Hagler, G.S.W.; Shelow, D.; Hindin, D.A.; Kilaru, V.J.; Preuss, P.W. The changing paradigm of air pollution monitoring. *Environmental Science and Technology* **2013**, *47*, 11369–11377.
13. White, R.M.; Paprotny, I.; Doering, F.; Cascio, W.E.; Solomon, P.A.; Gundel, L.A. Sensors and "apps" for community-based: Atmospheric monitoring. *EM: Air and Waste Management Association's Magazine for Environmental Managers* **2012**, 36–40.
14. Williams, R.; Kilaru, V.; Snyder, E.; Kaufman, A.; Dye, T.; Rutter, A.; Russell, A.; Hafner, H. *Air Sensor Guidebook, United States Environmental Protection Agency (US-EPA)*; 2014;
15. Zhou, X.; Lee, S.; Xu, Z.; Yoon, J. Recent Progress on the Development of Chemosensors for Gases. *Chemical Reviews* **2015**, *115*, 7944–8000.
16. Williams, R.; Duvall, R.; Kilaru, V.; Hagler, G.; Hassinger, L.; Benedict, K.; Rice, J.; Kaufman, A.; Judge, R.; Pierce, G.; et al. Deliberating performance targets workshop: Potential paths for emerging PM2.5 and O3 air sensor progress. *Atmospheric Environment: X* **2019**, *2*, 100031.
17. AQ-SPEC; South Coast Air Quality Management District; South Coast Air Quality Management District Air Quality Sensor Performance Evaluation Reports. Available online: http://www.aqmd.gov/aq-spec/evaluations#&MainContent_C001_Col00=2 (accessed on Dec 29, 2015).
18. 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. *Sensors and Actuators B: Chemical* **2015**, *215*, 249–257.
19. Lewis, A.; Edwards, P. Validate personal air-pollution sensors. *Nature News* **2016**, *535*, 29.
20. Aleixandre, M.; Gerboles, M. Review of small commercial sensors for indicative monitoring of

ambient gas. *Chemical Engineering Transactions* **2012**, *30*, 169–174.

21. Karagulian, F.; Borowiak, A.; Barbiere, M.; Kotsev, A.; van der Broecke, J.; Vonk, J.; Signorini, M.; Gerboles, M. *Calibration of AirSensEUR units during a field study in the Netherlands*; European Commission - Joint Research Centre: Ispra (ITALY), 2019;
22. Gerboles, M.; Spinelle, L.; Signorini, M. *AirSensEUR: an open data/software/hardware multi-sensor platform for air quality monitoring. Part A: sensor shield*; Publications Office of the European Union, 2015;
23. 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 CO2. *Sensors and Actuators B: Chemical* **2017**, *238*, 706–715.
24. Spinelle, L.; Gerboles, M.; Aleixandre, M. Performance Evaluation of Amperometric Sensors for the Monitoring of O 3 and NO 2 in Ambient Air at ppb Level. *Procedia Engineering* **2015**, *120*, 480–483.
25. Spinelle, L.; Gerboles, M.; Aleixandre, M.; Bonavitacola, F. Evaluation of metal oxides sensors for the monitoring of O3 in ambient air at ppb level. *Chemical Engineering Transactions* **2016**, 319–324.
26. Spinelle, L.; Gerboles, M.; Kotsev, A.; Signorini, M. *Evaluation of low-cost sensors for air pollution monitoring: Effect of gaseous interfering compounds and meteorological conditions*; Publications Office of the European Union, 2017;
27. US-EPA *Air Sensor Toolbox; Evaluation of Emerging Air Pollution Sensor Performance. US-EPA: https://www.epa.gov/air-sensor-toolbox/evaluation-emerging-air-pollution-sensor-performance*; 2018;
28. Williams, R.; Long, R.; Beaver, M.; Kaufman, A.; Zeiger, F.; Heimbinder, M.; Acharya, B.R.; Grinwald, B.A.; Kupcho, K.A.; Tobinson, S.E. *Sensor Evaluation Report*; Office of Research and Development National Exposure Research Laboratory; 2014;
29. Williams, R.; Kaufman, A.; Hanley, T.; Rice, J.; Garvey, S. *Evaluation of Field-deployed Low Cost PM Sensors*; Office of Research and Development National Exposure Research Laboratory; 2014;
30. Vaughn, D.L.; Dye, T.S.; Roberts, P.T.; Ray, A.E.; DeWinter, J.L. *Characterization of low-Cost NO2 Sensors*; U.S. Environmental Protection Agency Research Triangle Park, NC, 2010;
31. Crunaire, S.; Redon, N.; Spinelle, L. *1ER Essai national d'Aptitude des Microcapteurs EAµC) pour la Surveillance de la Qualité de l'Air: Synthèse des Résultas*; LCSQA, 2018; p. 38;.
