



Supplement of

Long-term evaluation of commercial air quality sensors: an overview from the QUANT (Quantification of Utility of Atmospheric Network Technologies) study

Sebastian Diez et al.

Correspondence to: Sebastian Diez (sebastian.diez@udd.cl) and Pete M. Edwards (pete.edwards@york.ac.uk)

The Supplement has been replaced due to an error in Eq. (S6). The third term has been corrected from $(y_i - b_0 - b_1 x_i)^2$ to $(x_i - b_0 - b_1 x_i)^2$. This mistake could otherwise cause confusion/misuse by readers, since "x" (reference monitors) represents a different dataset than "y" (sensors).

The copyright of individual parts of the supplement might differ from the article licence.

1

2

3 Sect. S1. Co-location sites

For the main QUANT deployment, 3 field sites were chosen: Manchester, London, and York, all providing
extensive reference measurements across a range of chemical environments representative of UK urban
atmospheres. On the other hand, only the Manchester site was used for the WPS colocation.

.

7 The Manchester Air Quality Supersite (MAOS, 53° 26' 39.2"N, 2° 12' 51.9"W) stands as one of the largest air 8 quality research facilities in the UK. Situated in an urban background setting approximately four kilometres south 9 of Manchester city center — the UK's second-largest metropolitan area with around 3.3 million residents — 10 MAQS benefits from a strategic location on the University of Manchester's Fallowfield Campus. This location is 11 notably distanced from direct traffic emissions, surrounded by student accommodations, university administrative 12 buildings, and sports facilities. The campus's vicinity to shops, bars, and restaurants introduces a range of human 13 activities, including varying levels of foot traffic and associated vehicular movement. Additionally, the presence 14 of these commercial and recreational spaces, alongside residential buildings, contributes to the area's ambient air 15 quality through emissions from heating and cooking, among other sources. For a visual representation of MAOS's 16 surroundings, please refer to Figure S1 (panel a). The site experiences an average winter temperature of 17 approximately 4-5°C with relative humidity around 87%, and an average summer temperature of about 16-17°C 18 with relative humidity near 88%. Detailed information on MAQS's reference instrumentation and the 19 methodologies employed for air quality measurements can be found in section S2. Data from MAQS are provided 20 with a 1-minute time resolution, facilitating a granular temporal analysis of air quality metrics.

21 The London Air Quality Supersite (LAQS, 51° 26' 58.9"N 0° 02' 14.6"W) serves as an urban background 22 monitoring site, nestled within Honor Oak Park in Greater London. Situated 9 km southeast of the city center of 23 the third-largest European urban conglomeration, LAQS offers a unique window into the air quality challenges of 24 an area inhabited by approximately 14.8 million people. Nestled within the serene King's College sports grounds, 25 is surrounded by middle-class neighbourhoods, abundant parks, and green spaces. This tranquil setting, is 26 distanced from major roads and pollution sources, provides a representative snapshot of the ambient air quality 27 typical of residential London. LAQS's surroundings are marked by a low level of commercial activity, with local 28 shops and restaurants contributing minimally to the area's overall noise and bustle. Figure S1 (panel b) offers an 29 aerial view of LAQS, illustrating the overall urban layout. The area is characterised by a temperate climate, 30 experiencing average winter temperatures of around 5°C with RH of approx. 84%, and milder summers with 31 temperatures averaging 17°C and RH of around 72%. Gas measurements at LAQS are conducted with a 1-minute

32 time resolution, while PM data are collected at a 15-minute resolution (see section S2 for more details).

The <u>York</u> Fishergate roadside site (YoFi, 53° 57' 06.9"N, 1° 04' 33.1"W), in the historic city of York, which is
home to approximately 210,000 inhabitants (avg. temp. in winter of ~4°C and RH ~87 %, avg. temp. in summer

35 around 15 °C and RH ~80 %). Situated just about 1 km from the city center on a traffic island, YoFi stands amidst

36 a predominantly residential area that also encompasses commercial and light industrial elements. Unique to its

37 location, the site is sandwiched between two lanes of Fishergate Road, a major avenue that bifurcates to facilitate

