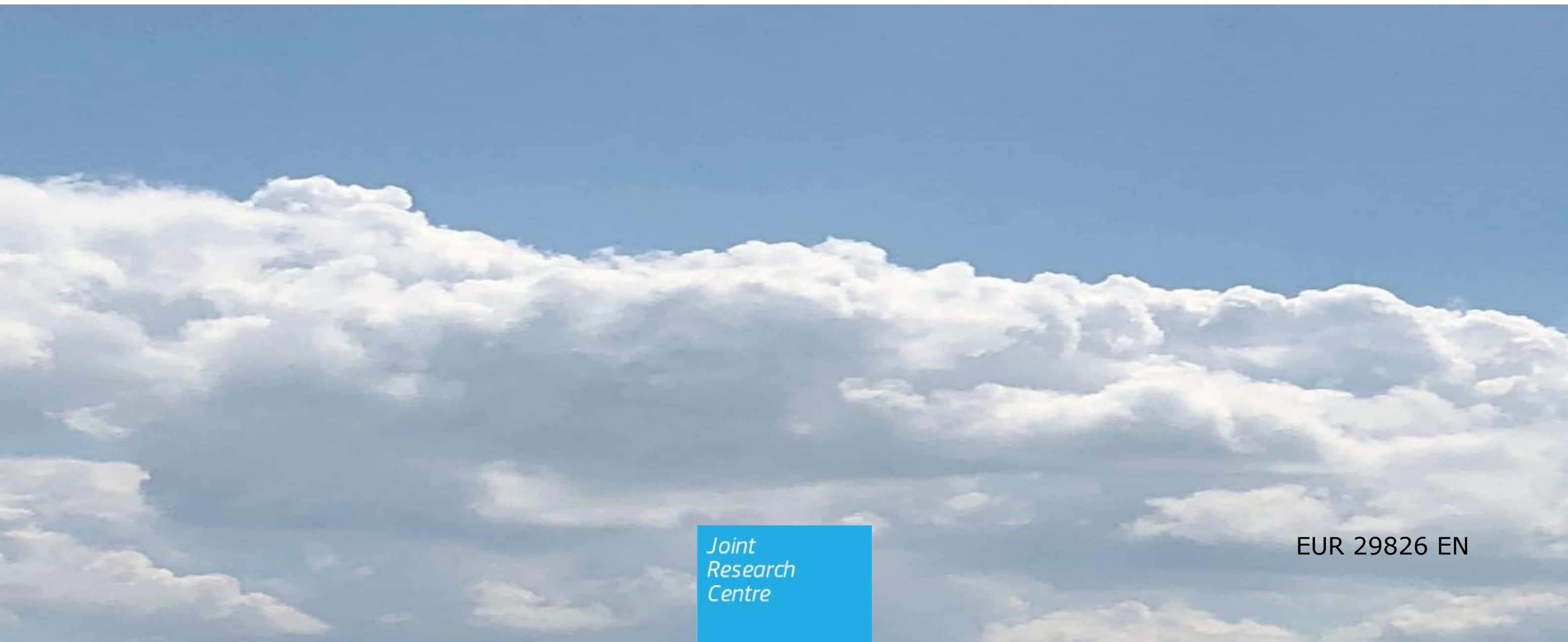
European Commission

JRC TECHNICAL REPORTS

Review of sensors for air quality monitoring

Karagulian, F., Gerboles, M.,
Barbiere, M., Kotsev, A., Lagler, F.,
Borowiak, A.

2019



Joint Research Centre

EUR 29826 EN

This publication is a Technical report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication.

Contact information
Name: Federico Karagulian and Michel Gerboles
Address: European Commission, Joint Research Centre, Via Enrico Fermi, 2749, I - 21027 Ispra (VA) Italy
Email: federico.karagulian@ec.europa.eu; michel.gerboles@ec.europa.eu
Tel.: +39 0332 78 5652

EU Science Hub
https://ec.europa.eu/jrc

JRC116534

EUR 29826 EN

PDF ISBN 978-92-76-09255-1 ISSN 1831-9424 doi:10.2760/568261

Luxembourg: Publications Office of the European Union, 2019

How to cite this report: Karagulian, F., Gerboles, M., Barbiere, M., Kotsev, A., Lagler, F., Borowiak, A., *Review of sensors for air quality monitoring*, EUR 29826 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-09255-1, doi:10.2760/568261, JRC116534.

Contents

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. One benefit of 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, the scientific literature currently reports independent information about the performance of sensor systems against reference measurements for about 110 sensor systems. In fact, the data quality of low-cost sensors is often questionable. It is affected by atmospheric conditions, pollutant concentration levels and therefore by the site location where the measurements are carried out.

This report presents the collected results of quantitative studies of the performance of low-cost sensors against reference measurements. The information collected was an attempt to assess the following issues:

- Agreement between sensor and reference measurements;
- Availability of raw data, transparency of data treatment and possibility of a-posteriori calibration;
- Capability to measure multiple pollutants;
- Affordability of sensor systems taking into consideration the number of provided sensors.

This information gathered in this report comes from research institutes having a LCS testing programme in place, e.g. the California Board - Air Quality Sensor Performance Evaluation Center (AQ-SPEC), the European Union Joint Research Centre (EU JRC) and the United States Environmental Protection Agency (US EPA). Other information was drawn from peer-reviewed journals that tested different types of sensors in research studies. Finally, this information has been linked with scripts able to perform a statistical analysis in the form of an electronic report. This work represents an important review to classify commercial sensors based on their agreement with reference systems.

There are only a few available commercial sensor systems that are consistent with all issues mentioned above that show a good agreement with reference measurements (coefficient of determination, R^2, higher than 0.75 and slope of regression line within 1±0.5) and total price lower than 3 k€. The conclusion from this market analysis is that the only sensor system satisfying the requirements of multipollutant, availability of raw data, transparency of all applied data treatment, availability of evaluation of the performance of sensor system in literature with high coefficient of determination (>0.85) has been found to be the AirSensEUR v.2.

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[45].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 sufficient reference Air Quality Monitoring Stations (AQMS)[27].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[14–18]. 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 including low-cost sensors among their routine methods 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 offers 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[49].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 OEM users need 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[1,12,38,45,63,85,86,91], quantitative data for comparing and evaluating the agreement between LCS and reference data are mostly missing from the existing reviews. Additionally, there is no commonly accepted protocol for testing LCS[87] and the metrics reported are generally diverse making it difficult to compare the performance of LCS in different evaluation studies. Among the available tests of LCS, there are clear indications that the accuracy of LCS measurements can be questionable[3,68] 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[49]. 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 of 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 (O3), nitric dioxide (NO2) and carbon monoxide (CO), the pollutants that are included into the European Union Air Quality Directive[27] 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. In total, 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. Instead, 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)3, the European Union Joint Research Centre (EU JRC)[1,34,41,66–72] and the United States Environmental Protection Agency (US EPA)[79,86–89].

A significant portion of the data comes from the first French field intercomparison exercise (Crunaire23) 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: NO2, O3 and PM2.5/PM10. This exercise involved nearly 5 French laboratories in charge of air pollution monitoring and 10 companies (manufacturers or distributors/sellers), 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 evaluating OEMs and/or SSys and reporting quantitative comparisons of LCS data and reference measurements were identified. They include the Air Quality Egg, Air Quality Station, AirCasting[3,9,31,57], Carnegie Mellon[31,92], CitiSense[89] Cairsense[39], Developer Kit[3], HKEPD/14-02771[74], making-sense.eu[55], communitysensing.org[79], MacPoll.eu[68], OpenSense II[8,56], Proof of Concept AirSensEUR[41], SNAQ Heathrow[54,62]). Out of the *1423* r Records collected from literature, we identified *1192* Records (*201* OEM and *991* SSys) from *90* alive sensors (*25* OEM and *65* SSys) and *231* Records (*119* OEM and *112* SSys) from *22* "non active" (or discontinued) LCSs (*6* OEM and *16* SSys). "Low-cost" refers to the purchase price of LCS[48] compared to the purchase and operating cost of reference analysers[54] for the monitoring of regulated inorganic pollutants and particulate matter that can easily be an order of magnitude more costly. More recently, ultra-affordable OEMs are starting to appear on the market for PM monitoring.[5,11,47]. For the detection of $PM_{2.5}$, some of these sensors are starting to achieve performances comparable to low-cost OEMs manufactured in the Western world. 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[76]. Some of these LCS can achieve similar performance to more expensive OEMs[3,32,37,79,86–89]. The data treatment of LCSs can be classified in two distinct categories:

1. Processing of LCS data performed by "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 234 Records identified comprising 108 OEMs and 126 SSys using open source software for data management. These 234 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 know the complete chain of data treatment. 1189 Records made up of 212 and 977 SSys not using open source software for data treatment were identified. These 1189 Records came from 34 unique LCSs. In most cases, these SSys are pre-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 concentrations using the available counts per bin. 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 principles of operation used by the different types of sensor (electrochemical, metal oxides, optical particulate counter, optical sensors) are reported in a recent work by WMO48. This work also describes observed 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 is a limitation as the SSys might need a-posteriori calibration other than the one provided by the manufacturer, but 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 a commercial supplier selling them. Typically, this type of LCS is used for research and publication and 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.

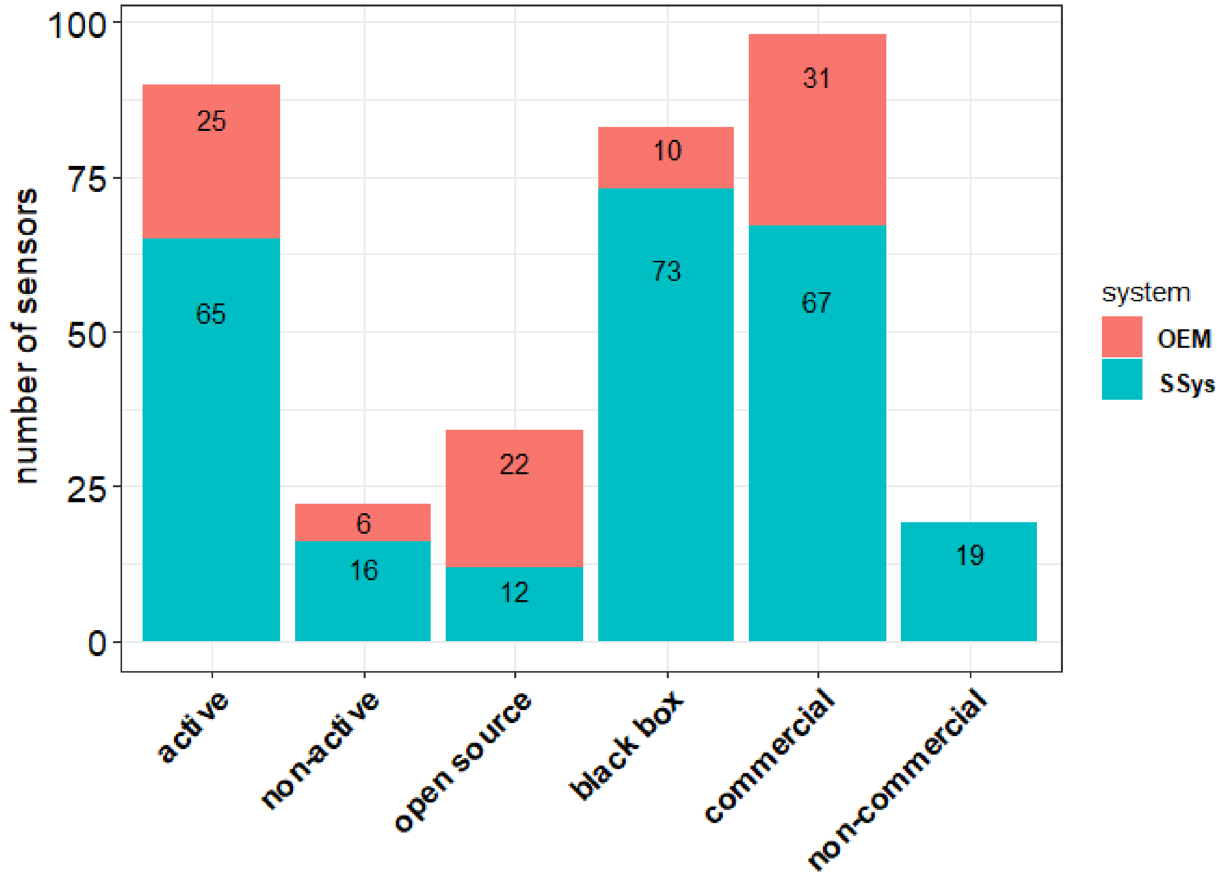


Figure 1. Number of sensor models gathered from the literature review. Sensors has ben classified by their type of technology, availability, openness and commerciality.

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. Agreement between sensor and reference measurements Table 1. Number of analysed records for OEMs/Sensor Systems by pollutant and by type of technology.

pollutant	type	Field/Lab	n. records	references
CO	electrochemical	FIELD	51	AQ-SPEC[3], Jiao[38], Sun[74], Wastine[82], Zimmerman[92], Marjovi[53], Karagulian[41], Popoola[62], Castell[12], Borrego[10], Cross[22], Gillooly[34]
CO	electrochemical	LAB	9	Sun[74], Mead[54], Castell[12], Gerboles[33], Wei[84], Zimmerman[92]
CO	MOs	FIELD	27	AQ-SPEC[3], Spinelle[69], Borrego[10], Piedrahita[60]
CO	MOs	LAB	2	AQ-SPEC[3], Piedrahita[60]
NO	electrochemical	FIELD	44	AQ-SPEC[3], Jiao[38], Bigi[8], Wastine[82], Spinelle[69], Karagulian[41], Mead[54], Popoola[62], Castell[12], Borrego[10], Cross[22], Gillooly[34], LCSQA[47]
NO	electrochemical	LAB	6	Castell[12], Gerboles[33], Wei[84]
NO	MOs	FIELD	1	LCSQA[47]
NO_2	electrochemical	FIELD	137	AQ-SPEC[3], Jiao[38], Sun[74], Mijling[55], Spinelle[68], Mueller[56], Bigi[8], Marjovi[53], Cordero[20], Karagulian[41], Wastine[82], Wastine[83], Mead[54], Popoola[62], Borrego[10], Castell[12], Cross[22], Duvall[27], Gillooly[34], Zimmerman[92], LCSQA[47]
NO_2	electrochemical	LAB	21	Williams[87], Sun[74], Vaughn[79], Castell[12], Spinelle[66], Gerboles[33], Wei[84], Sun[75], Zimmerman[92]
NO_2	MOs	FIELD	28	AQ-SPEC[3], US-EPA[78], Borrego[10], Piedrahita[60], Spinelle[68], Lin[50], LCSQA[47]
NO_2	MOs	LAB	10	Vaughn[79], Williams[87], Piedrahita[60]
O_3	electrochemical	FIELD	65	Jiao[38], Spinelle[68], Mueller[56], Karagulian[41], Wastine[82], AQ-SPEC[3], Borrego[10], Castell[12], Cross[22], Duvall[27], Feinberg[30], LCSQA[47]
O_3	electrochemical	LAB	10	Spinelle[66], Castell[12], Gerboles[33], Wei[84]
O_3	MOs	FIELD	54	AQ-SPEC[3], Jiao[38], Marjovi[53], Borrego[10], Feinberg[30]
O_3	MOs	LAB	3	AQ-SPEC[3], Spinelle[67], Vaughn[79]
O_3	UV	FIELD	9	AQ-SPEC[3]
O_3	UV	LAB	1	Sun[74]
$PM_{2.5}$	Electrical	FIELD	6	AQ-SPEC[3]
$PM_{2.5}$	nephelometer	FIELD	129	Borghi[9], Jiao[38], Feinberg[30], US-EPA[78], Williams[86], AQ-SPEC[3], Zikova[91], Chakrabarti[19], Borrego[10], Olivares[59], Holstius[36], Gao[32], Karagulian[40], LCSQA[47]
$PM_{2.5}$	nephelometer	LAB	24	Manikonda[52], AQ-SPEC[3], Wang[81], Alvarado[2], Sousan[64], Holstius[36], Kelly[42], Austin[4]
$PM_{2.5}$	OPC	FIELD	428	AQ-SPEC[3], Mukherjee[57], Feinberg[30], Jiao[38], Cavaliere[13], Williams[86], Borrego[10], Viana[80], Northcross[58], Holstius[36], Steinle[73], Han[35], Jovasevic[39], Gillooly[34], Sun[74], Dacunto[23],

				Crilley[21], Di-Antonio[25], Badura[5], Pillarisetti[61], Kelly[42], Zheng[90], Laquai[46], Budde[11], Liu[51], LCSQA[47]
$PM_{2.5}$	OPC	LAB	27	AQ-SPEC[3], Cavaliere[13], Manikonda[52], Northcross[58], Sousan[65], Pillarisetti[61], Kelly[42]
PM_1	Electrical	FIELD	6	AQ-SPEC[3]
PM_1	nephelometer	FIELD	1	LCSQA[47]
PM_1	OPC	FIELD	102	AQ-SPEC[3], Williams[86], Crilley[21], Di-Antonio[25], LCSQA[47]
PM_1	OPC	LAB	8	AQ-SPEC[3], Sousan[65]
PM_{10}	nephelometer	FIELD	26	AQ-SPEC[3], Borrego[10], LCSQA[47]
PM_{10}	nephelometer	LAB	1	Alvarado[2]
PM_{10}	OPC	FIELD	176	AQ-SPEC[3], Cavaliere[13], Borrego[10], Feinberg[30], Han[35], Jovasevic[39], Williams[86], Crilley[21], Budde[11], LCSQA[47]
PM_{10}	OPC	LAB	11	AQ-SPEC[3], Cavaliere[13], Manikonda[52], Sousan[64], Sousan[65]

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* 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 NO_2, NO, CO and O_3,the largest number of tests were performed using electrochemical sensors with 343 Records, followed by metal oxides sensors (MOs) 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 OEMs models currently used to monitor Particulate Matter and gaseous pollutants (NO_2, NO, CO and O_3) according to their type of technology. SSys models 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 a few cases, the same model of SSys was tested using different types of OEMs when performing validation tests[34,41] .