32. Fishbain, B.; Lerner, U.; Castell, N.; Cole-Hunter, T.; Popoola, O.; Broday, D.M.; Iñiguez, T.M.; Nieuwenhuijsen, M.; Jovasevic-Stojanovic, M.; Topalovic, D.; et al. An evaluation tool kit of air quality micro-sensing units. *Science of The Total Environment* **2017**, *575*, 639–648.
33. Borghi, F.; Spinazzè, A.; Campagnolo, D.; Rovelli, S.; Cattaneo, A.; Cavallo, D.M. Precision and Accuracy of a Direct-Reading Miniaturized Monitor in PM2.5 Exposure Assessment. *Sensors* **2018**, *18*, 3089.
34. Feinberg, S.; Williams, R.; Hagler, G.S.W.; Rickard, J.; Brown, R.; Garver, D.; Harshfield, G.; Stauffer, P.; Mattson, E.; Judge, R.; et al. Long-term evaluation of air sensor technology under ambient conditions in Denver, Colorado. *Atmospheric Measurement Techniques* **2018**, *11*, 4605–4615.
35. Mukherjee, A.; Stanton, L.G.; Graham, A.R.; Roberts, P.T. Assessing the Utility of Low-Cost Particulate Matter Sensors over a 12-Week Period in the Cuyama Valley of California. *Sensors* **2017**, *17*, 1805.
36. Zikova, N.; Masiol, M.; Chalupa, D.C.; Rich, D.Q.; Ferro, A.R.; Hopke, P.K. Estimating Hourly Concentrations of PM2.5 across a Metropolitan Area Using Low-Cost Particle Monitors. *Sensors (Basel)* **2017**, *17*.
37. Jiao, W.; Hagler, G.; Williams, R.; Sharpe, R.; Brown, R.; Garver, D.; Judge, R.; Caudill, M.; Rickard, J.; Davis, M.; et al. Community Air Sensor Network (CAIRSENSE) project: evaluation of low-cost sensor performance in a suburban environment in the southeastern United States. *Atmospheric Measurement Techniques* **2016**, *9*, 5281–5292.
38. Sun, L.; Wong, K.C.; Wei, P.; Ye, S.; Huang, H.; Yang, F.; Westerdahl, D.; Louie, P.K.K.; Luk,

C.W.Y.; Ning, Z. Development and Application of a Next Generation Air Sensor Network for the Hong Kong Marathon 2015 Air Quality Monitoring. *Sensors (Basel)* **2016**, *16*.
39. Mijling, B.; Jiang, Q.; de Jonge, D.; Bocconi, S. Practical field calibration of electrochemical NO2 sensors for urban air quality applications. *Atmos. Meas. Tech. Discuss.* **2017**, *2017*, 1–25.
40. 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. *Atmospheric Measurement Techniques Discussions* **2018**, 1–29.
41. Mueller, M.; Meyer, J.; Hueglin, C. Design of an ozone and nitrogen dioxide sensor unit and its long-term operation within a sensor network in the city of Zurich. *Atmos. Meas. Tech.* **2017**, 17.
42. Mead, M.I.; Popoola, O.A.M.; Stewart, G.B.; Landshoff, P.; Calleja, M.; Hayes, M.; Baldovi, J.J.; McLeod, M.W.; Hodgson, T.F.; Dicks, J.; et al. The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmospheric Environment* **2013**, *70*, 186–203.