- 38 traffic flow into and out of the city's southern part. Directly across from YoFi, a primary school adds to the daily
- 39 human activity around the site, while the nearby River Ouse, located merely 300 metres to the west, contributes
- 40 to the area's environmental characteristics. A vibrant commercial zone, featuring pubs and restaurants, is found
- 41 just 100 metres to the north. Moreover, the site is flanked by Walmgate Stray, an expanse of recreational fields,
- 42 located about 300 metres to the southeast, offering a green respite amidst the urban setting. Additional details can
- 43 be visualised in Figure S1 (panel c), providing an aerial perspective of the site's key features and its urban context.
- 44 This self-contained air quality monitoring station was specifically selected for the QUANT study to assess sensors'
- 45 responses to the greater pollutant variability typical of traffic-related sites, contrasting with the urban background
- 46 settings of MAQS and LAQS. YoFi provides data on PM and NOx with a 1-hour time resolution. Additionally,
- 47 in a targeted effort to enhance our understanding of air quality dynamics, O₃ measurements (deployed on the 15th
- 48 of May 2020, specifically as part of the QUANT study), utilising a 1-minute time resolution to offer detailed
- 49 insights into temporal variations (refer to section S2 for more details).



- 51 Figure S1: Aerial views of the air quality monitoring sites: a) MAQS, b) LAQS, and c) YoFi, captured from Google
- 52 Earth. These images illustrate the diverse urban settings of each site, emphasising aspects such as their proximity to
- 53 traffic sources, presence of green spaces, and the general urban layout. Image credits: © Google Earth.

54 Sect. S2. Reference instrumentation, QA/QC, and data-sharing periods

- 55 Table S1 summarises the reference instrumentation at each site, Table S2 describes some of the QA/QC processes
- at the supersites, and Table S3 shows the data periods shared with the suppliers.
- 57 Table S1. Research grade instrumentation used for the QUANT study.

Analyte	Manchester	London	York
NO	Thermo 42i-y (Chem)	Teledyne T200U (Chem)	Teledyne T200UP
NO ₂	*Teledyne T500U (CAPS)	*Teledyne T500U (CAPS)	(Chem)
O 3	*Thermo 49i (UV)	*Teledyne 400E (UV)	*2B 205 (UV)
PM	*Palas FIDAS200 (OAS)	*Palas FIDAS200 (OAS)	*Met One BAM 1020 (BA)

*Equivalent to reference (as defined in the European Air Quality Directive 2008/50/EC)

Acronyms: Chem: Chemiluminescence; CAPS: Cavity Attenuated Phase Shift Spectroscopy; UV: Ultraviolet; OAS:
 Optical aerosol spectrometer; BA: Beta attenuation.

61 Table S2. Summary of Quality Assurance processes in MAQS and LAQS

Instrument	Frequency	*Process	
NOy	At least monthly	Zero and span checks using standard cylinder and scrubber. Corrections to zero and span values.	
NO_2	Daily	Automatic zero and span checks using internal NO ₂ diffusion tube and scrubber. Zero corrections, span monitored.	
O ₃	Daily	Automatic zero and span checks using internal O_3 lamp and scrubber. Corrections to zero, span monitored.	
СО	Every three hours & monthly	Zero checks every three hours and span checks monthly using onsite cylinder. Adjustments to zero and span values.	
CO ₂ and CH ₄	Regular	Stability checks using onsite cylinder, no corrections made.	
*PM	Semiannual	Sizing response verified with Mono dust, flow rate checked with Gilibrator.	

62 *Checked with external standards by NPL every 6 months. These external standards are also used to provide a certification of the on-site

standard cylinders. Final corrections to the data are provided by using the audit data to define the concentration of the on-site standards, with
 zero and span values interpolated between the calibration points.

65 **Sizing and flow checked every 6-month NPL audit process.1

66 Table S3. Reference data is shared with the sensor manufacturers.

	QUANT main study		Wider Participation Study		
Reference dataset	Period	Released	Reference dataset	Period	Released
1	10-12-2019 - 17-02-2020	15-04-2020	1	17-06-2021 - 16-07-2021	23-07-2021
2	18-02-2020 - 17-08-2020	27-10-2020	2	01-12-2021 - 31-12-2021	26-01-2022
3	18-08-2020 - 17-02-2021	15-04-2021	3	01-05-2022 - 31-05-2022	15-06-2022

67 Sect. S3. QUANT main study devices

68 In this section, a brief description of the QUANT main study systems' components is offered.

69 PurpleAir (PA) (https://www2.purpleair.com) devices (PA-II-SD model, firmware v4.11) reports particulate 70 matter (PM1, PM2.5, and PM10), and it was chosen for its penetration around the world. Two identical Plantower 71 PMS5003 (Plantower) sensors (channels A and B) are found in each PA. It offers two data products (2-min avg. 72 time): the "cf atm" (for outdoor applications) and the "cf 1" (for indoor or controlled environment applications). 73 The PMS behaves like a nephelometer rather than an optical particle counter to measure the light scattered by the 74 PM (Ouimette et al., 2022) and is composed of a laser, a photodiode, a fan, and a microprocessor control unit. 75 They also measure temperature (Temp), relative humidity (RH), and atmospheric pressure (Pres) (Bosch). The 76 data can be communicated via Wi-Fi or stored locally (microSD card), which was the preferred way during the 77 colocation. No calibrated products are offered by the company.