"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 fewer OEMs (24) than SSys (65) and therefore different SSys are using the same sets 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. 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 the situation where parameters that have no effect on the sensor data, but that are correlated with other variables that do have an effect, are taken into account in the calibration. The performance of such calibration models can be poor when LCSs are used at sites other 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[29,51,56]. 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 between 2010 and 2014. In 2015, we recorded the largest 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[3], Records were evaluated per model of LCS.

Overall, *34* references reporting field tests with LCSs co-located at urban sites were found, as well as *7* 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 *248* Records for *42* LCSs using daily data. Therefore, hourly data were considered statistically more significant.

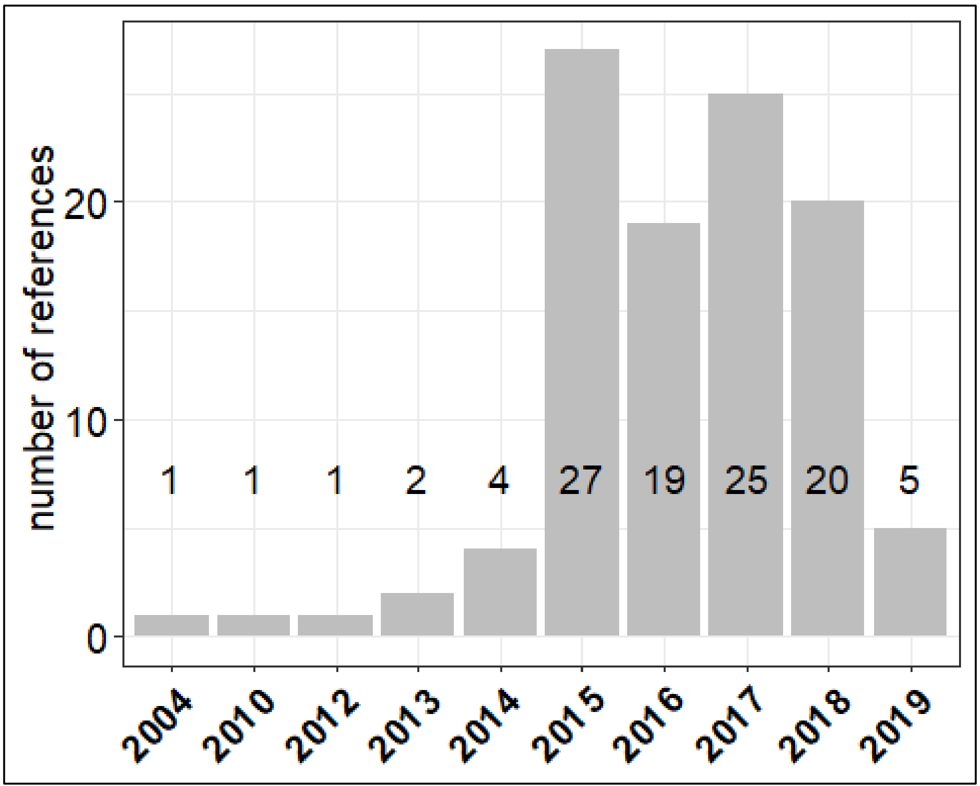


Figure 2. Number of references per year of publication.

3. Method of evaluation

The European Union Air Quality Directive[27] indicates that measurement uncertainty shall be the main indicator used for the evaluation of the data quality objective of air pollution measurement methods[27].However, the evaluation of this metric is cumbersome[30] and it is not included in the majority of sensor studies (see Table 2). For the performance criteria used to evaluate air quality modelling applications[77], the set of statistical indicators includes the Root Mean Square Error (RMSE), the bias, the Standard Deviation (SD) and the correlation coefficient (R), of which RMSE is thought to be the most explicative one. The statistical indicators can be better visualised in a target diagramme[68].Unfortunately, Table 2 also shows that RMSE is mainly unreported in the literature. As already mentioned above, integrated indicators like the IPI[32] would breach our objective to use solely quantitative and comparable indicators. Additionally, it is impossible to compute IPIs a posteriori since time series are mainly not available in literature.

We therefore 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[6]. 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[1] 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 ones.

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 report 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[8,12,20,22,31,35,37,41,51,55,60,68,90] which clearly indicates the magnitude of the error in LCS data unit and is also sensitive to extreme values and outliers. Only a few studies report the measurement uncertainty[12,41,47,68,84,88,93]. 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 metric used in this work.

metrics	n. Field Tests	n. Laboratory Tests
	1290	133
R^2 **from calibrations**	218	60
R^2 **from comparisons**	1164	72
slope of reg. line	1063	55
intercept	1027	54
RMSE	285	5
Uncertainity (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, the reported coefficient of determination R2 was used 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 being close to 1. In any case, 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 and 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 based calibration experiments were found.

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 *1423* Records with details of the 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 the 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. LCS data time-drift was rarely included into the list of calibration covariates[39,60].

Table 3. Types of calibration models used for the calibration of sensors at different time resolutions (ANN: artificial neural network, exp: exponential; log: logarithmic; MLR: multilinear regression; quad: quadratic; RF: random forest; SVM: support vector machine; SVR: support vector regression)

pollutant	calibration model	n. records	references	Median R^2 calib	Median R^2
CO	ANN	2	Wastine[82], Spinelle[69]	NA	0.58
CO	linear	12	Sun[74], Wastine[82], Castell[12], Cross[22], Gerboles[33], Spinelle[69], Zimmerman[92]	0.85	0.15
CO	MLR	21	Jiao[38], Karagulian[41], Wastine[82], Wei[84], Piedrahita[60], Spinelle[69], Zimmerman[92]	0.89	0.83
CO	quad	12	AQ-SPEC[3]	0.63	NA
CO	RF	1	Zimmerman[92]	0.91	NA
NO	ANN	2	Wastine[82], Spinelle[69]	NA	0.57
NO	linear	8	Wastine[82], Castell[12], Cross[22], Spinelle[69], Gerboles[33], LCSQA[47]	0.96	0.032
NO	MLR	20	Jiao[38], Bigi[8], Karagulian[41], Wastine[82], Spinelle[69], Wei[84]	0.92	0.91
NO	RF	2	Bigi[8]	NA	0.9
NO	SVR	2	Bigi[8]	NA	0.90

NO2	ANN	7	Spinelle[68], Cordero[20], Wastine[82], Wastine[83]	0.87	0.94
NO2	linear	25	Sun[74], Vaughn[79], Spinelle[68], Wastine[82], Wastine[83], Castell[12], Cross[22], Gerboles[33], Zimmerman[92], Lin[50], LCSQA[47]	0.29	0.17
NO2	log	1	Vaughn[79]	0.89	NA
NO2	MLR	48	Jiao[38], Sun[74], Mijling[55], Spinelle[68], Mueller[56], Bigi[8], Cordero[20], Karagulian[41], Wastine[82], Wastine[83], Piedrahita[60], Wei[84], Sun[75], Zimmerman[92]	0.81	0.81
NO2	quad	6	AQ-SPEC[3]	0.61	NA
NO2	RF	7	Bigi[8], Cordero[20], Zimmerman[92]	0.86	0.91
NO2	SVM	4	Cordero[20]	0.85	0.94
NO2	SVR	2	Bigi[8]	NA	0.78
O3	ANN	2	Spinelle[68], Wastine[82]	NA	0.89
O3	linear	13	AQ-SPEC[3], Sun[74], Spinelle[68], Wastine[82], Castell[12], Cross[22], Gerboles[33], LCSQA[47]	0.84	0.53
O3	log	1	Vaughn[79]	0.88	NA
O3	MLR	20	Jiao[38], Spinelle[68], Karagulian[41], Wastine[82], Spinelle[67], Wei[84]	0.91	0.88
O3	quad	9	AQ-SPEC[3]	0.72	NA
PM1	Kholer	2	Di-Antonio[25]	NA	0.74
PM1	log	6	AQ-SPEC[3]	0.76	NA
PM10	exp	6	AQ-SPEC[3]	0.59	NA
PM10	Kholer	2	Crilley[21]	NA	NA
PM10	linear	3	AQ-SPEC[3], Cavaliere[13], Jovasevic[39]	0.77	0.63
PM10	log	7	AQ-SPEC[3]	0.58	NA
PM10	quad	1	Alvarado[2]	0.65	NA
PM10-2.5	linear	4	Sousan[64], Han[35], Jovasevic[39]	0.63	0.98
PM2.5	exp	3	Dacunto[23], Kelly[42], Austin[4]	0.91	0.97
PM2.5	Kholer	4	Crilley[21], Di-Antonio[25]	NA	0.78
PM2.5	linear	36	Mukherjee[57], Wang[81], Alvarado[2], Cavaliere[13], Jovasevic[39], Olivares[59], Kelly[42], Zheng[90], Holstius[36]	0.84	0.67
PM2.5	log	7	AQ-SPEC[3], Laquai[46]	0.73	NA
PM2.5	MLR	17	Jiao[38], Sun[74], Zheng[90], Holstius[36], Liu[51]	0.81	0.65
PM2.5	quad	8	Chakrabarti[19], Alvarado[2], Zheng[90], Gao[32]	0.87	0.88
PM2.5	RF	3	Liu[51]	NA	0.79
PM2.5-0.5	linear	9	Northcross[58], Steinle[73], Han[35], Jovasevic[39]	0.84	0.98
PM2.5-0.5	MLR	1	Jiao[38]	0.6	0.45
PM2.5-0.5	quad	6	AQ-SPEC[3], Manikonda[52]	0.82	NA

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 to be expected, 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 worst R^2 for field comparison. Linear calibration should thus be avoided for gas LCSs. For CO and CO, the calibration method giving the highest $R^2$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) either 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[25,40]. 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[25]. However, the number of records is much higher for MLR than for supervised learning techniques. MLR was applied to both MOs sensor and electrochemical sensors which resulted in scattered R² when looking at individual studies. Additionally, supervised learning techniques may be more sensitive to re-location than MLR[29].
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). These 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% above which, humidity is responsible for particle growth[21,26]. 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 ([4,24,43]) with a median R^2 of 98 or using the Kölher theory on PM mass concentration[36] 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, making it difficult to compare the results of the different studies. Calibration of LSC against a reference analyser was found to be carried out using different averaging times. 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 $PM_{2.5}$, R^2 ~ 1 close to 1 were found for hourly data of **PMS1003** by **Plantower**[43] and for the **PMS3003**, **Dylos DC1100 PRO** and **DC1700** by **Dylos** for minute data[3,73,90].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**[3] reported values of R^2 falling within the range of 0.7 - 1.0. The same OEM sensor OPC-N2, reported values of R^2 just above 0.7 when measuing PM_1 while it did not show a good performance when measuring PM_{10}[3].We need to stress that optical sensors, such as OPCs and nephelometers, are somewhat limited in coping with gravity effects 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)[66]. On the other hand, for minute data, values of R^2 close to 1 were found for **AirSensEUR (v.2)** by **LiberaIntentio**[41] as well as for the **S-500** by **Aeroqual**[3] (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[84]), **AirSensEUR (v.2) by LiberaIntentio**[41] and for the minute values **MAS**[74] (see Figure 3). The NO2 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**[34] and **CO-B4** by **Alphasense**[84] reported R^2 ~ 1 for hourly data. High values of R^2 were also reported for the SSys **AirSensEUR (v.2)**, when calibrating CO minute data[41] (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 and **SGX Sensortech**[60], the **Smart Citizen Kit** by **Acrobotic**[3] and the **RAMP**[93]

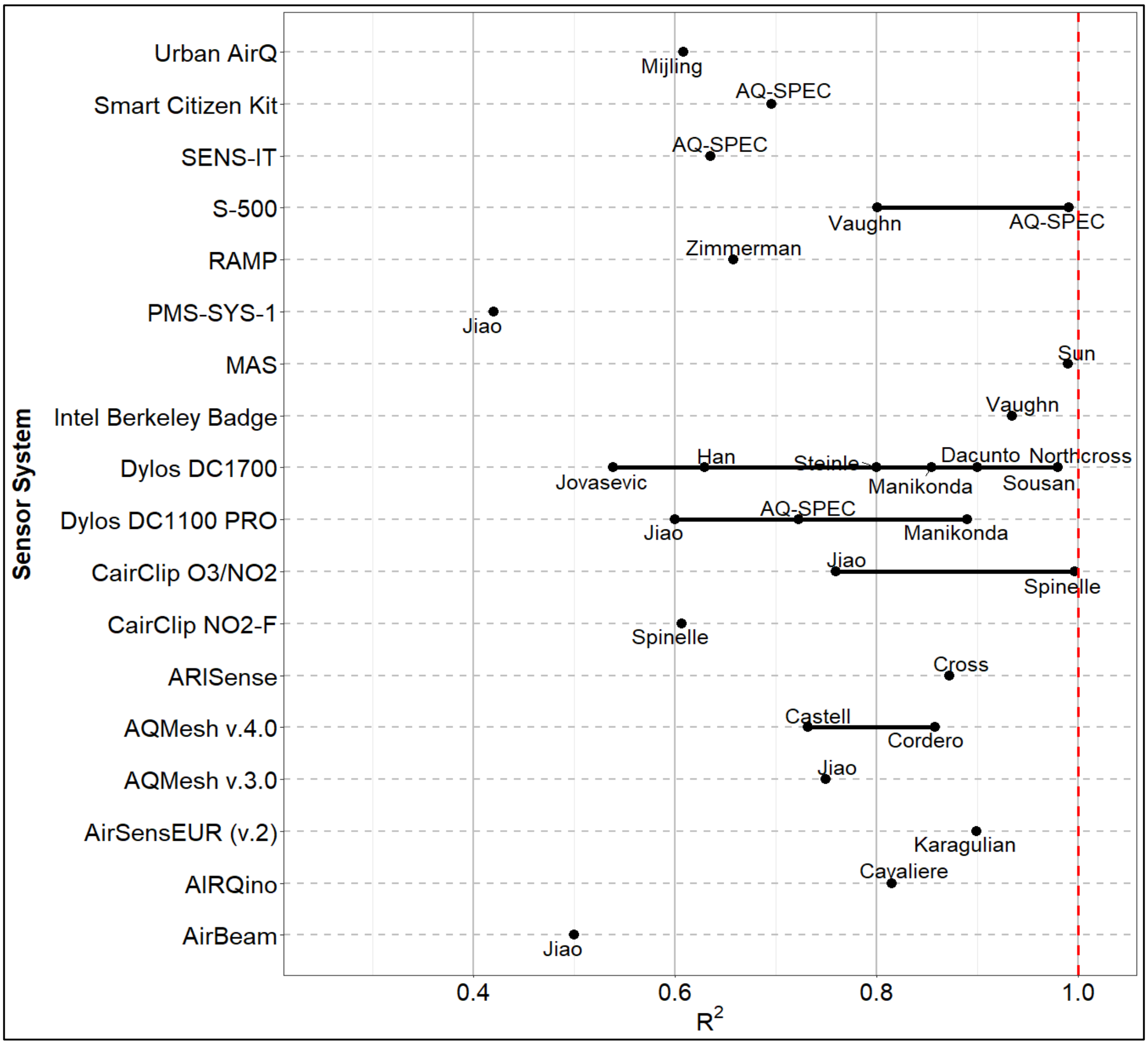


Figure 3. Mean R^2 for obtained from the calibration of sensor systems against reference measurements.

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 $PM_{2.5}$, PM_{10}, PM_1, O_3, NO_2 and CO against reference measurements:

- For the SSys, **PA-II** by **PurpleAir**[3] and **PATS+** by **Belkley Air**[61] 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[23] and, the **Air Quality Station** by **AS LUNG**[3]. Nevertheless, we need to point out that the performance of LCSs measuring PM10, on average, was very poor.
- For the hourly PM measurements of OEMs (Figure A5), the **OPC-N2**, **OPC-N3**[3,5,21,31,57] the **SDS011** by **Nova Fitness**[5] showed R^2 values in the range 0.7 - 1.0. For the daily PM measurements of OEMs (Figure A6), we found R^2 within the range 0.7 - 1.0 for the **OPC-N2**, **OPC-N3**[3].
- For the daily PM measurements of SSys (Figure A7), **PA-II**,[3] **AirQUINO** by **CNR**[13] 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 (v.2)** by **LiberaIntentio**[41], 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**[28,31,68,89], the **Aeroqual Series 500 (and SM50)**[31], the **O3-3E1F** by **CityTechnology**[31,34,68,70] and the **NO2-B43F** by **Alphasense**[75,93]. 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 a SSys, except for PM_{10}.