43. Popoola, O.A.M.; Stewart, G.B.; Mead, M.I.; Jones, R.L. Development of a baseline-temperature correction methodology for electrochemical sensors and its implications for long-term stability. *Atmospheric Environment* **2016**, *147*, 330–343.
44. Mooney, D.; Willis, P.; Stevenson, K. *A Guide for Local Authorities Purchasing Air Quality Monitoring Equipment*; 2006;
45. Laquai, B. Particle Distribution Dependent Inaccuracy of the Plantower PMS5003 low-cost PM-sensor. **2017**.
46. Badura, M.; Batog, P.; Drzeniecka-Osiadacz, A.; Modzel, P. Optical particulate matter sensors in PM2.5 measurements in atmospheric air. *E3S Web of Conferences* **2018**, *44*, 00006.
47. The World Air Quality Index *Sensing the Air Quality: Research on Air Quality Sensors. http://aqicn.org/sensor/.*; 2019;
48. Holstius, D.M.; Pillarisetti, A.; Smith, K.R.; Seto, E. Field calibrations of a low-cost aerosol sensor at a regulatory monitoring site in California. *Atmospheric Measurement Techniques* **2014**, *7*, 1121–1131.
49. Liu, H.-Y.; Schneider, P.; Haugen, R.; Vogt, M. Performance Assessment of a Low-Cost PM2.5 Sensor for a near Four-Month Period in Oslo, Norway. *Atmosphere* **2019**, *10*, 41.
50. Borrego, C.; Costa, A.M.; Ginja, J.; Amorim, M.; Coutinho, M.; Karatzas, K.; Sioumis, Th.; Katsifarakis, N.; Konstantinidis, K.; De Vito, S.; et al. Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise. *Atmospheric Environment* **2016**, *147*, 246–263.
51. Manikonda, A.; Zíková, N.; Hopke, P.K.; Ferro, A.R. Laboratory assessment of low-cost PM monitors. *Journal of Aerosol Science* **2016**, *102*, 29–40.
52. Zikova, N.; Hopke, P.K.; Ferro, A.R. Evaluation of new low-cost particle monitors for PM2.5 concentrations measurements. *Journal of Aerosol Science* **2017**, *105*, 24–34.
53. Sousan, S.; Koehler, K.; Hallett, L.; Peters, T.M. Evaluation of the Alphasense optical particle counter (OPC-N2) and the Grimm portable aerosol spectrometer (PAS-1.108). *Aerosol Science and Technology* **2016**, *50*, 1352–1365.
54. Dacunto, P.J.; Klepeis, N.E.; Cheng, K.-C.; Acevedo-Bolton, V.; Jiang, R.-T.; Repace, J.L.; Ott, W.R.; Hildemann, L.M. Determining PM2.5 calibration curves for a low-cost particle monitor: common indoor residential aerosols. *Environmental Science: Processes & Impacts* **2015**, *17*, 1959–1966.
55. Viana, M.; Rivas, I.; Reche, C.; Fonseca, A.S.; Pérez, N.; Querol, X.; Alastuey, A.; Álvarez-Pedrerol, M.; Sunyer, J. Field comparison of portable and stationary instruments for outdoor urban air exposure assessments. *Atmospheric Environment* **2015**, *123, Part A*, 220–228.
56. Marjovi, A.; Arfire, A.; Martinoli, A. Extending Urban Air Quality Maps Beyond the Coverage of a Mobile Sensor Network: Data Sources, Methods, and Performance Evaluation. In Proceedings of the Proceedings of the 2017 International Conference on Embedded Wireless Systems and Networks; Junction Publishing: USA, 2017; pp. 12–23.
57. 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.
58. 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. *Atmos. Meas. Tech.* **2017**, 14.
59. Wei, P.; Ning, Z.; Ye, S.; Sun, L.; Yang, F.; Wong, K.; Westerdahl, D.; Louie, P. Impact Analysis of Temperature and Humidity Conditions on Electrochemical Sensor Response in Ambient Air Quality Monitoring. *Sensors* **2018**, *18*, 59.