78 *Note: For this study, only Channel A and the data product "cf atm" were included in the analysis and shown in 79 the plots.

80 AQMesh (https://www.aqmesh.com) reports NO₂, NO, O₃ using electrochemical (EC) sensors (Alphasense), CO₂ 81 with a non-dispersive infrared sensor (NDIR, Alphasense), PM₁, PM_{2.5}, and PM₁₀ through a light-scattering sensor 82 (Nephelometer, Environmental Instr.) with 1-minute time resolution (algorithm v5.1 for gases and v3.0 for PM). 83 This instrument also registers Temp, RH, and Pres (Solid-State sensors) (Zauli-Sajani et al., 2022) and the 84 sampling mechanism employs a pump. The collected data is sent to the company server via a cellular network and 85 post-processed (Temp, RH, and cross-interference correction) in the cloud by a proprietary algorithm. Finally, the 86 data is released to the final user via secure web login or through its Application Programming Interface (API).

- 87 Although the first 4 months of the deployment the data had a 15-min resolution, since then the provided resolution 88 is 1-min average.
- 89 AQY (v.1.0) is also a multi-species device (https://www.aeroqual.com) and measures O₃, NO₂, PM_{2.5}, PM₁₀,
- 90 Temp, and RH. This is the only device system that does not use Alphasense sensors for gases. While O_3 is
- 91 quantified using a metal oxide sensor (WO₃-based, Aeroqual Ltd), the NO₂ is measured by an EC sensor
- 92 (Membrapore type O₃/M5, Aeroqual Ltd) (Weissert et al., 2019). For PM it uses a light scattering method (Nova)
- 93 to convert size and particle count to a mass fraction and behaves like a nephelometer (Myklebust et al., 2022).

- 94 These LCS devices send their data (1-min time resolution) to the Aeroqual server via cellular (WiFi could also be
- 95 used for this purpose) or stored locally (microSD card). The non-local data access is through a web portal or via 96 API.
- 97 Zephyr units (https://www.earthsense.co.uk) measure PM (Nephelometer, Plantower), Temp & RH (Sensirion),
- 98 and Press (Bosch) (the sample uptake uses a fan). As most of the commercial units tested here, it used Alphasense
- 99 EC sensors (the "A series", a smaller version than the B series) for gases (NO, NO₂, and O₃). These devices send
- 100 their raw data to the server via a cellular network, where they pre-process the raw signals. We have secure access
- 101 to the measurements with a time resolution of 1-min per species through the website or via its API.
- 102 ARIsense v200 devices (https://quant-aq.com) measure NO, NO₂, O₃, CO (EC, Alphasense), CO₂ (NDIR,
- 103 Alphasense), Temp & RH (Sensirion), and Press (Bosch) (Cross et al., 2017). Of all the devices tested, this is the
- 104 only one that uses an Optical Particle Counter (OPC) for PM (Particles Plus). Communication is carried out
- 105 through a cellular network and the data products are accessed through a web portal or API (1-minute time
- 106 resolution). According to the company policy, only the gas data products are subjected to calibrations (if
- 107 colocation data is available).

108 109 Table S4. Summary of sensor measurements and the time resolution data provided by participating companies in the Main OUANT study.

System Measurands		Time Resol.
PA	PM1, PM2.5, PM10	2min
AQM	PM1, PM2.5, PM10, NO, NO2, O3, CO2	1min/15min
AQY	PM _{2.5} , PM ₁₀ , NO ₂ , O ₃	1 min
Zep	PM1, PM2.5, PM10, NO, NO2, O3	1min
Ari	PM ₁ , PM _{2.5} , PM ₁₀ , NO, NO ₂ , O ₃ , CO; CO ₂	1min



111 Figure S2. Data product for each of the participating companies during Main QUANT. The top panels are for NO₂,

112 the middle panels for O₃ and the bottom panels for PM_{2.5}. The y-axis represents the different products: "out-of-

box", call and cal2. The x-axis shows the dates for which each company provided the mentioned products.