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** (O_3, NO_2), 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.

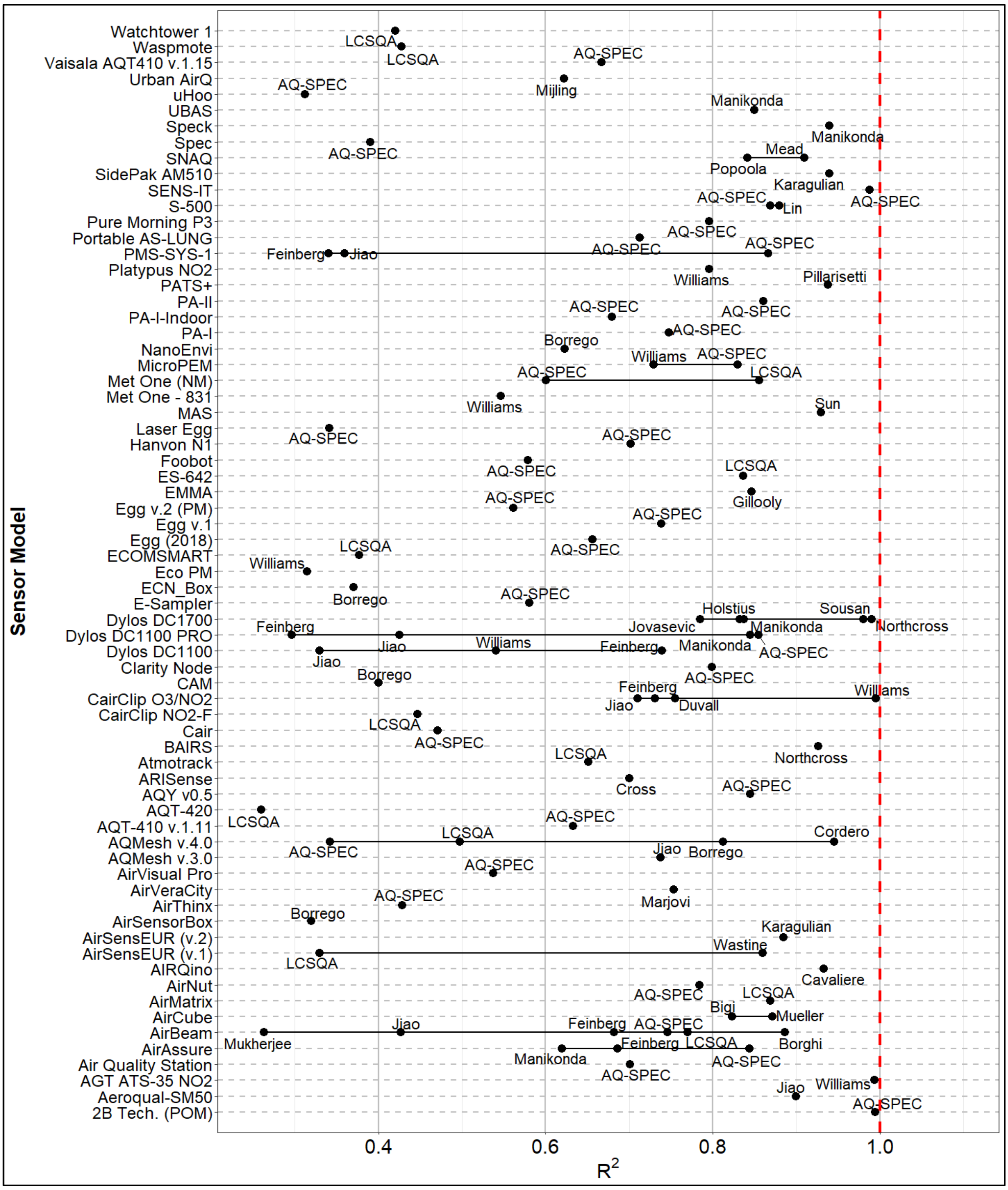


Figure 4. Mean R^2 for obtained from the comparison of sensor systems against reference measurements.

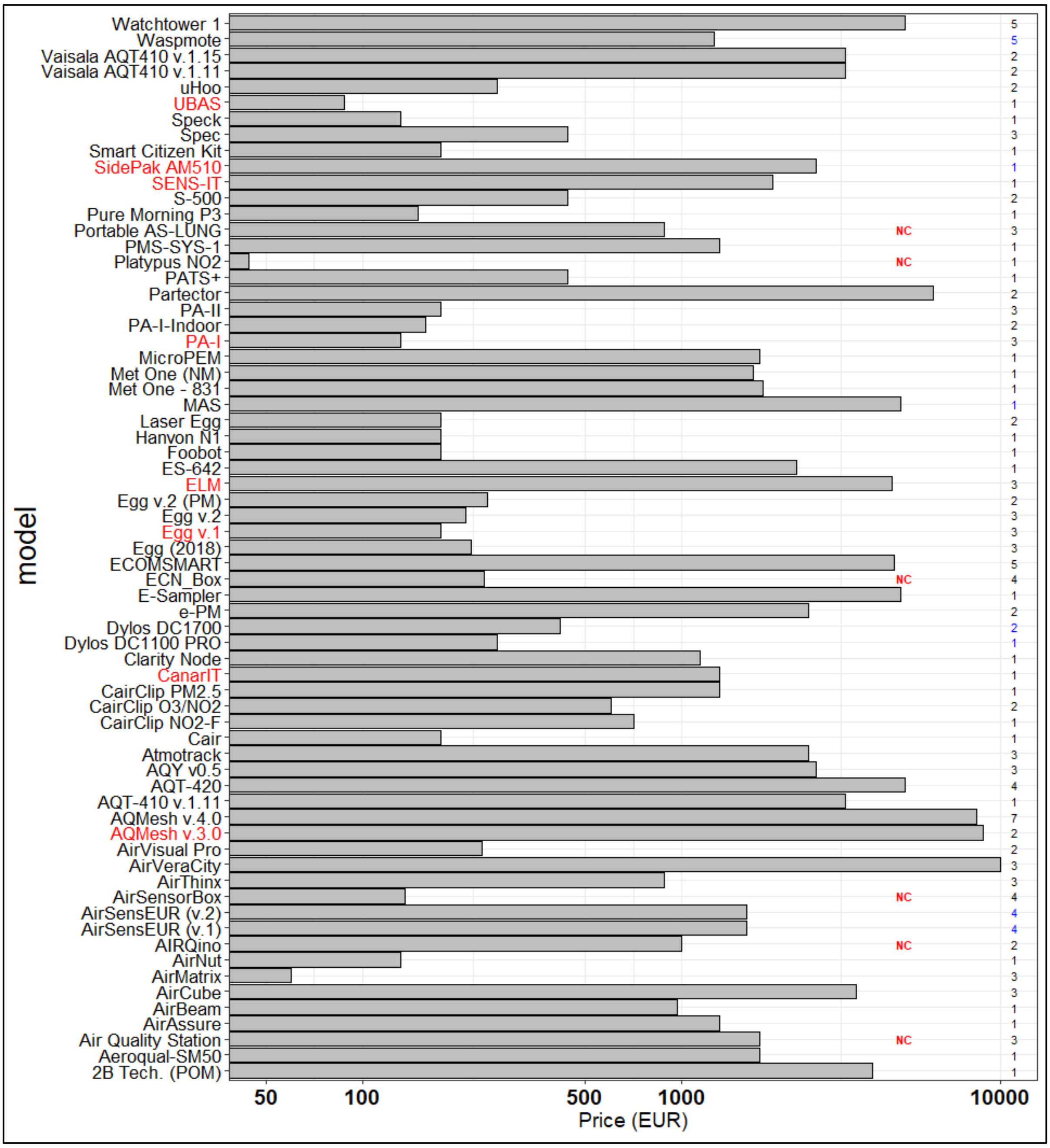


Figure 5. Prices of SS grouped by model. (Numbers in bold indicates the number of pollutants measured by each sensor. x-axis uses logarithmic scale). Numbers in bold indicate the number of open source (blue) and black box (black) records. Names of 'living' & 'updated' and 'non-living' sensors are indicated in black and red colour, respectively. NC indicates non-commercially available sensor.

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 are a large number of SSys measuring one pollutant and only a few 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 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, NO, CO, O_3 and PM). However, only data for PM were available at the time of this review.

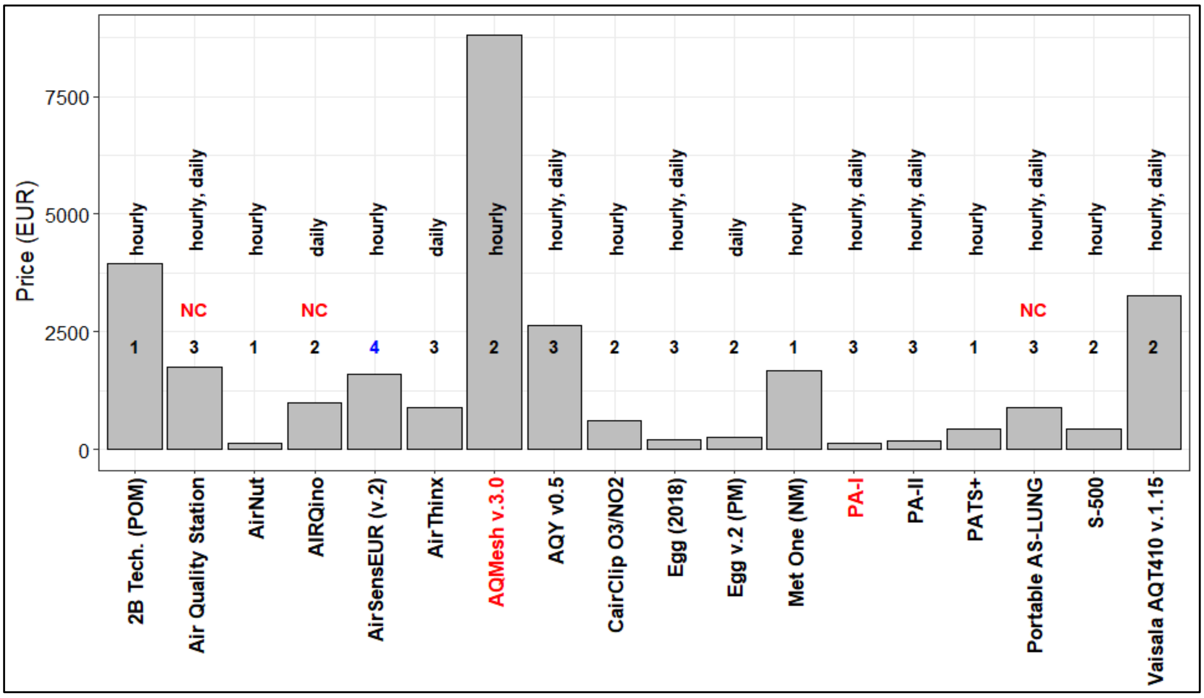


Figure 6. Price of low-cost sensor systems associated to measurements performed at different averaging times. Numbers in bold indicate the number of pollutants measured by open source (blue) and black box (black) sensors. Only records with R^2 > 0.85 and 0.8 < $slope$ < 1.2 are shown. Names of 'living' & 'updated' and 'non-living' sensors are indicated in black and red colour, respectively. NC indicates non-commercially available sensor.

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

Model	**pollutant**	**mean R^2**	**mean slope**	**mean 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	0.87	1	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	black box	Y	non commercial	880
AirSensEUR (v.2)	NO_2, O_3, CO, NO	0.91	0.98	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	0.9	0.85	black box	Y	non commercial	1760
AQY v0.5	$PM_{2.5}$	0.87	0.97	4	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	1	0.74	black box	Y	commercial	3960
AQMesh v.3.0	NO	0.87	0.88	0.76	black box	N	commercial	8800

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 increase in 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.

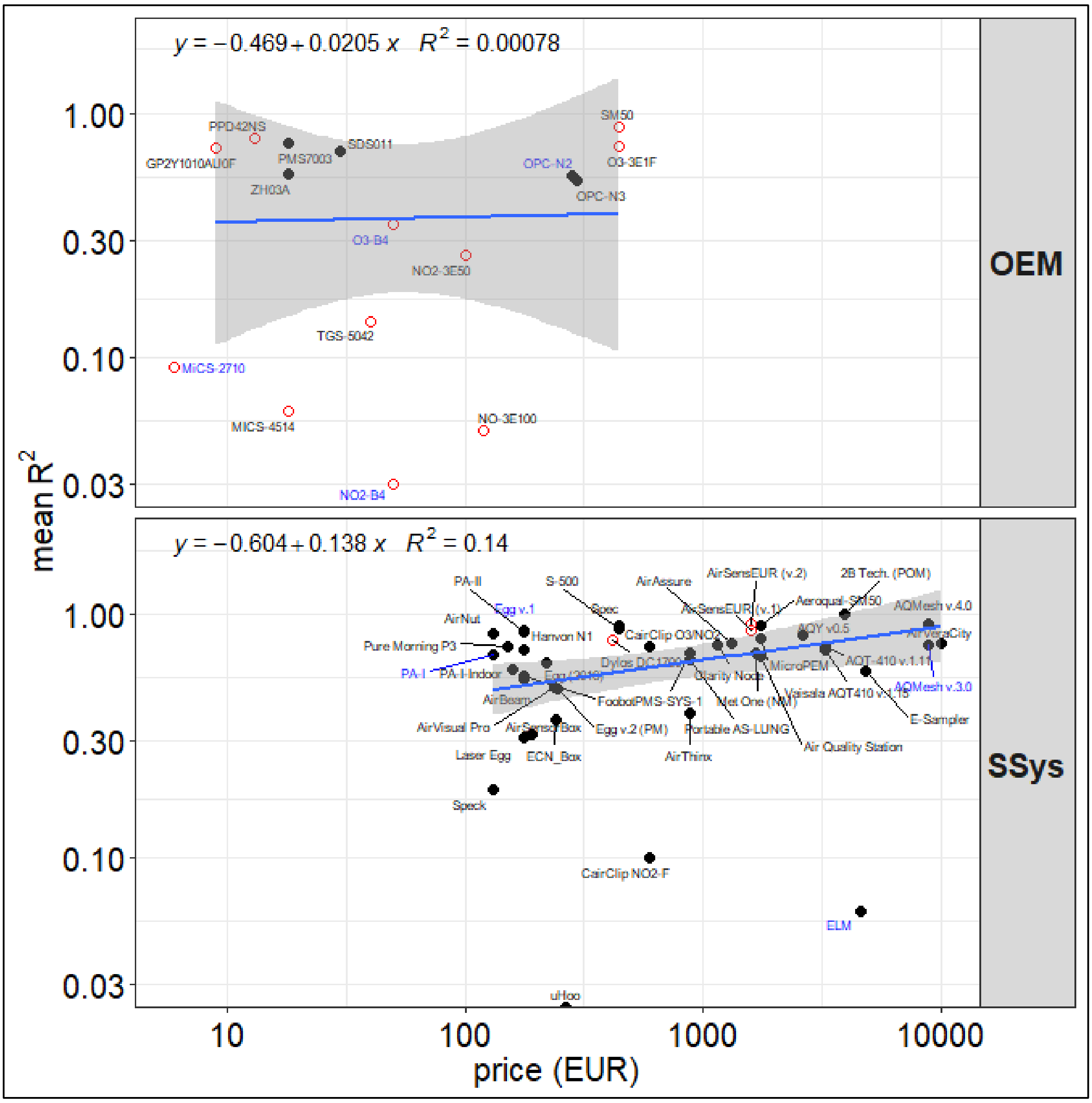


Figure 7. Relation between prices of OEMs/Sensor Systems (SS) and R^2 for field test only. Logarithmic scale has been set for both axis. Open source and black box models are indicated with open and full circles, respectively. Names of 'living' and 'non-living' sensors are indicated in black and blue colour, 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

According to the European Air Quality Directive, a sensor system can be considered "Equivalent" when it meets the Data Quality Objectives (DQOs) set for data capture and uncertainty. In order for sensor system measurement to be incorporated into the legal framework set by the Air Quality Directive in Europe, they shall satisfy one of the data quality objectives (DQOs) of the Directive. DQOs, defined as the maximum allowed relative uncertainty, are defined either for reference and indicative measurements or for objective estimations. For inorganic gaseous pollutants, they correspond to 15, 25 to 30 and 75 %, respectively. Although, the objective of sensor systems is to provide the most accurate air pollution measurements, it is most likely that the DQO for reference measurements is out of reach while it is believed that by improving the sensor calibration procedures the DQO of "Indicative Measurements" could be met at fixed monitoring sites.