60. Gillooly, S.E.; Zhou, Y.; Vallarino, J.; Chu, M.T.; Michanowicz, D.R.; Levy, J.I.; Adamkiewicz, G. Development of an in-home, real-time air pollutant sensor platform and implications for community use. *Environmental Pollution* **2019**, *244*, 440–450.
61. Zimmerman, N.; Presto, A.A.; Kumar, S.P.N.; Gu, J.; Hauryliuk, A.; Robinson, E.S.; Robinson, A.L.; R. Subramanian A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring. *Atmos. Meas. Tech.* **2018**, *11*, 291–313.
62. Piedrahita, R.; Xiang, Y.; Masson, N.; Ortega, J.; Collier, A.; Jiang, Y.; Li, K.; Dick, R.P.; Lv, Q.; Hannigan, M.; et al. The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. *Atmospheric Measurement Techniques* **2014**, *7*, 3325–3336.
63. 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. *Atmospheric Measurement Techniques* **2018**, *11*, 3717–3735.
64. Cordero, J.M.; Borge, R.; Narros, A. Using statistical methods to carry out in field calibrations of low cost air quality sensors. *Sensors and Actuators B: Chemical* **2018**, *267*, 245–254.
65. Duvall, R.M.; Long, R.W.; Beaver, M.R.; Kronmiller, K.G.; Wheeler, M.L.; Szykman, J.J. Performance Evaluation and Community Application of Low-Cost Sensors for Ozone and Nitrogen Dioxide. *Sensors* **2016**, *16*, 1698.
66. Sun, L.; Westerdahl, D.; Ning, Z. Development and Evaluation of A Novel and Cost-Effective Approach for Low-Cost NO2 Sensor Drift Correction. *Sensors* **2017**, *17*, 1916.
67. Lin, C.; Gillespie, J.; Schuder, M.D.; Duberstein, W.; Beverland, I.J.; Heal, M.R. Evaluation and calibration of Aeroqual series 500 portable gas sensors for accurate measurement of ambient ozone and nitrogen dioxide. *Atmospheric Environment* **2015**, *100*, 111–116.
68. Vaughn, D.L.; Dye, T.S.; Roberts, L.; Ray, A.E.; DeWinter, J.L. *Characterization of Low-Cost NO2 Sensors*; U.S. Environmental Protection Agency Research Triangle Park, North Carolina: North Carolina, 2010; p. 10;.
69. US-EPA Evaluation of Elm and Speck Sensors. **2015**, *EPA/600/R-15/314 | November 2015 | www.epa.gov/ord*.
70. Wang, Y.; Li, J.; Jing, H.; Zhang, Q.; Jiang, J.; Biswas, P. Laboratory Evaluation and Calibration of Three Low-Cost Particle Sensors for Particulate Matter Measurement. *Aerosol Science and Technology* **2015**, *49*, 1063–1077.
71. Alvarado, M.; Gonzalez, F.; Fletcher, A.; Doshi, A.; Alvarado, M.; Gonzalez, F.; Fletcher, A.; Doshi, A. Towards the Development of a Low Cost Airborne Sensing System to Monitor Dust Particles after Blasting at Open-Pit Mine Sites. *Sensors* **2015**, *15*, 19667–19687.
72. Chakrabarti, B.; Fine, P.M.; Delfino, R.; Sioutas, C. Performance evaluation of the active-flow personal DataRAM PM2.5 mass monitor (Thermo Anderson pDR-1200) designed for continuous personal exposure measurements. *Atmospheric Environment* **2004**, *38*, 3329–3340.
73. Olivares, G.; Edwards, S. The Outdoor Dust Information Node (ODIN) – development and performance assessment of a low cost ambient dust sensor. *Atmospheric Measurement Techniques Discussions* **2015**, *8*, 7511–7533.
74. Pillarisetti, A.; Allen, T.; Ruiz-Mercado, I.; Edwards, R.; Chowdhury, Z.; Garland, C.; Hill, L.D.; Johnson, M.; Litton, C.D.; Lam, N.L.; et al. Small, Smart, Fast, and Cheap: Microchip-Based Sensors to Estimate Air Pollution Exposures in Rural Households. *Sensors (Basel)* **2017**, *17*.