114 Sect. S4. WPS devices

115 A short description of the WPS devices' components is shown in this section

116 <u>Modulair-PM</u> instruments (<u>https://quant-aq.com</u>) employ two different techniques to obtain PM mass 117 concentration (it samples the air using a fan), an OPC (Alphasense, OPC-N3) and a nephelometer (Plantower, 118 PMS5003). This system provides 1-min time resolution data for PM₁, PM_{2.5}, and PM₁₀, plus size-resolved particle 119 number concentration (range 350 nm to 40 μ m) (Meyer et al., 2022; Westgate and Ng, 2022). Temp, RH, and 120 Press are also measured, but no data was found about the sensing elements it uses. The post-processed data can 121 be accessed locally (microSD card) or through its server (cellular network comm) via its web portal or API.

122 <u>AQMesh</u> (see earlier description).

The <u>Atmos</u> device (<u>http://urbansciences.in/</u>) reports PM₁, PM_{2.5}, PM₁₀ (Plantower, PMS7003) plus Temp and RH (Adafruit), employing a fan as a means to sample the air. The system transmits the data (1-min time resolution) to a cloud server (only via Wi-Fi) and also stores it locally (Puttaswamy et al., 2022). The data can be accessed via a web dashboard or API. Unfortunately, and due to the meteorological conditions at the Manchester supersite these co-located devices only survived for about 2 months.

128 The IMB instrument (https://www.bosch-mobility-solutions.com) measures NO2, O3 PM2.5 and PM10,

129 (Alphasense sensors), plus Press, RH an Temp (no details were found about the brand and model). The raw data

- 130 is transmitted to their cloud using cellular connectivity (3G or LTE). The final data is 1-min resolution (accessed
- 131 only via API).

- 132 Polludrone (https://oizom.com) uses Alphasense sensors for gas measurements (B4 series for NO, NO₂, O₃. No
- 133 data available about CO, CO_2 and SO_2) and a Wuhan Cubic PM3006S for PM ($PM_{2.5}$ and PM_{10}) (Oizom -
- 134 Polludrone Smart, 2023). It also registers RH and Temp, but no data was found in regards to sensor model/brand.
- 135 The sampling mechanism uses a fan and data transmission is wireless. The final product (time res is 10-min) can
- be obtained through the Oizom webpage and/or via API.
- 137 <u>Kunak Air Pro (https://www.kunak.es/)</u> uses a fan for sampling and all sensors are from Alphasense (EC, B series
- for CO, NO, NO₂ and O₃; an NDIR sensor for CO₂; and an OPC-N3 for PM₁, PM_{2.5}, and PM₁₀) (Hofman et al.,
- 139 2022). It also provides Temp, RH, and Press (no data was found in regards to environmental sensor model/brand).
- 140 The raw data is transmitted via a multi-band network, and the final data (time res is 5-min) can be accessed through
- 141 their website or via API.
- 142 The <u>Silax Air (https://vortexiot.com</u>) system measures NO₂, O₃, PM₁₀ and PM_{2.5}. Their webpage mentions that for
- 143 PM an optical scattering sensor is used and EC sensors for the gases. Further details weren't found. The raw data
- 144 is transmitted via 4G or WiFi and the final user accesses the final product (5-min time res) through API or website.
- 145 The <u>Node-S</u> system (<u>https://www.clarity.io</u>) holds a nephelometer (Plantower PMS6003) to measure 3 PM size
- 146 cuts (PM₁, PM_{2.5}, PM₁₀) (Liu et al., 2022) and EC sensors for NO₂ (Alphasense) (Miech et al., 2021). The air is
- 147 dragged into the system by a fan and a Bosch sensor is used for press, RH, and temp. The data is communicated
- 148 to Clarity's cloud via cellular signal (4G) and the final product is ~3-min time res (something unusual for sensor
- systems). Access to the final data is via the web portal or through API.
- 150 Praxis/Urban (https://www.southcoastscience.com) system employs EC sensors for NO, NO₂, O₃ (Alphasense, A
- series), an NDIR for CO₂ (Alphasense), and particle counter (Alphasense, OPC-N3) for PM₁, PM₁₀ and PM_{2.5}.
- 152 The Temp/RH is Sensirion and the Press sensor is TDK. The raw data is communicated to the company server
- using 4G and the user can access it and post-processed data through an API (1-min time res).