There is little information available in the literature regarding calibration of LCSs. Nevertheless, it was possible to list the calibration methods giving the highest R^2 when applied to the results of field tests. For CO and NO our review 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 $PM_{2.5}$ when high relative humidity data (> 75-80%) 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 R^2, 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.

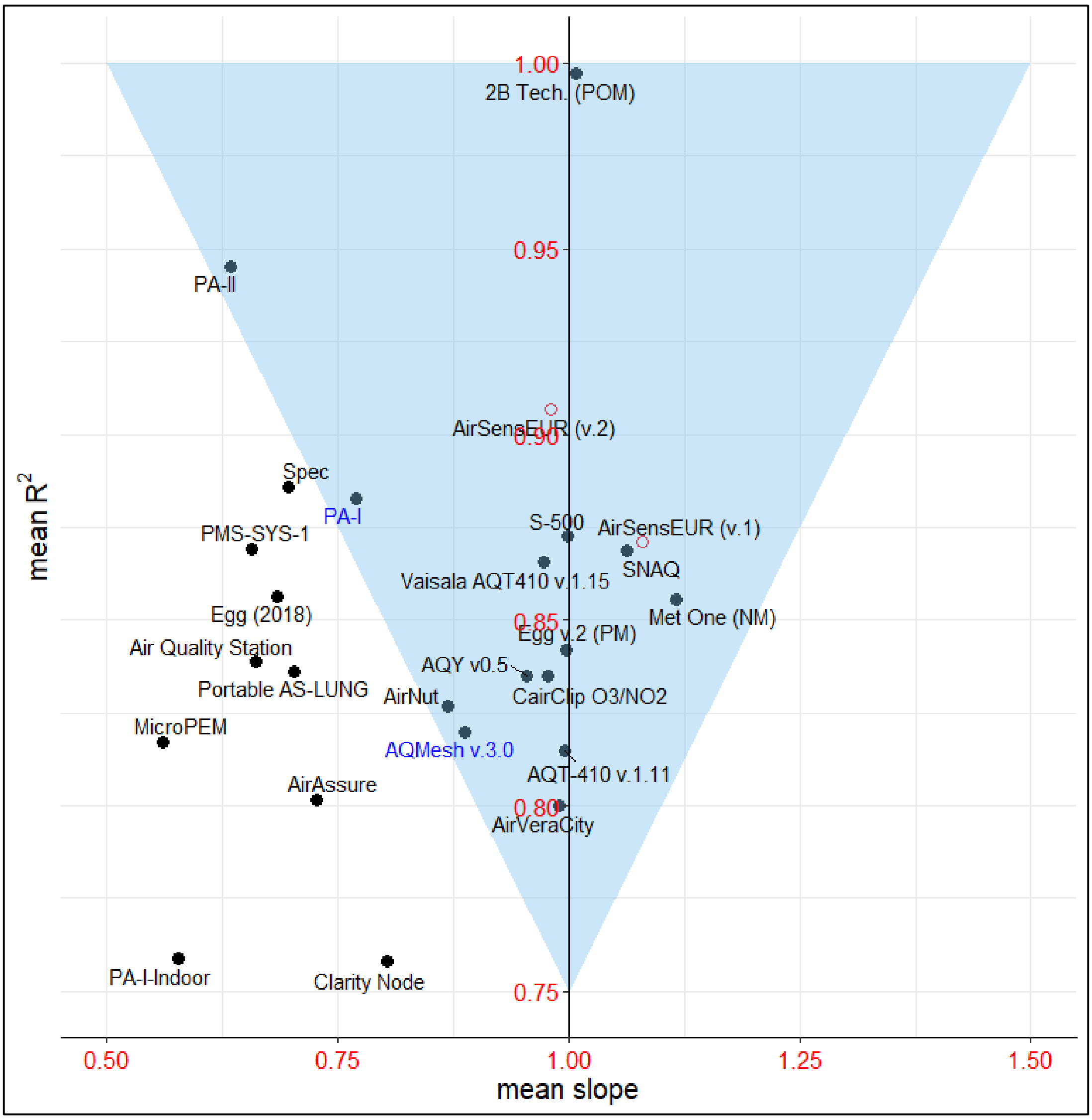


Figure 8. Correspondence between R^2 and slope for sensor systems (SS) for 1 hour averaging time. Only sensor models with $mean\ R^2$ > 0.75 and 0.5 < $mean\ slope$ < 1.2 are shown. Names of 'living' and 'non-living' sensors are indicated in black and blue color, respectively.

Although this report gives an exhaustive survey of the independent LCS evaluations, the concept of comparing LCS field tests from different studies can be difficult and may result in misleading conclusions. It is difficult because of the lack of uniformity in the metrics representing LCS data quality between studies makes them 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 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 for evaluation of LCSs at international level a high priority, while intercomparison exercises where LCSs are gathered at the same test sites and at the same time are also greatly needed.

Finally, among open source sensor systems we could identify the **AirSensEUR (v.2)** by **LiberaIntentio** and the **AIRQuino** by the **CNR** for the detection of NO_2, CO, O_3, NO and PM, respectively. As we can see, 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, the **AIRQuino** can measure up to five pollutants

($PM_{2.5}$, PM_{10}, NO_2, O_3, CO and NO, CO_2 and $VOCs$), however, only data fror PM were available at the time of this review. On the other hand, the **AirSensEUR (v2)** is a complete sensor system that can also measure PM beside gaseous pollutants including "CO_2 and Rn (radon)". This sensor system is already operative and has undergone multiple calibrations and field tests where measurements of gaseous pollutants showing good agreement with reference measurements.

To conclude this market analysis, the only sensor system satisfying the requirements of multipollutant, availability of raw data, transparency of all applied data treatment, availability of evaluation of the performance of sensor system in literature with high coefficient of determination (>0.85) has been found to be the AirSensEUR v.2.

Appendix A

Table A1. Number of analysed records and sensor models by averaging time.

Averaging time	n. records	n. OEMs & SS
1 hour	610	86
5 min	253	40
24 hour	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[84]	open source	Y	50
CO 3E300	CO	electrochemical	Gerboles[33]	open source	Y	100
DataRAM pDR-1200	PM2.5	nephelometer	Chakrabarti[19]	black box	N	
DiscMini	PM	OPC	Viana[80]	open source	Y	11000
DN7C3CA006	PM2.5	nephelometer	Sousan[64]	open source	Y	10
DSM501A	PM2.5	nephelometer	Wang[81], Alvarado[2]	open source	Y	15
FIS SP-61	O3	MOs	Spinelle[67]	open source	Y	50
GP2Y1010AU0F	PM2.5, PM10	nephelometer	Olivares[59], Manikonda[52], Sousan[64], Alvarado[2], Wang[81]	open source	Y	10
MiCS-2710	NO2	MOs	Spinelle[68], Williams[87]	open source	N	7
MICS-4514	CO, NO2	MOs	Spinelle[69], Spinelle[68]	open source	Y	20
NO-3E100	NO	electrochemical	Spinelle[69], Gerboles[33]	open source	Y	120
NO-B4	NO	electrochemical	Wei[84]	open source	Y	50
NO2-3E50	NO2	electrochemical	Spinelle[68], Spinelle[66], Gerboles[33]	open source	Y	100
NO2-A1	NO2	electrochemical	Williams[87]	black box	Y	50
NO2-B4	NO2	electrochemical	Spinelle[66], Spinelle[68]	open source	N	50
NO2-B42F	NO2	electrochemical	Wei[84]	open source	N	50
NO2-B43F	NO2	electrochemical	Sun[75]	open source	Y	50
O3-3E1F	O3	electrochemical	Spinelle[66], Spinelle[68], Gerboles[33]	open source	Y	500
O3-B4	O3	electrochemical	Spinelle[66], Spinelle[68], Wei[84]	open source	N	50
OPC-N2	PM1, PM2.5, PM10	OPC	AQ-SPEC[3], Mukherjee[57], Sousan[65], Feinberg[30], Crilley[21], Di-	black box, open source	N	310

			Antonio[25], Badura[5], LCSQA[47]			
OPC-N3	PM1, PM2.5, PM10	OPC	AQ-SPEC[3]	black box	Y	338
PMS1003	PM2.5	OPC	Kelly[42]	black box	Y	20
PMS3003	PM2.5	OPC	Zheng[90], Kelly[42]	black box	Y	30
PMS5003	PM2.5	OPC	Laquai[46]	black box	Y	15
PMS7003	PM2.5	OPC	Badura[5]	black box	Y	20
PPD42NS	PM2.5, PM3, PM2	nephelometer	Wang[81], Holstius[36], Gao[32], Kelly[42], Austin[4]	open source	Y	15
SDS011	PM2.5, PM10	OPC	Budde[11], Badura[5], Liu[51], Laquai[46]	black box	Y	30
SM50	O3	MOs	Feinberg[30]	open source	Y	500
TGS-5042	CO	MOs	Spinelle[69]	open source	Y	40
TZOA-PM Research Sensors	PM	nephelometer	Feinberg[30]	open source	Y	90
ZH03A	PM2.5	OPC	Badura[5]	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)	O3	UV	AQ-SPEC[3]	black box	Y	4500
Aeroqual-SM50	O3	MOs	Jiao[38]	black box	Y	2000
AGT ATS-35 NO2	NO2	MOs	Williams[87]	black box	N	
Air Quality Station	PM1, PM2.5, PM10	OPC	AQ-SPEC[3]	black box	Y	2000
AirAssure	PM2.5	nephelometer	Feinberg[30], Manikonda[52], AQ-SPEC[3]	black box	Y	1500
AirBeam	PM2.5	OPC, nephelometer	Mukherjee[57], Feinberg[30], Borghi[9], Jiao[38], AQ-SPEC[3], LCSQA[47]	black box	Y	200
AirCube	NO2, O3, NO	electrochemical	Mueller[56], Bigi[8]	black box	Y	3538
AirMatrix	PM1, PM10, PM2.5	nephelometer	LCSQA[47]	black box	Y	60
AirNut	PM2.5	OPC	AQ-SPEC[3]	black box	Y	150
AIRQino	PM2.5, PM10	OPC	Cavaliere[13]	black box	Y	1000
AirSensEUR (v.1)	CO, NO, NO2, O3	electrochemical	Wastine[82], Wastine[83],	open source, black box	Y	1600

			LCSQA[47]			
AirSensEUR (v.2)	CO, NO, NO2, O3	electrochemical	Karagulian[41]	open source	Y	1600
AirSensorBox	NO2, CO, O3, PM10	electrochemical, MOs, nephelometer	Borrego[10]	black box	Y	280
AirThinx	PM1, PM2.5, PM10	OPC	AQ-SPEC[3]	black box	Y	1000
AirVeraCity	CO, NO2, O3	electrochemical, MOs	Marjovi[53]	black box	Y	10000
AirVisual Pro	PM2.5, PM10	OPC	AQ-SPEC[3]	black box	Y	270
AQMesh v.3.0	CO, NO	electrochemical	Jiao[38]	black box	N	10000
AQMesh v.4.0	CO, NO2, NO, O3, PM1, PM10, PM2.5	electrochemical, OPC	AQ-SPEC[3], Cordero[20], Castell[12], Borrego[10], LCSQA[47]	black box	Y	10000
AQT-410 v.1.11	O3	electrochemical	AQ-SPEC[3]	black box	Y	3700
AQT-420	NO2, O3, PM10, PM2.5	electrochemical, OPC	LCSQA[47]	black box	Y	5000
AQY v0.5	PM2.5, NO2, O3	OPC, electrochemical, MOs	AQ-SPEC[3]	black box	updated	3000
ARISense	NO2, CO, NO, O3	electrochemical	Cross[22]	black box	Y	
Atmotrack	PM1, PM10, PM2.5	OPC	LCSQA[47]	black box	Y	2500
BAIRS	PM2.5-0.5	OPC	Northcross[58]	open source	N	475
Cair	PM2.5, PM10-2.5	OPC	AQ-SPEC[3]	black box	Y	200
CairClip NO2-F	NO2	electrochemical	Spinelle[66], Spinelle[68], Duvall[27], LCSQA[47]	black box	Y	600
CairClip O3/NO2	O3, NO2	electrochemical	Jiao[38], Spinelle[66], Williams[87], Duvall[27], Feinberg[30]	black box	Y	600
CairClip PM2.5	PM2.5	OPC	Williams[86]	black box	Y	1500
CAM	PM10, PM2.5, NO2, CO, NO	OPC, electrochemical	Borrego[10]	black box	Y	
CanarIT	PM	OPC	Williams[86]	black box	N	1500
Clarity Node	PM2.5	OPC	AQ-SPEC[3]	black box	Y	1300
Dylos DC1100	PM2.5-0.5	OPC	Jiao[38], Williams[86], Feinberg[30]	black box, open source	Y	300
Dylos DC1100 PRO	PM2.5-0.5, PM10-2.5,	OPC	AQ-SPEC[3], Jiao[38], Feinberg[30],	black box, open source	Y	300

	PM10		Manikonda[52]			
Dylos DC1700	PM2.5-0.5, PM10, PM10-2.5, PM3, PM2, PM2.5	OPC	Manikonda[52], Sousan[64], Northcross[58], Holstius[36], Steinle[73], Han[35], Jovasevic[39], Dacunto[23]	open source	Y	475
e-PM	PM10, PM2.5	OPC	LCSQA[47]	black box	Y	2500
E-Sampler	PM2.5	OPC	AQ-SPEC[3]	black box	Y	5500
ECN_Box	PM10, PM2.5, NO2, O3	nephelometer, electrochemical	Borrego[10]	black box	Y	274
Eco PM	PM1	OPC	Williams[86]	black box	N	
ECOMSMART	NO2, O3, PM1, PM10, PM2.5	electrochemical, OPC	LCSQA[47]	black box	Y	4650
Egg (2018)	PM1, PM2.5, PM10	OPC	AQ-SPEC[3]	black box	Y	249
Egg v.1	CO, NO2, O3	MOs	AQ-SPEC[3]	black box	N	200
Egg v.2	CO, NO2, O3	electrochemical	AQ-SPEC[3]	black box	Y	240
Egg v.2 (PM)	PM2.5, PM10	nephelometer	AQ-SPEC[3]	black box	Y	280
ELM	NO2, PM10, O3	MOs, nephelometer	AQ-SPEC[3], US-EPA[78]	black box	N	5200
EMMA	PM2.5, CO, NO2, NO	OPC, electrochemical	Gillooly[34]	black box	Y	
ES-642	PM2.5	OPC	LCSQA[47]	black box	Y	2600
Foobot	PM2.5	OPC	AQ-SPEC[3]	black box	Y	200
Hanvon N1	PM2.5	nephelometer	AQ-SPEC[3]	black box	Y	200
Intel Berkeley Badge	NO2, O3	electrochemical, MOs	Vaughn[79]	open source	N	
ISAG	NO2, O3	MOs	Borrego[10]	black box	N	
Laser Egg	PM2.5, PM10	nephelometer	AQ-SPEC[3]	black box	Y	200
M-POD	CO, NO2	MOs	Piedrahita[60]	black box	N	
MAS	CO, NO2, O3, PM2.5	electrochemical, UV, OPC	Sun[74]	black box, open source	N, Y	5500
Met One - 831	PM10	OPC	Williams[86]	black box	Y	2050
Met One (NM)	PM2.5	OPC	AQ-SPEC[3], LCSQA[47]	black box	Y	1900
MicroPEM	PM2.5	OPC	AQ-SPEC[3], Williams[86]	black box	Y	2000
NanoEnvi	NO2, O3, CO	electrochemical, MOs	Borrego[10]	black box	Y	
PA-I	PM1, PM2.5,	OPC	AQ-SPEC[3]	black box	N	150

	PM10					
PA-I-Indoor	PM2.5, PM10	OPC	AQ-SPEC[3]	black box	Y	180
PA-II	PM1, PM2.5, PM10	OPC	AQ-SPEC[3]	black box	Y	200
Partector	PM1, PM2.5	Electrical	AQ-SPEC[3]	black box	Y	7000
PATS+	PM2.5	OPC	Pillarisetti[61]	black box	Y	500
Platypus NO2	NO2	MOs	Williams[87]	black box	Y	50
PMS-SYS-1	PM2.5	nephelometer	Jiao[38], AQ-SPEC[3], Williams[86], Feinberg[30]	black box	Y	1000
Portable AS-LUNG	PM1, PM2.5, PM10	OPC	AQ-SPEC[3]	black box	Y	1000
Pure Morning P3	PM2.5	OPC	AQ-SPEC[3]	black box	Y	170
RAMP	CO, NO2	electrochemical	Zimmerman[92]	open source	Y	
S-500	O3, NO2	MOs	AQ-SPEC[3], Lin[50], Vaughn[79]	black box	Y	500
SENS-IT	O3, CO, NO2	MOs	AQ-SPEC[3]	black box	N, Y	2200
SidePak AM510	PM2.5	nephelometer	Karagulian[40]	open source	N	3000
Smart Citizen Kit	CO	MOs	AQ-SPEC[3]	black box	Y	200
SNAQ	NO2, CO, NO	electrochemical	Mead[54], Popoola[62]	black box	Y	
Spec	CO, NO2, O3	electrochemical	AQ-SPEC[3]	black box	Y	500
Speck	PM2.5	nephelometer	Feinberg[30], US-EPA[78], Williams[86], Manikonda[52], Zikova[91], AQ-SPEC[3]	black box	Y	150
UBAS	PM2.5	nephelometer	Manikonda[52]	black box	N	100
uHoo	PM2.5, O3	nephelometer, MOs	AQ-SPEC[3]	black box	Y	300
Urban AirQ	NO2	electrochemical	Mijling[55]	open source	N	
Vaisala AQT410 v.1.11	CO, NO2	electrochemical	AQ-SPEC[3]	black box	Y	3700
Vaisala AQT410 v.1.15	CO, NO2	electrochemical	AQ-SPEC[3]	black box	Y	3700
Waspmote	NO, NO2, PM1, PM10, PM2.5	MOs, OPC	LCSQA[47]	open source	Y	1270
Watchtower 1	NO2, PM1, PM10, PM2.5, O3	electrochemical, OPC	LCSQA[47]	black box	Y	5000