75. Austin, E.; Novosselov, I.; Seto, E.; Yost, M.G. Laboratory Evaluation of the Shinyei PPD42NS Low-Cost Particulate Matter Sensor. *PLOS ONE* **2015**, *10*, e0137789.

76. Gao, M.; Cao, J.; Seto, E. A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of PM2.5 in Xi'an, China. *Environmental Pollution* **2015**, *199*, 56–65.
77. Kelly, K.E.; Whitaker, J.; Petty, A.; Widmer, C.; Dybwad, A.; Sleeth, D.; Martin, R.; Butterfield, A. Ambient and laboratory evaluation of a low-cost particulate matter sensor. *Environmental Pollution* **2017**, *221*, 491–500.
78. Karagulian, F.; Belis, C.A.; Lagler, F.; Barbiere, M.; Gerboles, M. Evaluation of a portable nephelometer against the Tapered Element Oscillating Microbalance method for monitoring PM2.5. *J. Environ. Monit.* **2012**, *14*, 2145–2153.
79. Mukherjee, A.; Stanton, L.G.; Graham, A.R.; Roberts, P.T. Assessing the Utility of Low-Cost Particulate Matter Sensors over a 12-Week Period in the Cuyama Valley of California. *Sensors* **2017**, *17*, 1805.
80. Cavaliere, A.; Carotenuto, F.; Di Gennaro, F.; Gioli, B.; Gualtieri, G.; Martelli, F.; Matese, A.; Toscano, P.; Vagnoli, C.; Zaldei, A. Development of Low-Cost Air Quality Stations for Next Generation Monitoring Networks: Calibration and Validation of PM2.5 and PM10 Sensors. *Sensors* **2018**, *18*, 2843.
81. Viana, M.; Rivas, I.; Reche, C.; Fonseca, A.S.; Pérez, N.; Querol, X.; Alastuey, A.; Álvarez-Pedrerol, M.; Sunyer, J. Field comparison of portable and stationary instruments for outdoor urban air exposure assessments. *Atmospheric Environment* **2015**, *123, Part A*, 220–228.
82. Steinle, S.; Reis, S.; Sabel, C.E.; Semple, S.; Twigg, M.M.; Braban, C.F.; Leeson, S.R.; Heal, M.R.; Harrison, D.; Lin, C.; et al. Personal exposure monitoring of PM 2.5 in indoor and outdoor microenvironments. *Science of The Total Environment* **2015**, *508*, 383–394.
83. Han, I.; Symanski, E.; Stock, T.H. Feasibility of using low-cost portable particle monitors for measurement of fine and coarse particulate matter in urban ambient air. *Journal of the Air & Waste Management Association* **2017**, *67*, 330–340.
84. Jovašević-Stojanović, M.; Bartonova, A.; Topalović, D.; Lazović, I.; Pokrić, B.; Ristovski, Z. On the use of small and cheaper sensors and devices for indicative citizen-based monitoring of respirable particulate matter. *Environmental Pollution* **2015**, *206*, 696–704.
85. Sousan, S.; Koehler, K.; Thomas, G.; Park, J.H.; Hillman, M.; Halterman, A.; Peters, T.M. Inter-comparison of low-cost sensors for measuring the mass concentration of occupational aerosols. *Aerosol Science and Technology* **2016**, *50*, 462–473.
86. Crilley, L.R.; Shaw, M.; Pound, R.; Kramer, L.J.; Price, R.; Young, S.; Lewis, A.C.; Pope, F.D. Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring. *Atmospheric Measurement Techniques* **2018**, *11*, 709–720.
87. Zheng, T.; Bergin, M.H.; Johnson, K.K.; Tripathi, S.N.; Shirodkar, S.; Landis, M.S.; Sutaria, R.; Carlson, D.E. Field evaluation of low-cost particulate matter sensors in high and low concentration environments. *Atmospheric Measurement Techniques Discussions* **2018**, 1–40.