154	Table S5. Summary of sensor measurements and the time resolution data provided by participating companies in the
155	WPS study.

System	Measurands	Time Resol.
Mod	PM ₁ , PM _{2.5} , PM ₁₀	1min
AQM	PM1, PM2.5, PM10, NO, NO2, O3, CO; CO2	15min
Atm	PM ₁ , PM _{2.5} , PM ₁₀	2min
IMB	PM ₁ , PM _{2.5} , PM ₁₀ , NO ₂ , O ₃	1min
Poll	PM1, PM2.5, PM10, NO, NO2, O3	10min
AP	PM1, PM2.5, PM10, NO, NO2, O3, CO; CO2	5min







Figure S3. Data product for each of the participating companies in the WPS. The top panels are for NO₂, the middle panels for O₃ and the bottom panels for PM_{2.5}. The y-axis represents the different products: "out-of-box", call and

160 cal2. The x-axis shows the dates for which each company provided the mentioned products.

161 Sect. S5. Performance Metrics

162 In the assessment of sensor measurement error, it is standard practice to employ a linear additive model, described163 by the following equation:

$$164 y_i = b_1 x_i + b_0 + \varepsilon_i (S1)$$

In this model, the dependent variable "y" represents the sensor measurements, while the independent variable "x" denotes the reference measurements. The coefficient b_1 corresponds to the slope of the regression line (the response sensitivity of the sensor relative to the reference) and b_0 is the ordinate at the origin (the sensor's output when the reference measurement is zero). ε_i , assumed to have a mean of zero and a standard deviation of σ_{ε} , captures the portion of "y" that cannot be explained by "x". For a sensor to perfectly match the reference measurements (i.e., y = x), b_1 would equal one, with both b_0 and ε_i being zero.

172 Coefficient of Determination (R^2)

173 R² is an adimensional metric that quantifies the proportion of variance in the sensor measurements ("y") that can
174 be explained by its linear relationship with the reference measurements ("x"):

175
$$R^{2} = \frac{\sum_{i=1}^{n} (x_{i} - \hat{y})^{2}}{\sum_{i=1}^{n} (y_{i} - \hat{y})^{2}}$$
(S2)

As a bounded metric, R^2 varies between zero and one ($0 \le R^2 \le 1$), where a value closer to 176 177 one indicates a stronger linear association between the sensor and reference 178 data. Despite being one of the most widely used metrics in sensor evaluation, as 179 highlighted by Karagulian et al. (2019), R^2 comes with limitations that warrant careful 180 consideration. Notably, R² does not account for bias in the data; a regression line diverging from the ideal 1:1 181 relationship between "x" and "y" does not affect its value. Additionally, R² is influenced by the dynamic range of 182 the measurements, which can skew its interpretation. Given these nuances, it is prudent to report R² alongside 183 complementary metrics that can offer a more rounded view of sensor performance. For a more in-depth analysis 184 of the limitations and proper use of R², readers are directed to the discussion in Legates and McCabe Jr. (1999).

185 Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE)

MAE and RMSE (both dimensional metrics, expressed in the same units as the measured variable), also stand as
 very popular metrics for performance evaluation, as they offer insights into the accuracy of sensors, presenting a
 fuller picture than the R² alone. These metrics can be estimated as follows:

189
$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - x_i|$$
 (S3)

190
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - x_i)^2}$$
 (S4)

191

192 However, both MAE and RMSE quantify average errors. MAE does so by calculating the average magnitude of 193 errors without directionality, utilising absolute differences, while RMSE gauges the standard deviation of these 194 differences, highlighting the squared differences between sensor readings and reference grade measurements. 195 Although MAE and RMSE are both valued for their measure of accuracy, they bear distinct implications in 196 practice. MAE treats all errors equally, allocating proportional weight across the board. Conversely, RMSE 197 disproportionately penalises larger errors due to its squaring of difference values, an aspect noted by (Willmott 198 and Matsuura, 2005). This characteristic makes RMSE particularly sensitive to outliers, shaping its utility in 199 identifying and rectifying significant deviations.