Table A4. Shortlist of sensor systems showing good agreement with reference systems (R^2 > 0.85; 0.8 < slope < 1.2) for daily data.

model	pollutant	mean R^2	mean slope	mean intercept	open/close	living	commercial	price (EUR)
PA-I	PM_1	0.99	0.91	0.47	black box	N	commercial	132
PA-II	PM_1	0.99	0.83	1.8	black box	Y	commercial	176
Egg (2018)	PM_1	0.88	0.81	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	0.85	1.3	black box	Y	commercial	880
Portable AS-LUNG	PM_1	0.93	0.88	1.5	black box	Y	non commercial	880
AIRQino	$PM_{2.5}$, PM_{10}	0.93	1	1.1	black box	Y	non commercial	1000
Air Quality Station	PM_1	0.94	0.89	1.1	black box	Y	non commercial	1760
AQY v0.5	$PM_{2.5}$	0.91	0.94	4	black box	updated	commercial	2640
Vaisala AQT410 v.1.15	CO	0.86	0.91	0.25	black box	Y	commercial	3256

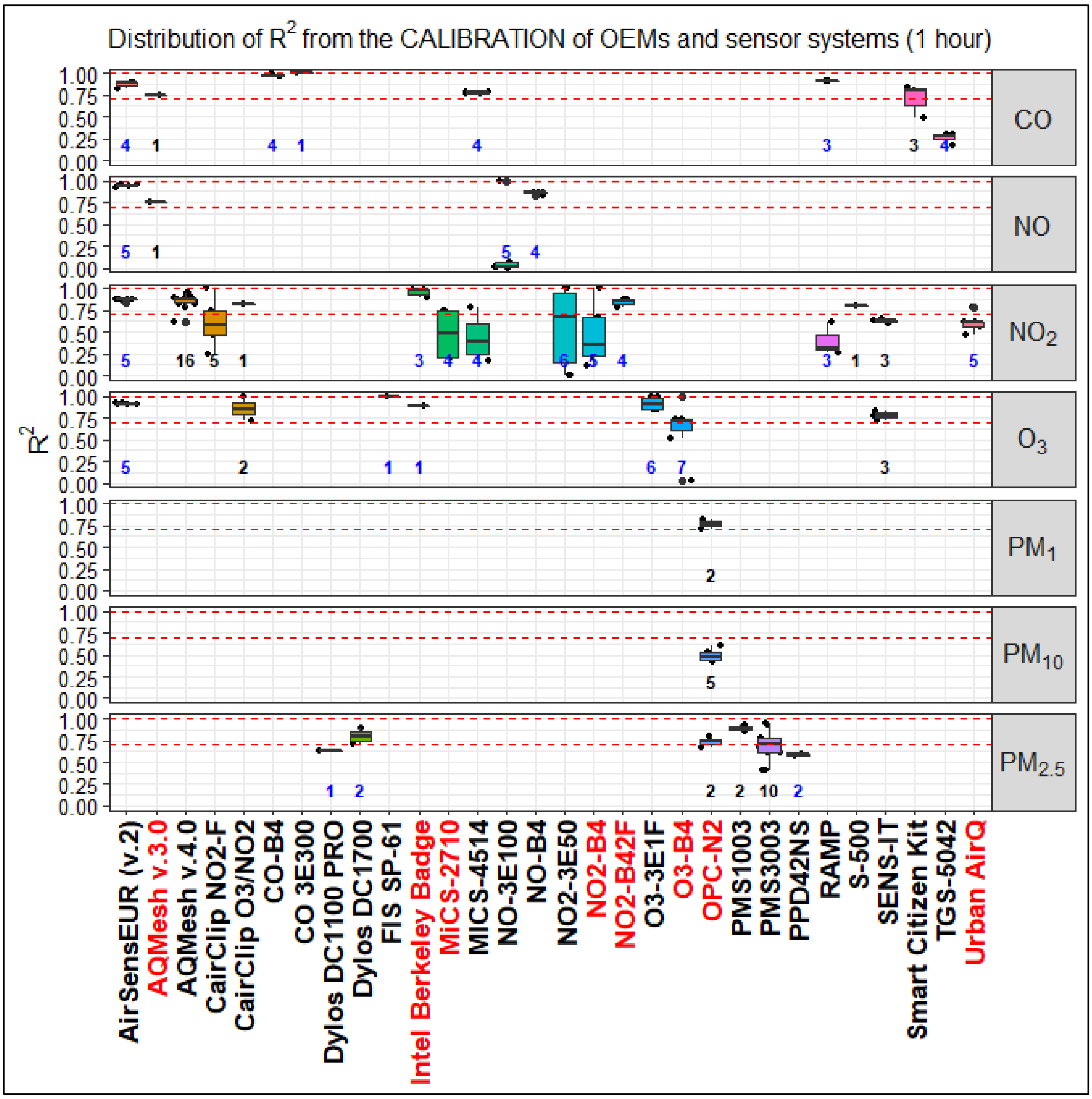


Figure A1. Distribution of R^2 for OEMs and sensor systems obtained from the calibration against the reference. Records were averaged over a time-scale of 1 hour. Dashed lines indicate the value of 0.7 and 1.0. Numbers in bold indicate the number of open source (blue) and black box (black) records. Names of 'living' and 'non-living' sensors are indicated in black and red colour, respectively.

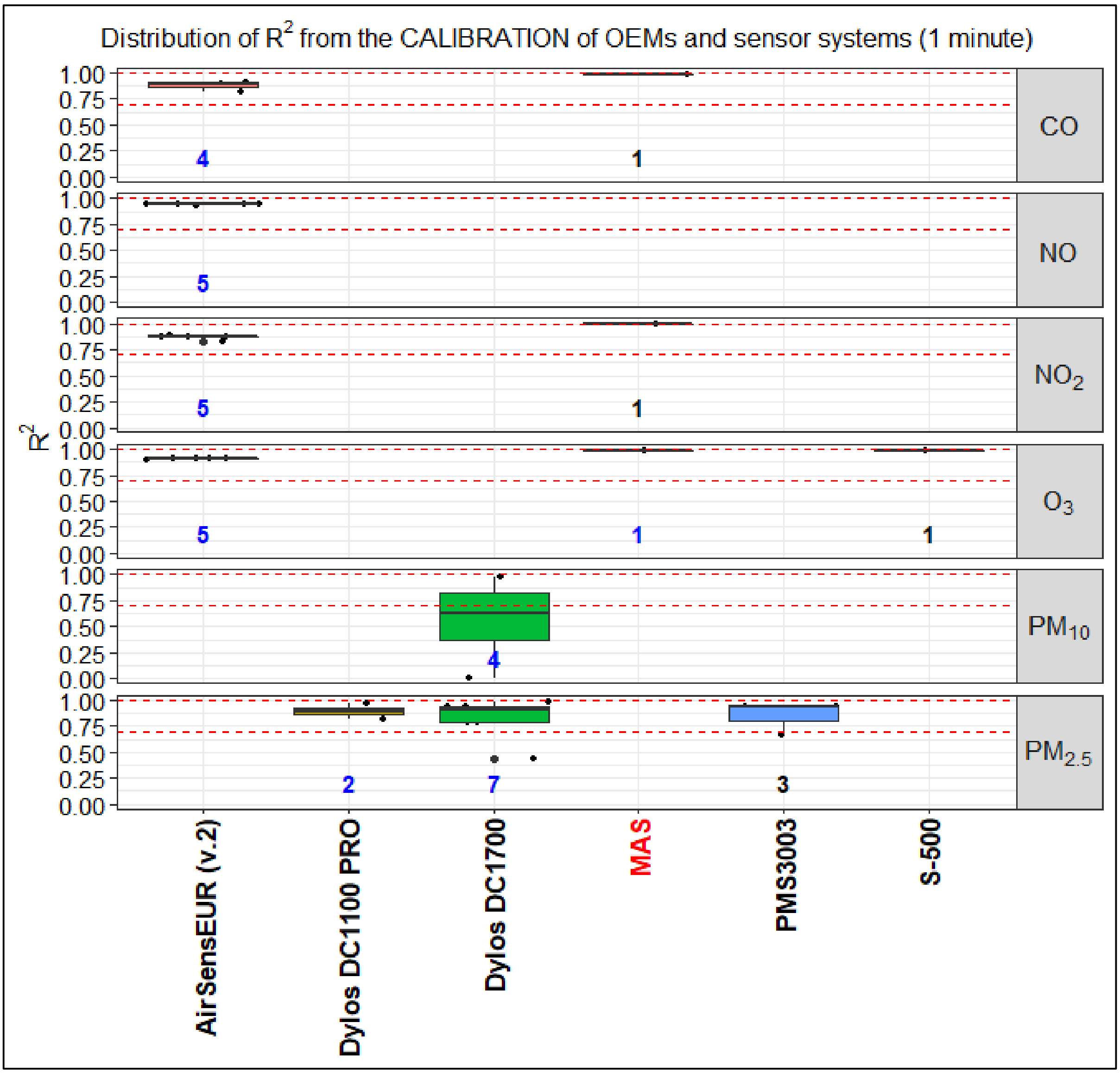


Figure A2. Distribution of R^2 for OEMs and ensor systems obtained from the calibration against the reference. Records were averaged over a time-scale of 1 minute. Dashed lines indicate the value of 0.7 and 1.0. Numbers in bold indicate the number of open source (blue) and black box (black) records. Names of 'living' and 'non-living' sensors are indicated in black and red colour, respectively.

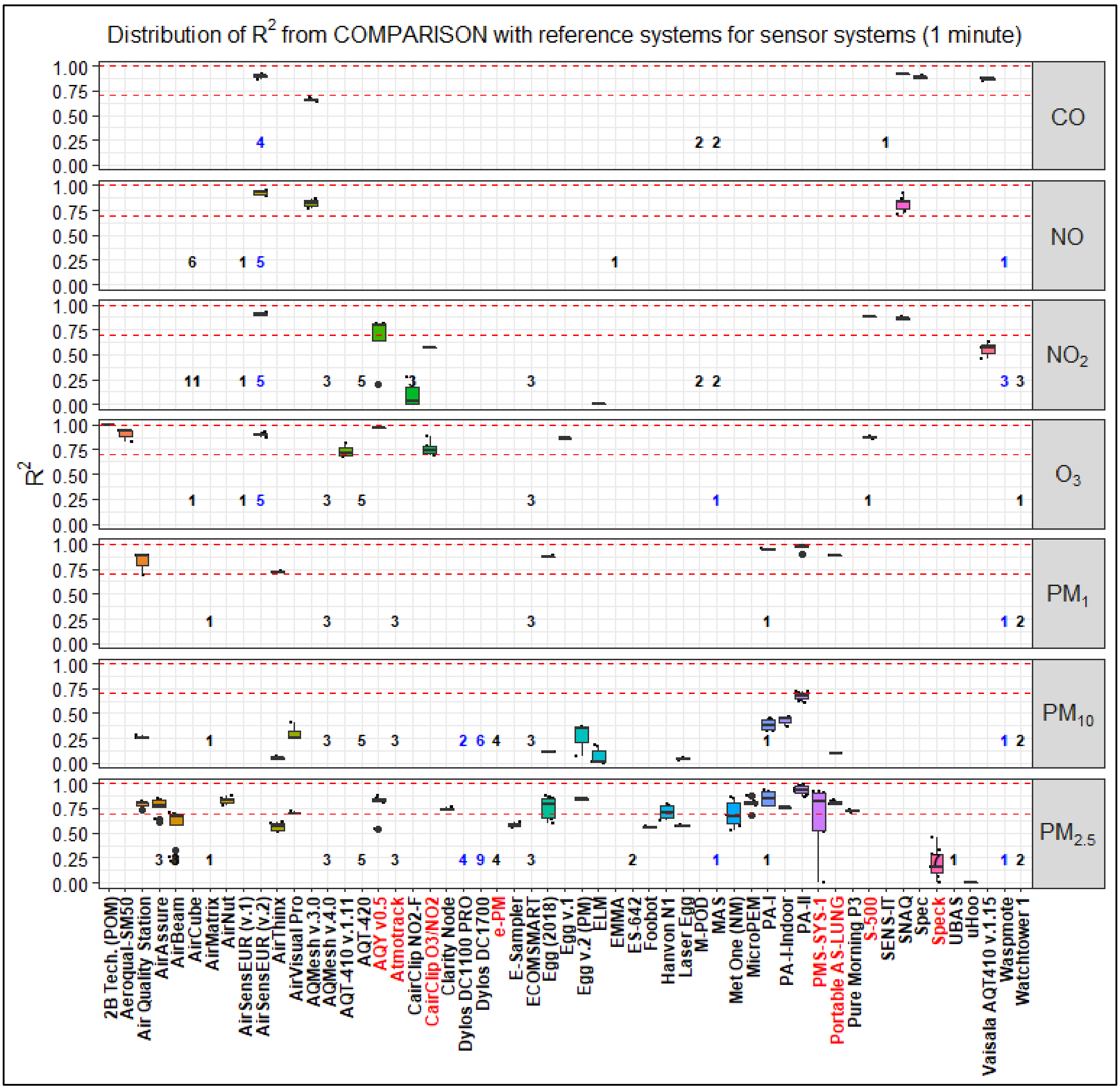


Figure A3. Distribution of R^2 from the comparison of sensor systems against reference systems. Records were averaged over a time-scale of 1 minute. Numbers in bold indicate the number of open source (blue) and black box (black) records. Names of 'living' and 'non-living' sensors are indicated in black and red colour, respectively.

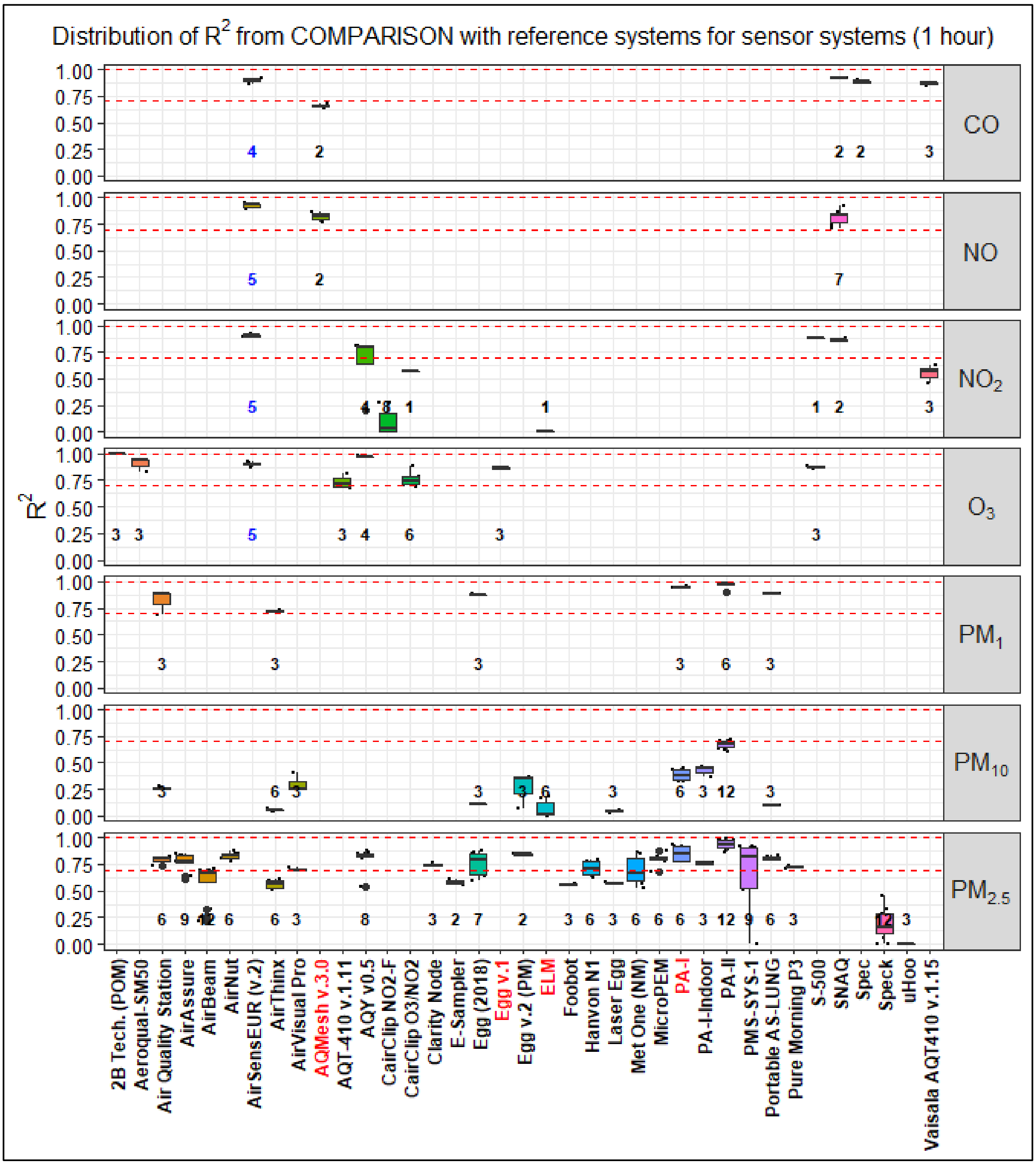


Figure A4. Distribution of R^2 from the comparison of sensor systems against reference systems. Records were averaged over a time-scale of 1 hour. Numbers in bold indicate the number of open source (blue) and black box (black) records. Names of 'living' and 'non-living' sensors are indicated in black and red colour, respectively.