88. Liu, H.-Y.; Schneider, P.; Haugen, R.; Vogt, M. Performance Assessment of a Low-Cost PM2.5 Sensor for a near Four-Month Period in Oslo, Norway. *Atmosphere* **2019**, *10*, 41.
89. Esposito, E.; Salvato, M.; Vito, S.D.; Fattoruso, G.; Castell, N.; Karatzas, K.; Francia, G.D. Assessing the Relocation Robustness of on Field Calibrations for Air Quality Monitoring Devices. In Proceedings of the Sensors and Microsystems; Leone, A., Forleo, A., Francioso, L., Capone, S., Siciliano, P., Di Natale, C., Eds.; Springer International Publishing, 2018; pp. 303–312.
90. Eposito, E.; Vito, S.D.; Salvato, M.; Fattoruso, G.; Castell, N.; Karatzas, K.; Francia, G.D. Is on field calibration strategy robust to relocation? In Proceedings of the 2017 ISOCS/IEEE International Symposium on Olfaction and Electronic Nose (ISOEN); 2017; pp. 1–3.
91. BIPM - Guide to the Expression of Uncertainty in Measurement (GUM) Available online: https://www.bipm.org/en/publications/guides/gum.html (accessed on Jul 2, 2019).
92. European Commission *Guide to the demonstration of equivalence of ambient air monitoring methods, Report by an EC Working, Group on Guidance*; European Commission, 2010;
93. Gerboles, M.; Lagler, F.; Rembges, D.; Brun, C. Assessment of uncertainty of NO2

measurements by the chemiluminescence method and discussion of the quality objective of the NO2 European Directive. *Journal of Environmental Monitoring* **2003**, *5*, 529.
94. Thunis, P.; Pederzoli, A.; Pernigotti, D. Performance criteria to evaluate air quality modeling applications. *Atmospheric Environment* **2012**, *59*, 476–482.
95. Barrett, J.P. The Coefficient of Determination—Some Limitations. *The American Statistician* **1974**, *28*, 19–20.
96. Alexander, D.L.J.; Tropsha, A.; Winkler, D.A. Beware of R2: simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models. *J Chem Inf Model* **2015**, *55*, 1316–1322.
97. Piedrahita, R.; Xiang, Y.; Masson, N.; Ortega, J.; Collier, A.; Jiang, Y.; Li, K.; Dick, R.P.; Lv, Q.; Hannigan, M.; et al. The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. *Atmospheric Measurement Techniques* **2014**, *7*, 3325–3336.
98. Wastine, B. *Essai d'Aptitude AirSensEUR du 12-janv au 22-fev 2018 réalisé par Atmo Normandie pour l'exercise d'intercomparaison n°1 du LCSQA, https://db-airmontech.jrc.ec.europa.eu/download/181114_ASE_ICP_1_v4.pdf*; Atmo Normandie, 2019; p. 28;.
99. Wastine, B. *AirSensEur: Point sur les expérimentations menées depuis 2018, https://db-airmontech.jrc.ec.europa.eu/download/181114_ASE_ICP_2_v3.pdf*; Atmo-Normandie, 2019;
100. Di Antonio, A.; Popoola, O.A.M.; Ouyang, B.; Saffell, J.; Jones, R.L. Developing a Relative Humidity Correction for Low-Cost Sensors Measuring Ambient Particulate Matter. *Sensors* **2018**, *18*, 2790.
101. 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**.
102. De Vito, S.; Esposito, E.; Salvato, M.; Popoola, O.; Formisano, F.; Jones, R.; Di Francia, G. Calibrating chemical multisensory devices for real world applications: An in-depth comparison of quantitative machine learning approaches. *Sensors and Actuators B: Chemical* **2018**, *255*, 1191–1210.
103. Northcross, A.L.; Edwards, R.J.; Johnson, M.A.; Wang, Z.-M.; Zhu, K.; Allen, T.; Smith, K.R. A low-cost particle counter as a realtime fine-particle mass monitor. *Environ. Sci.: Processes Impacts* **2013**, *15*, 433–439.