200 Mean Bias Error (MBE)

The MBE quantifies the average bias in sensor measurements relative to reference values. Expressed in the sameunits as the variable being measured, MBE reflects the systematic error, offering a straightforward indication of a

sensor's tendency to overestimate or underestimate the reference:

204
$$MBE = \frac{1}{n} \sum_{i=1}^{n} (y_i - x_i)$$
 (S5)

A zero value of MBE indicates no consistent over- or underestimation, while positive or negative values signal systematic bias in measurement. This simplicity in interpretation makes MBE particularly valuable for initial assessments of sensor accuracy and for guiding calibration efforts to correct for systematic bias. However, the MBE does not capture the precision of the measurements. For this reason, MBE is most effective when used in conjunction with other metrics, such as RMSE and MAE, to gain a comprehensive understanding of sensor performance, encompassing both systematic and random errors.

211 Relative Expanded Uncertainty (REU)

212 In contrast to single-value metrics such as R², RMSE, and MAE, which assess data sets as a whole, REU offers a

213 "point by point" metric. This allows for graphical representations (like the REU in the concentration space or as

a time series), offering detailed insights into measurement performance variability. The REU's mathematical

215 framework is outlined in the "Guidance for the Demonstration of Equivalence of Ambient Air Monitoring

216 Methods" (European Commission, 2010), as follows:

217
$$U(y_i) = \sqrt{\frac{RSS}{n-2} - u^2(x_i) + (x_i - b_0 - b_1 x_i)^2}$$
(S6)

218
$$REU(y_i) = \frac{k \cdot U(y_i)}{\hat{x}}$$
(S7)

219
$$RSS = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2$$
 (S8)

here, U(y_i) represents the measurement uncertainty [concentration units]; REU(y_i) denotes the REU [percentage];
u(x_i) is the random uncertainty of the reference monitor [concentration units]; "n" stand for the number of
collocated data points considered; RSS is the Residual Sum of Squares; k is the coverage factor (set at 2 for a 95%
confidence level).

A distinctive feature of REU is its incorporation of the uncertainty associated with the reference method (i.e., u(x_i)). This aspect recognizes that all measurements, including those from reference methods, are subject to inherent uncertainties. While calculating REU is more complex than traditional metrics, it's essential to acknowledge that, like any metric, REU is based on specific assumptions and considerations. These factors must be thoughtfully evaluated when interpreting data to ensure that conclusions are firmly rooted in the context of the study.

230 Current guidance and normalisation efforts

Table S6 summarises the key metrics addressed in some of the most recent guidance documents and technical standards. These metrics have been categorised under various labels: linearity, bias, error, uncertainty, data coverage, and inter-sensor precision. Each of these guidelines and regulations has its own set of procedures, protocols, and thresholds. Therefore, it is advisable for readers to consult the original documents for a detailed understanding of these specificities.

237	Table S6. Summary of field evaluation metrics for senso	rs according to different guidelines and technical standards
-----	---	--

EPA ^{1&2}	CEN ³	ASTM ^{4&5}
PM _{2.5} & O ₃	NO ₂ , O ₃ , CO, SO ₂ & Bencene	PM _{2.5} , PM ₁₀ NO ₂ , O ₃ , CO & SO ₂
\mathbb{R}^2		R ²
Slope	Slope	Slope
Intercept	Intercept	Intercept
		MAE
RMSE		RMSE
NRMSE		NRMSE
	REU	
Data	Data	Data
completeness	Capture	Capture Rate
SD	u _(bs,s)	$\mathbf{S}_{\mathrm{r,f}}$
CV		
	EPA ^{1&2} PM _{2.5} & O ₃ R ² Slope Intercept RMSE NRMSE Data completeness SD CV	EPA1&2CEN3PM2.5 & O3NO2, O3, CO, SO2 & BenceneR2SlopeSlopeInterceptInterceptInterceptInterceptRMSENRMSEREUDataDataDatacompletenessCaptureSDu(bs,s)CV

238 <u>References in the table:</u>

239 ¹EPA/600/R-20/279 Performance Testing Protocols, Metrics, and Target Values for Ozone Air Sensors.