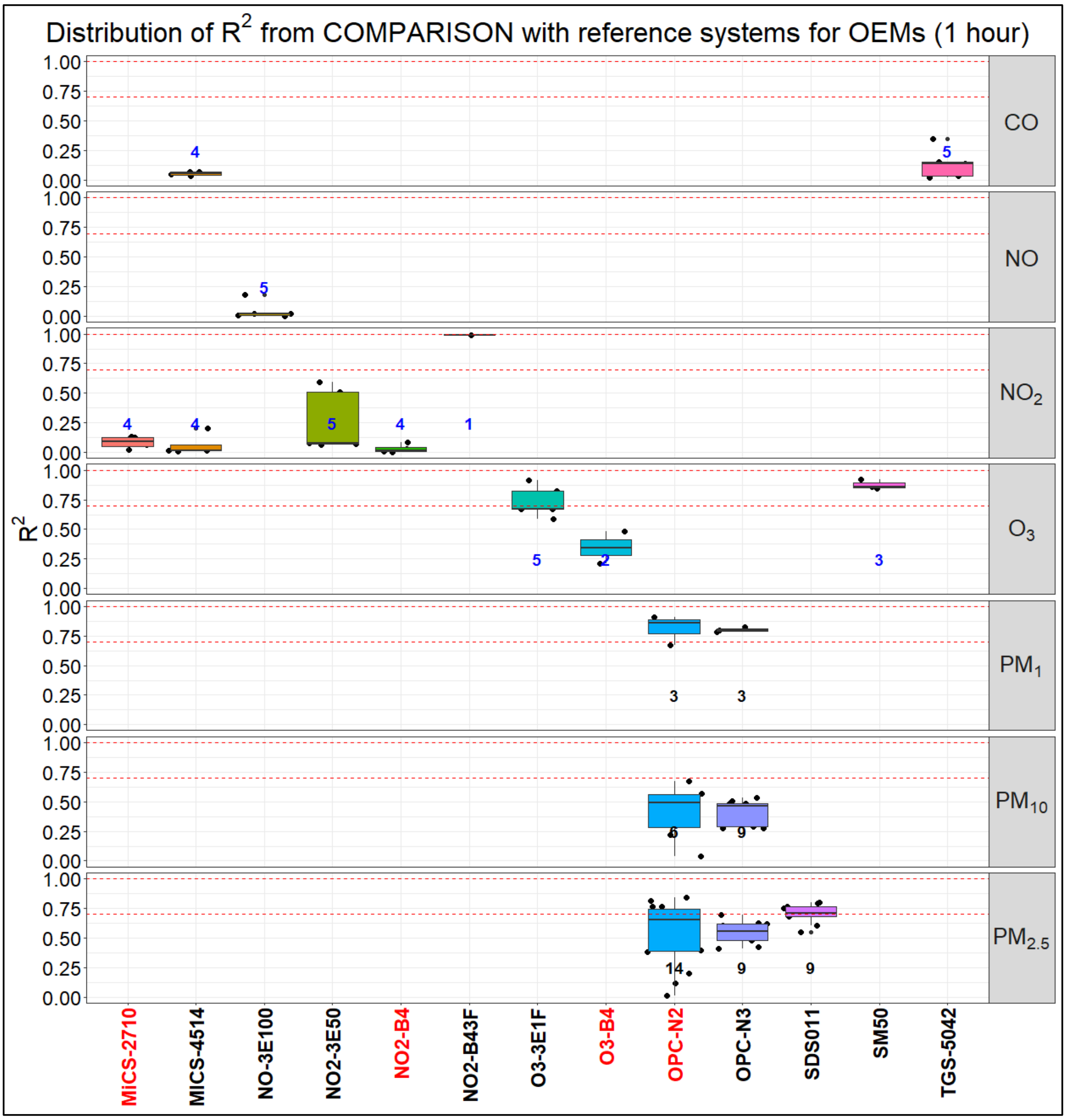


Figure A5. Distribution of R^2 from the comparison of OEMs against reference systems. Records were averaged over a time-scale of 1 hour. Numbers in bold indicate the number of open source (blue) and black box (black) records. Names of 'living' and 'non-living' sensors are indicated in black and red colour,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

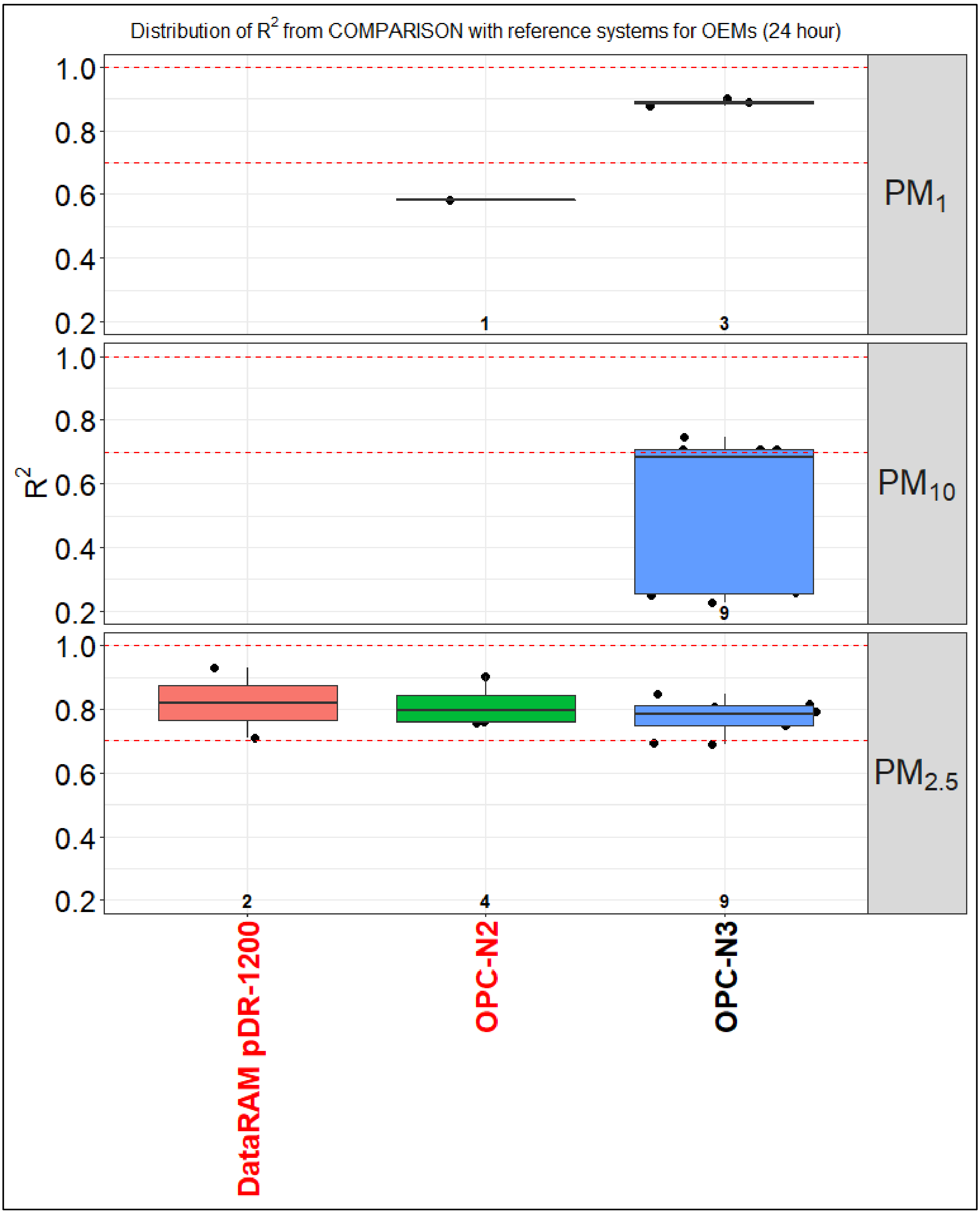


Figure A6. Distribution of R^2 from the comparison of OEMs against reference systems. Records were averaged over a time-scale of 24 hour. Numbers in bold indicate the number of open source (blue) and black box (black) records. Names of 'living' and 'non-living' sensors are indicated in black and red colour, respectively.

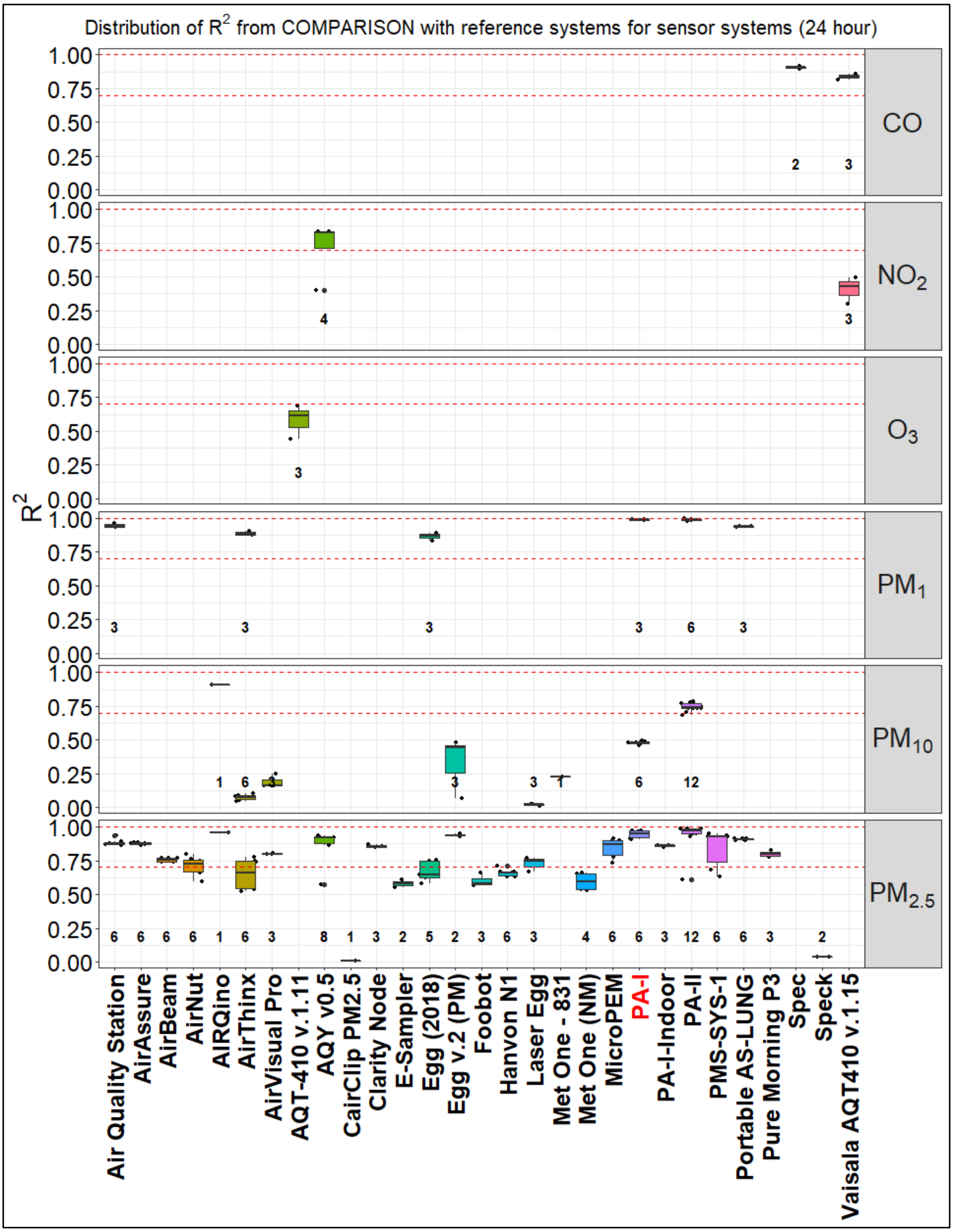


Figure A7. Distribution of R^2 from the comparison of sensor systems against reference systems. Records were averaged over a time-scale of 24 hour. Numbers in bold indicate the number of open source (blue) and black box (black) records. Names of 'living' and 'non-living' sensors are indicated in black and red colour, respectively.

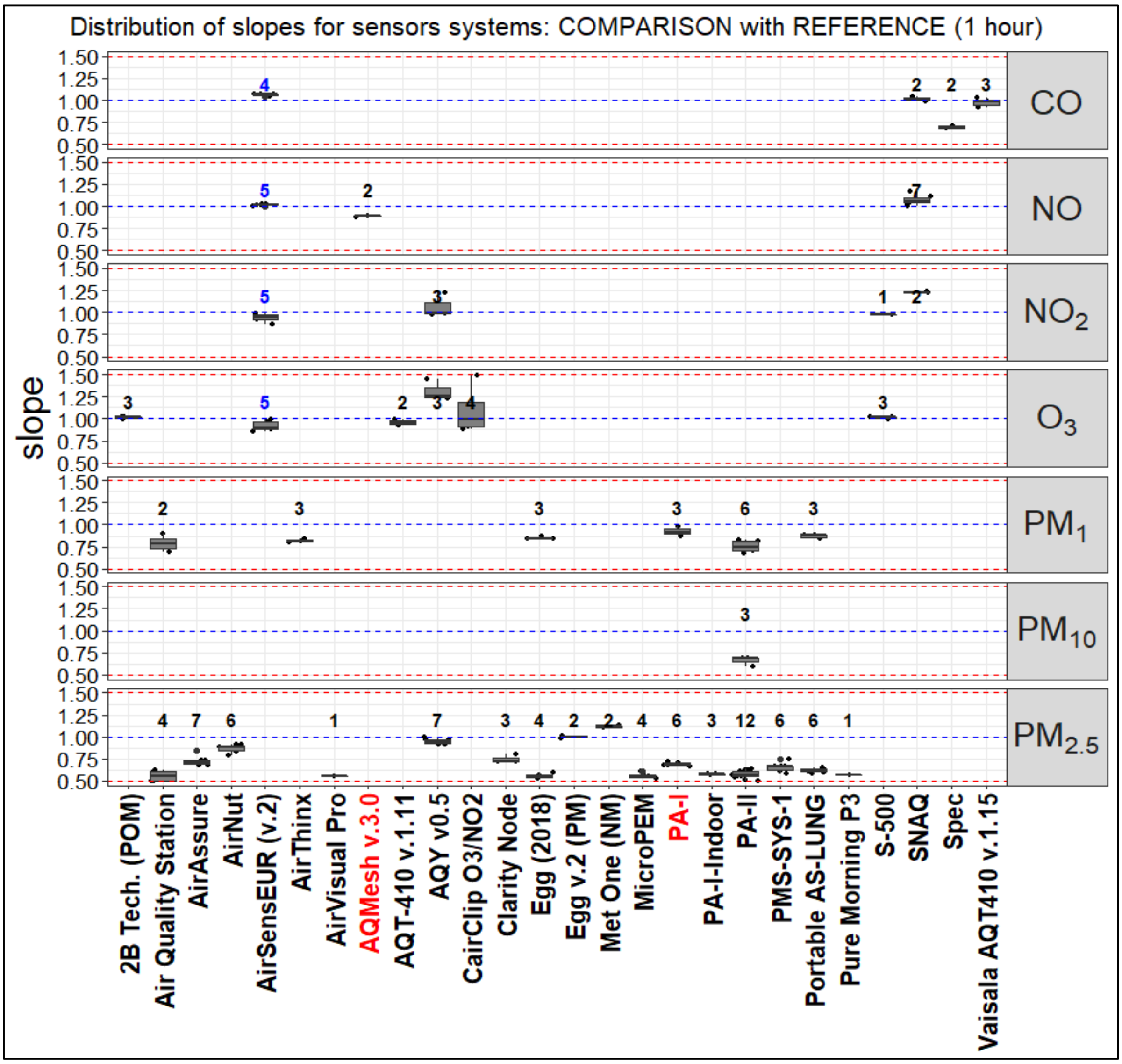


Figure A8. Distribution of slopes from the comparison of sensors systems against the reference. Only records with R^2 > 0.7 and 0.5 < slope < 1.5 are shown. Records were averaged over a time-scale of 1 hour. Numbers in bold indicate the number of open source (blue) and black box (black) records. Names of 'living' and 'non-living' sensors are indicated in black and red colour, respectively.