- ²EPA/600/R-20/280 Performance Testing Protocols, Metrics, and Target Values for Fine Particulate Matter
 Air Sensors.
- ³CEN/TS 17660-1: Air quality Performance evaluation of air quality sensor systems Part 1 Gaseous pollutants in ambient air.
- ⁴ASTM D8406-22: Standard Practice for Performance Evaluation of Ambient Outdoor Air Quality Sensors
 and Sensor-based Instruments for Portable and Fixed-point Measurement.
- ⁵ASTM WK74812: Standard Specification for Ambient Outdoor Air Quality Sensors and Sensor-based
 Instruments for Portable and Fixed-Point Measurement.
- Acronyms: EPA: U.S. Environmental Protection Agency; CEN: European Committee for Standardization;
 ASTM: American Society for Testing and Material. CV: Coefficient of Variation; SD: Standard Deviation
 (see the definition in the EPA Performance Testing Protocols); u_(bs,s): Between sensor system uncertainty
 (see the definition in the CEN TS 17660-1); S_{r,f}: field reproducibility standard deviation (see the definition
 in the ASTM protocols).



Figure S4. Inter-device precision of NO₂ measurements from "identical" devices across the 4 companies participating in QUANT is assessed using the "between sensor system uncertainty" metric (defined by the CEN/TS 17660-1:2021 as *u*(*bs*, *s*)). Each line represents this metric as a composite of all sensors per brand (excluding units with less than 75% data) within a 40-day sliding window.



Figure S5. The inter-device precision of O₂ measurements from "identical" devices across the 4 companies participating in QUANT is assessed using the "between sensor system uncertainty" metric (defined by the CEN/TS 17660-1:2021 as u(bs, s)). Each line represents this metric as a composite of all sensors per brand (excluding units with less than 75% data) within a 40-day sliding window.



Figure S6. Comparative analysis of "Sensor A" performance against two reference instruments for NO₂ measurements. The left plot shows the correlation with the Teledyne T500 (Cavity Attenuated Phase Shift Spectroscopy), while the right plot is against the Teledyne T200U (chemiluminescence) and specifically installed at the Manchester supersite for the QUANT study. The dashed red line represents the line of best fit for the sensor data against each reference, indicating a closer agreement with the T200U (slope=1.02) compared to the T500 (slope=0.73).



271

Figure S7. Comparative regression analysis and performance metrics of two distinct PM_{2.5} sensor systems benchmarked against a BAM for the top plots and a Fidas for the bottom plots. Each plot demonstrates the correlation and agreement between the sensor readings and the two equivalent-to-reference instruments in a roadside site located in York.

276 Sect. S7. NO₂ Diffusion tubes

277 A diffusion tube co-location study was carried out between November 2020 and November 2021 at the MAQS, 278 LAQS and York sites, using two types of diffusion tubes: the conventional (also known as LAQM, for Local Air 279 Quality Management) and UUNN (for UK Urban NO2 Network). LAQM tubes have an open end and capture 280 NO₂ which is converted to nitrite when reacting with triethanolamine (TEA) for subsequent analysis. On the other 281 hand, UUNN tubes, similar in the sampling process to LAOM, include an amorphous polyethylene filter at the 282 open end to further mitigate the effect of wind on NO₂ measurements. For more details refer to (Butterfield et al., 283 2021). Both types of tubes (conventional and UUNN) were installed in duplicates, either in shelters (to limit the 284 incidence of wind) or directly exposed without protection in mounting blocks. Figure S5 illustrates the

- 285 performance comparison of traditional diffusion tubes and a sensor system in Manchester. The data from these
- diffusion tubes have been used to correct the sensor shown here and explained in detail in Section 3.6 (Figures 9b)
- 287 and 9c).





Figure S8. The left plot displays the correlation between an air quality sensor's readings and those from a reference monitor for NO₂, while the right plot demonstrates the LAQM diffusion tube performance. The LAQM plot shows a tighter correlation with the 1:1 line, indicating a higher accuracy in measuring NO₂ concentrations for the period Nov 2020 - Nov 2021 at the Manchester supersite (blue dots represent monthly averages).