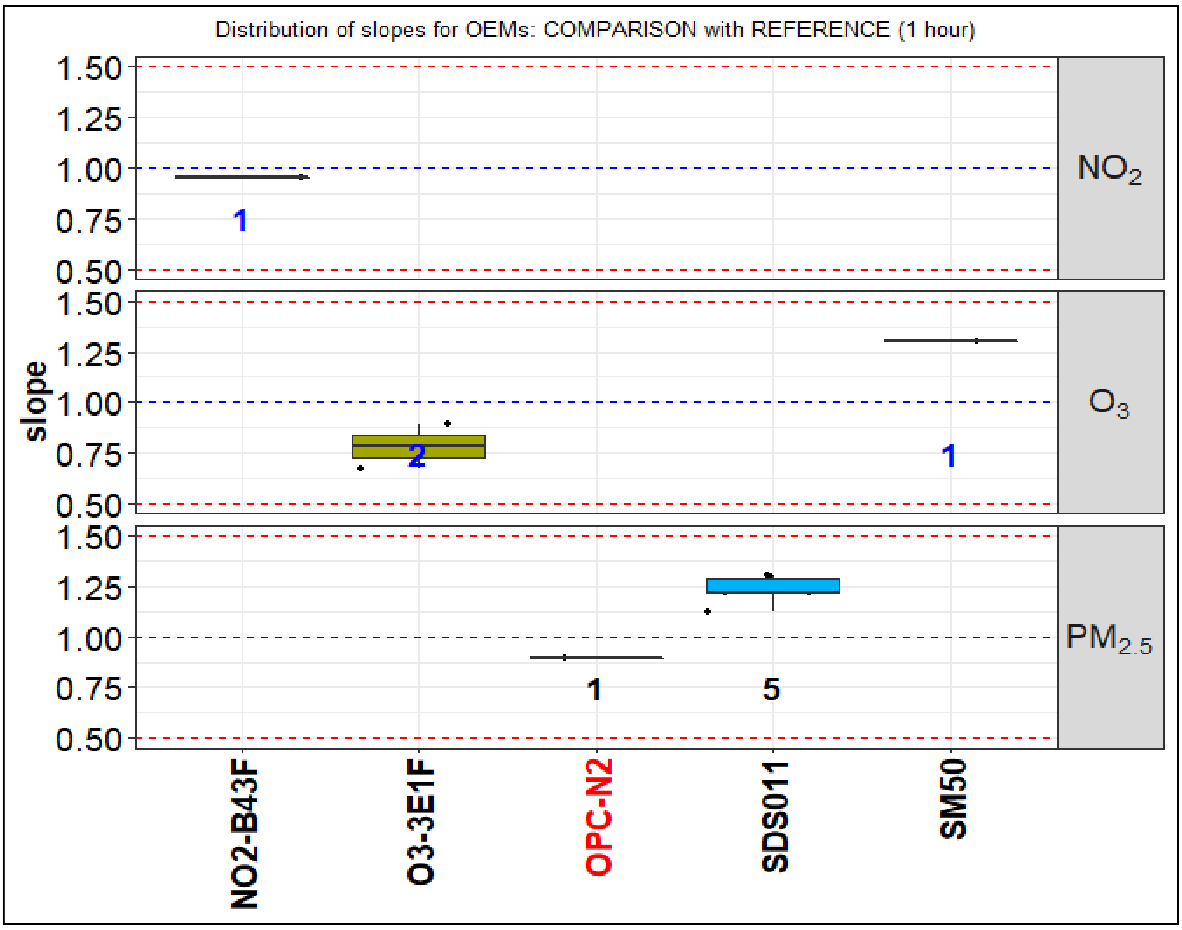


Figure A9. Distribution of slopes from the comparison of OEMs against the reference. Only records with R^2 > 0.7 and 0.5 < slope < 1.5 are shown. Records were averaged over a time-scale of 1 hour. Numbers in bold indicate the number of open source (blue) and black box (black) records. Names of 'living' and 'non-living' sensors are indicated in black and red colour, respectively.

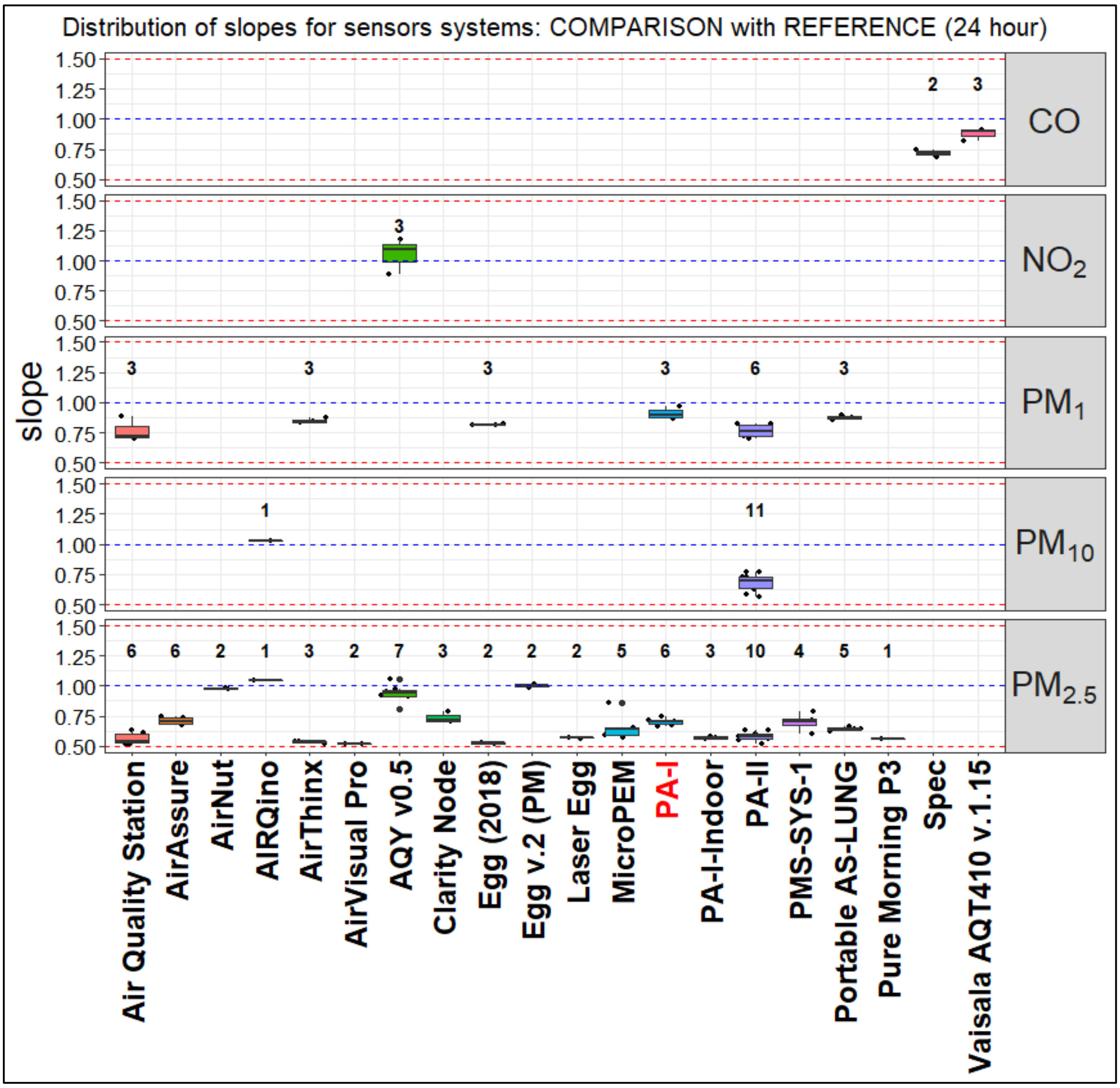


Figure A10. Distribution of slopes from the comparison of sensors systems against the reference. Only records with R^2 > 0.7 and 0.5 < slope < 1.5 are shown. Records were averaged over a time-scale of 24 hour. Numbers in bold indicate the number of open source (blue) and black box (black) records. Names of 'living' and 'non-living' sensors are indicated in black and red colour, respectively.

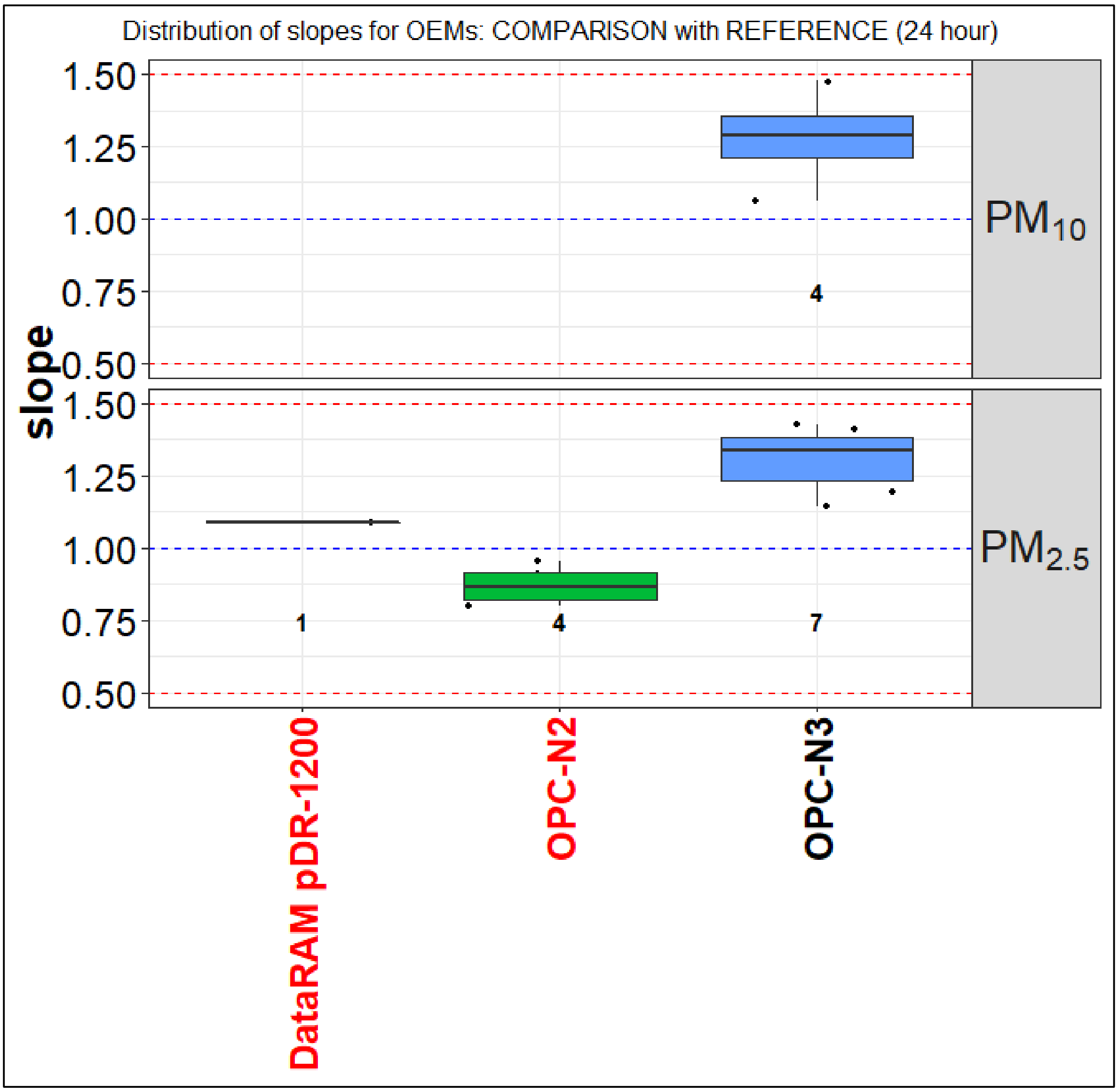


Figure A11. Distribution of slopes from the comparison of OEMs against the reference. Only records with R^2 > 0.7 and 0.5 < slope < 1.5 are shown. Records were averaged over a time-scale of 24 hour. Numbers in bold indicate the number of open source (blue) and black box (black) records. Names of 'living' and 'non-living' sensors are indicated in black and red colour, respectively.

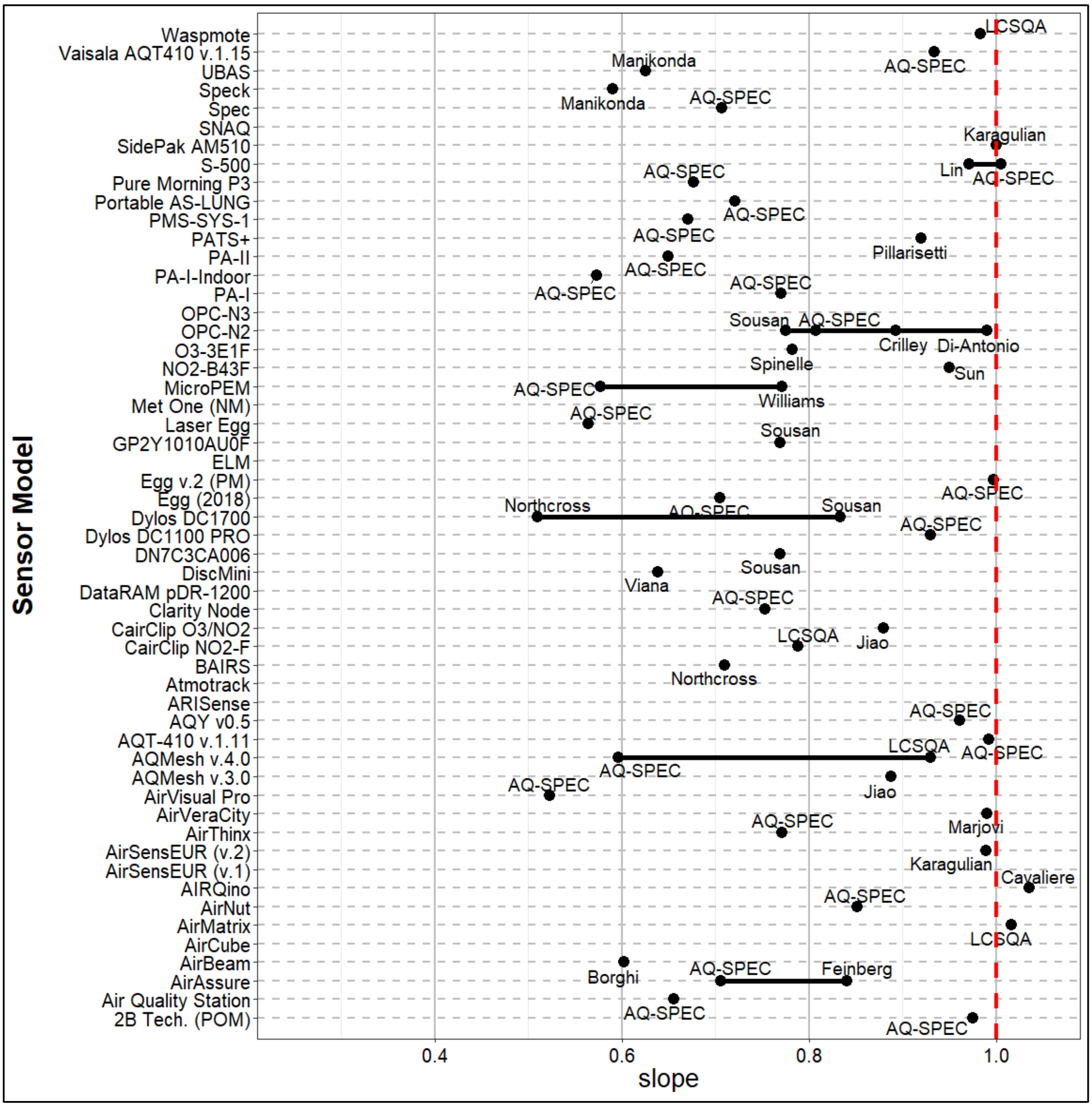


Figure A12. Mean *slope* for obtained from the comparison of OEMs and sensor systems against reference measurements.

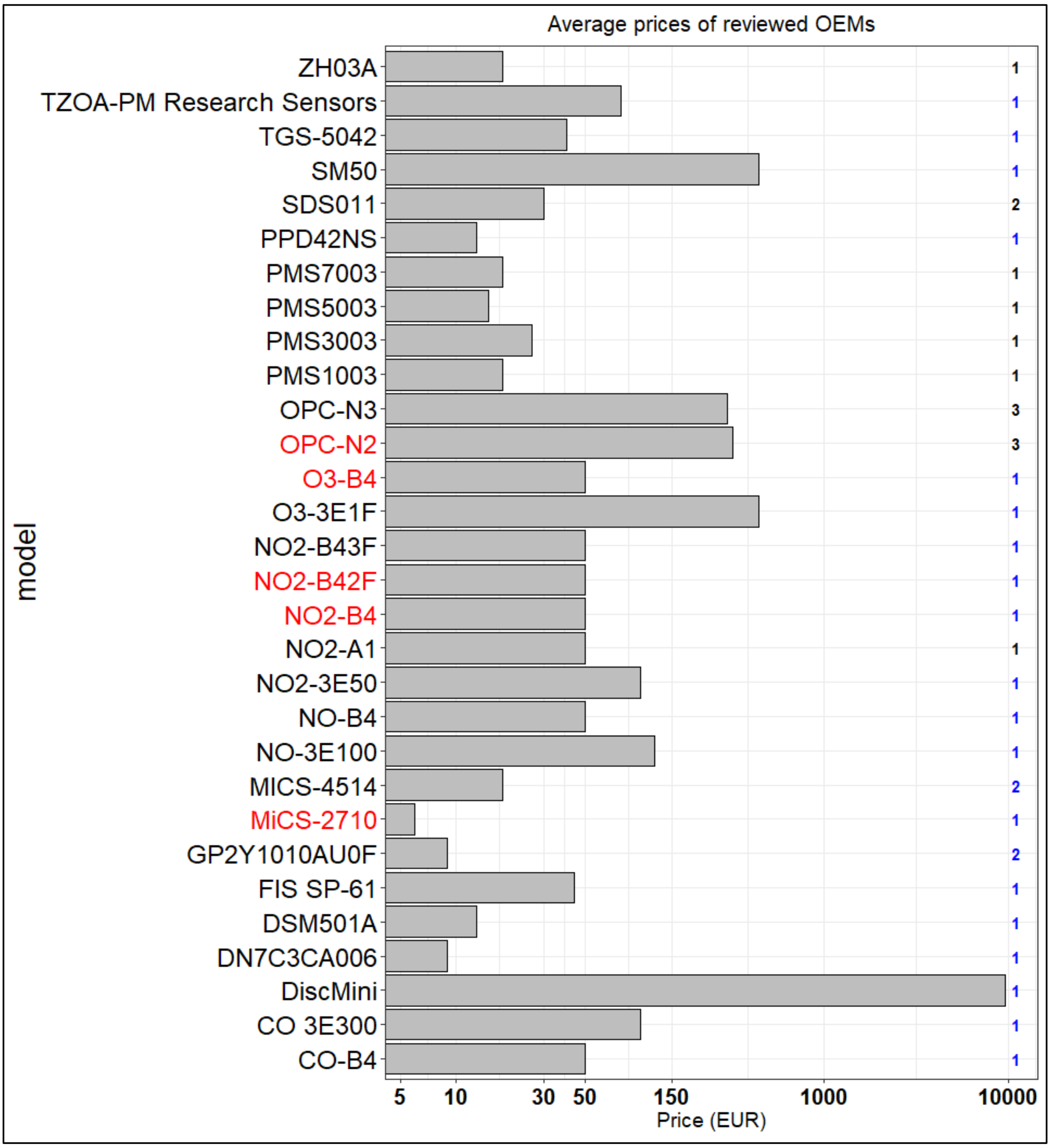


Figure A13. Prices of OEMs available on the market (Numbers in bold indicates the number of pollutants measured by each sensor. x-axis uses logarithmic scale). Numbers in bold indicate the number of open source (blue) and black box (black) records. Names of 'living' and 'non-living' sensors are indicated in black and red colour, respectively.

Bibliography

1. Alexandre, M. & Gerboles, M. Review of small commercial sensors for indicative monitoring of ambient gas. *Chemical Engineering Transactions* **30**, 169–174 (2012).

2. Alvarado, M. *et al.* Towards the Development of a Low Cost Airborne Sensing System to Monitor Dust Particles after Blasting at Open-Pit Mine Sites. *Sensors* **15**, 19667–19687 (2015).

3. AQ-SPEC, District, S. C. A. Q. M. & District, S. C. A. Q. M. *Air Quality Sensor Performance Evaluation Reports.* (2015).

4. Austin, E., Novosselov, I., Seto, E. & Yost, M. G. Laboratory Evaluation of the Shinyei PPD42NS Low-Cost Particulate Matter Sensor. *PLOS ONE* **10**, e0137789 (2015).

5. Badura, M., Batog, P., Drzeniecka-Osiadacz, A. & Modzel, P. Optical particulate matter sensors in $\mathrm{PM}_{2.5}$ measurements in atmospheric air. *E3S Web of Conferences* **44**, 00006 (2018).

6. Barrett, J. P. The Coefficient of Determination—Some Limitations. *The American Statistician* **28**, 19–20 (1974).

7. Bettair. *Bettair - Solving Cities Air Pollution bettair cities*.

8. Bigi, A., Mueller, M., Grange, S. K., Ghermandi, G. & Hueglin, C. Performance of NO, NO_2 low cost sensors and three calibration approaches within a real world application. *Atmospheric Measurement Techniques* **11**, 3717–3735 (2018).

9. Borghi, F. *et al.* Precision and Accuracy of a Direct-Reading Miniaturized Monitor in PM2.5 Exposure Assessment. *Sensors* **18**, 3089 (2018).

10. Borrego, C. *et al.* Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise. *Atmospheric Environment* **147**, 246–263 (2016).

11. Budde, M. *et al.* Suitability of the Low-Cost SDS011 Particle Sensor for Urban PM-Monitoring. (2018).

12. Castell, N. *et al.* Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? *Environment International* **99**, 293–302 (2017).

13. Cavaliere, A. *et al.* Development of Low-Cost Air Quality Stations for Next Generation Monitoring Networks: Calibration and Validation of PM2.5 and PM10 Sensors. *Sensors* **18**, 2843 (2018).

14. 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).

15. 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).

16. CEN. Ambient air - Standard method for the measurement of the concentration of ozone by ultraviolet photometry, (EN 14625:2012). European Committee for Standardization. (2012).

17. 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).

18. 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).

19. 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* **38**, 3329–3340 (2004).

20. 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* **267**, 245–254 (2018).

21. Crilley, L. R. *et al.* Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring. *Atmospheric Measurement Techniques* **11**, 709–720 (2018).

22. Cross, E. S. *et al.* Use of electrochemical sensors for measurement of air pollution: Correcting interference response and validating measurements. *Atmos. Meas. Tech.* 14 (2017).

23. Crunaire, S., Redon, N. & Spinelle, L. *1ER ESSAI NATIONAL D'APTITUDE DES MICRO-CAPTEURS(EAµC) POUR LA SURVEILLANCE DE LA QUALITE DE L'AIR : SYNTHESE DES RESULTATS*. 38 (LCSQA, 2018).

24. Dacunto, P. J. *et al.* Determining $\mathrm{PM}_{2.5}$ calibration curves for a low-cost particle monitor: Common indoor residential aerosols. *Environmental Science: Processes & Impacts* **17**, 1959–1966 (2015).

25. De Vito, S. *et al.* Calibrating chemical multisensory devices for real world applications: An in-depth comparison of quantitative machine learning approaches. *Sensors and Actuators B: Chemical* **255**, 1191–1210 (2018).

26. 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* **18**, 2790 (2018).

27. 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.

28. Duvall, R. M. *et al.* Performance Evaluation and Community Application of Low-Cost Sensors for Ozone and Nitrogen Dioxide. *Sensors* **16**, 1698 (2016).

29. Esposito, E. *et al.* Dynamic neural network architectures for on field stochastic calibration of indicative low cost air quality sensing systems. *Sensors and Actuators B: Chemical* **231**, 701–713 (2016).

30. Commission, E. *Guide to the demonstration of equivalence of ambient air monitoring methods, Report by an EC Working, Group on Guidance*. (European Commission, 2010).

31. Feinberg, S. *et al.* Long-term evaluation of air sensor technology under ambient conditions in Denver, Colorado. *Atmospheric Measurement Techniques* **11**, 4605–4615 (2018).

32. Fishbain, B. *et al.* An evaluation tool kit of air quality micro-sensing units. *Science of The Total Environment* **575**, 639–648 (2017).

33. 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* **199**, 56–65 (2015).

34. 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). doi:10.2788/30927

35. Gillooly, S. E. *et al.* Development of an in-home, real-time air pollutant sensor platform and implications for community use. *Environmental Pollution* **244**, 440–450 (2019).

36. 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* **67**, 330–340 (2017).

37. 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* **7**, 1121–1131 (2014).

38. iScape. *Summary of Air Quality sensors and recommendations for application. iScape project D1.5, February 2017*. (2017).

39. Jiao, W. *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* **9**, 5281–5292 (2016).

40. Jovašević-Stojanović, M. *et al.* On the use of small and cheaper sensors and devices for indicative citizen-based monitoring of respirable particulate matter. *Environmental Pollution* **206**, 696–704 (2015).

41. Karagulian, F. *et al. Calibration of AirSensEUR units during a field study in the Netherlands*. (European Commission - Joint Research Centre, 2019).

42. 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. *Journal of Environmental Monitoring* **14**, 2145–2153 (2012).

43. Kelly, K. E. *et al.* Ambient and laboratory evaluation of a low-cost particulate matter sensor. *Environmental Pollution* **221**, 491–500 (2017).

44. KEY-VOCs. *Metrology for VOC indicators in air pollution and climate change. Revision of the testing protocol and procedure for calibration based on laboratory validation, final report. ENV56 KEY-VOCs.* (2017).

45. Kumar, P. *et al.* The rise of low-cost sensing for managing air pollution in cities. *Environment International* **75**, 199–205 (2015).

46. Kunak. *Wireless Particle counter for Outdoors Kunak AIR P10*. *Kunak*

47. Laquai, B. Particle Distribution Dependent Inaccuracy of the Plantower PMS5003 low-cost PM-sensor. (2017).

48. Lewis, A. C., Schneidemesser, E. von & Peltier, R. *Low-cost sensors for the measurement of atmospheric composition: Overview of topic and future applications (World Meteorological Organization) WMO-No. 1215.* (2018).

49. Lewis, A. & Edwards, P. Validate personal air-pollution sensors. *Nature News* **535**, 29 (2016).

50. Lin, C. *et al.* Evaluation and calibration of Aeroqual series 500 portable gas sensors for accurate measurement of ambient ozone and nitrogen dioxide. *Atmospheric Environment* **100**, 111–116 (2015).

51. 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* **10**, 41 (2019).

52. Manikonda, A., Zíková, N., Hopke, P. K. & Ferro, A. R. Laboratory assessment of low-cost PM monitors. *Journal of Aerosol Science* **102**, 29–40 (2016).

53. 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 2017 International Conference on Embedded Wireless Systems and Networks* 12–23 (Junction Publishing, 2017).

54. Mead, M. I. *et al.* The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmospheric Environment* **70**, 186–203 (2013).

55. Mijling, B., Jiang, Q., Jonge, D. de & Bocconi, S. Practical field calibration of electrochemical NO2 sensors for urban air quality applications. *Atmos. Meas. Tech. Discuss.* **2017**, 1–25 (2017).

56. 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.* 17 (2017).

57. 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* **17**, 1805 (2017).

58. Northcross, A. L. *et al.* A low-cost particle counter as a realtime fine-particle mass monitor. *Environmental Science: Processes & Impacts* **15**, 433–439 (2013).

59. 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* **8**, 7511–7533 (2015).

60. Piedrahita, R. *et al.* The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. *Atmospheric Measurement Techniques* **7**, 3325–3336 (2014).

61. Pillarisetti, A. *et al.* Small, Smart, Fast, and Cheap: Microchip-Based Sensors to Estimate Air Pollution Exposures in Rural Households. *Sensors (Basel, Switzerland)* **17**, (2017).

62. 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* **147**, 330–343 (2016).

63. Snyder, E. *et al.* The changing paradigm of air pollution monitoring. *Environmental Science and Technology* **47**, 11369–11377 (2013).

64. 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* **50**, 1352–1365 (2016).

65. Sousan, S. *et al.* Inter-comparison of low-cost sensors for measuring the mass concentration of occupational aerosols. *Aerosol Science and Technology* **50**, 462–473 (2016).

66. 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* 319–324 (2016). doi:10.3303/CET1654054

67. 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). doi:10.2760/548327

68. 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* **215**, 249–257 (2015).

69. Spinelle, L., Gerboles, M., Villani, M. G., Aleixandre, M. & Bonavitacola. Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part. B: NO, CO and CO2. *Sensors and Actuators B: Chemical* **in preparation**,

70. 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* **120**, 480–483 (2015).

71. Spinelle, L., Gerboles, M. & Aleixandre, M. *Report of laboratory and in-situ validation of micro-sensor for monitoring ambient air pollution - O12: CairClipO3/NO2 of CAIRPOL (F)*. (Publications Office of the European Union, 2013).

72. Spinelle, L., Gerboles, M. & Aleixandre, M. *Report of laboratory and in-situ validation of micro-sensor for monitoring ambient air pollution - NO9: CairClipNO2 of CAIRPOL (F)*. (Publications Office of the European Union, 2013).

73. Steinle, S. *et al.* Personal exposure monitoring of PM 2.5 in indoor and outdoor microenvironments. *Science of The Total Environment* **508**, 383–394 (2015).

74. Sun, L. *et al.* Development and Application of a Next Generation Air Sensor Network for the Hong Kong Marathon 2015 Air Quality Monitoring. *Sensors (Basel, Switzerland)* **16**, (2016).

75. Sun, L., Westerdahl, D. & Ning, Z. Development and Evaluation of A Novel and Cost-Effective Approach for Low-Cost NO2 Sensor Drift Correction. *Sensors* **17**, 1916 (2017).

76. Index, T. W. A. Q. *Sensing the Air Quality: Research on Air Quality Sensors. http://aqicn.org/sensor/.* (2019).

77. Thunis, P., Pederzoli, A. & Pernigotti, D. Performance criteria to evaluate air quality modeling applications. *Atmospheric Environment* **59**, 476–482 (2012).

78. Agency, U. S. E. P. Evaluation of Elm and Speck Sensors. **EPA/600/R-15/314 November 2015 www.epa.gov/ord**, (2015).

79. Vaughn, D. L., Dye, T. S., Roberts, L., Ray, A. E. & DeWinter, J. L. *Characterization of Low-Cost NO2 Sensors*. 10 (U.S. Environmental Protection Agency Research Triangle Park, North Carolina, 2010).

80. Viana, M. *et al.* Field comparison of portable and stationary instruments for outdoor urban air exposure assessments. *Atmospheric Environment* **123, Part A**, 220–228 (2015).

81. Wang, Y. *et al.* Laboratory Evaluation and Calibration of Three Low-Cost Particle Sensors for Particulate Matter Measurement. *Aerosol Science and Technology* **49**, 1063–1077 (2015).

82. Wastine, B. *AirSensEur: Point sur les expérimentations menées depuis 2018*. (Atmo-Normadie, 2019).

83. 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*. 28 (Atmo Normandie, 2019).

84. Wei, P. *et al.* Impact Analysis of Temperature and Humidity Conditions on Electrochemical Sensor Response in Ambient Air Quality Monitoring. *Sensors* **18**, 59 (2018).

85. White, R. *et al.* Sensors and 'apps' for community-based: Atmospheric monitoring. *EM: Air and Waste Management Association's Magazine for Environmental Managers* 36–40 (2012).

86. Williams, R. *et al. Air Sensor Guidebook, United States Environmental Protection Agency (US-EPA)*. (2014).

87. Williams, R. *et al.* Deliberating performance targets workshop: Potential paths for emerging PM2.5 and O3 air sensor progress. *Atmospheric Environment: X* **2**, 100031 (2019).

88. Williams, R., Kaufman, A., Hanley, T., Rice, J. & Garvey, S. *Evaluation of Field-deployed Low Cost PM Sensors*. (2014).

89. Williams, R. *et al. Sensor Evaluation Report*. (2014).

90. Zheng, T. *et al.* Field evaluation of low-cost particulate matter sensors in high and low concentration environments. *Atmospheric Measurement Techniques Discussions* 1–40 (2018). doi:10.5194/amt-2018-111

91. Zhou, X., Lee, S., Xu, Z. & Yoon, J. Recent Progress on the Development of Chemosensors for Gases. *Chemical Reviews* **115**, 7944–8000 (2015).

92. Zikova, N. *et al.* Estimating Hourly Concentrations of PM2.5 across a Metropolitan Area Using Low-Cost Particle Monitors. *Sensors (Basel, Switzerland)* **17**, (2017).

93. Zimmerman, N. *et al.* A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring. *Atmos. Meas. Tech.* **11**, 291–313 (2018).

KJ-NA-29826-EN-N

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

EU Science, Research and Innovation

EU Science Hub

Publications Office
of the European Union

doi:10.2760/568261

ISBN 978-92-76-09255-1