293 References

- Butterfield, D., Martin, N. A., Coppin, G., and Fryer, D. E.: Equivalence of UK nitrogen dioxide diffusion tube
 data to the EU reference method, Atmos. Environ., 262, 118614,
- 296 https://doi.org/10.1016/j.atmosenv.2021.118614, 2021.
- 297 Cross, E. S., Williams, L. R., Lewis, D. K., Magoon, G. R., Onasch, T. B., Kaminsky, M. L., Worsnop, D. R.,
- and Jayne, J. T.: Use of electrochemical sensors for measurement of air pollution: correcting
- interference response and validating measurements, Atmospheric Meas. Tech., 10, 3575–3588,
- **300** https://doi.org/10.5194/amt-10-3575-2017, 2017.
- 301 European Commission: Guide to the demonstration of equivalence of ambient air monitoring methods, Report
 302 by an EC Working, Group on Guidance. European Commission, 2010.
- 303 Hofman, J., Peters, J., Stroobants, C., Elst, E., Baeyens, B., Van Laer, J., Spruyt, M., Van Essche, W., Delbare,
- 304 E., Roels, B., Cochez, A., Gillijns, E., and Van Poppel, M.: Air Quality Sensor Networks for Evidence-
- **305** Based Policy Making: Best Practices for Actionable Insights, Atmosphere, 13, 944,
- **306** https://doi.org/10.3390/atmos13060944, 2022.
- 307 Karagulian, F., Barbiere, M., Kotsev, A., Spinelle, L., Gerboles, M., Lagler, F., Redon, N., Crunaire, S., and
- 308 Borowiak, A.: Review of the Performance of Low-Cost Sensors for Air Quality Monitoring,
- 309 Atmosphere, 10, 506, https://doi.org/10.3390/atmos10090506, 2019.

- 310 Legates, D. R. and McCabe Jr., G. J.: Evaluating the use of "goodness-of-fit" Measures in hydrologic and
- 311 hydroclimatic model validation, Water Resour. Res., 35, 233–241,
- **312** https://doi.org/10.1029/1998WR900018, 1999.
- 313 Liu, G., Moore, K., Su, W.-C., Delclos, G. L., Gimeno Ruiz de Porras, D., Yu, B., Tian, H., Luo, B., Lin, S.,
- 314 Lewis, G. T., Craft, E., and Zhang, K.: Chemical explosion, COVID-19, and environmental justice:
- 315 Insights from low-cost air quality sensors, Sci. Total Environ., 849, 157881,
- **316** https://doi.org/10.1016/j.scitotenv.2022.157881, 2022.
- Meyer, M., Afshar-Mohajer, N., Cross, E., and Mudgett, P.: Feasibility of using Low-Cost COTS Sensors for
 Particulate Monitoring in Space Missions, 2022.
- Miech, J. A., Stanton, L., Gao, M., Micalizzi, P., Uebelherr, J., Herckes, P., and Fraser, M. P.: Calibration of
 Low-Cost NO2 Sensors through Environmental Factor Correction, Toxics, 9, 281,
- **321** https://doi.org/10.3390/toxics9110281, 2021.
- Myklebust, H., Aarhaug, T. A., and Tranell, G.: Use of a Distributed Micro-sensor System for Monitoring the
 Indoor Particulate Matter Concentration in the Atmosphere of Ferroalloy Production Plants, JOM, 74,
 4787–4797, https://doi.org/10.1007/s11837-022-05487-7, 2022.
- 325 Oizom Polludrone Smart: http://www.aqmd.gov/aq-spec/sensordetail/oizom---polludrone-smart, last
 326 access: 27 January 2023.
- 327 Ouimette, J. R., Malm, W. C., Schichtel, B. A., Sheridan, P. J., Andrews, E., Ogren, J. A., and Arnott, W. P.:
- Evaluating the PurpleAir monitor as an aerosol light scattering instrument, Atmospheric Meas. Tech.,
 15, 655–676, https://doi.org/10.5194/amt-15-655-2022, 2022.
- 330 Puttaswamy, N., Sreekanth, V., Pillarisetti, A., Upadhya, A. R., Saidam, S., Veerappan, B., Mukhopadhyay, K.,
- 331 Sambandam, S., Sutaria, R., and Balakrishnan, K.: Indoor and Ambient Air Pollution in Chennai, India
- during COVID-19 Lockdown: An Affordable Sensors Study, Aerosol Air Qual. Res., 22, 210170,
- **333** https://doi.org/10.4209/aaqr.210170, 2022.
- Weissert, L. F., Alberti, K., Miskell, G., Pattinson, W., Salmond, J. A., Henshaw, G., and Williams, D. E.: Low cost sensors and microscale land use regression: Data fusion to resolve air quality variations with high
- **336** spatial and temporal resolution, Atmos. Environ., 213, 285–295,
- **337** https://doi.org/10.1016/j.atmosenv.2019.06.019, 2019.
- Westgate, S. and Ng, N. L.: Using in-situ CO2, PM1, PM2.5, and PM10 measurements to assess air change
 rates and indoor aerosol dynamics, Build. Environ., 224, 109559,

- 340 https://doi.org/10.1016/j.buildenv.2022.109559, 2022.
- 341 Willmott, C. J. and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error
- 342 (RMSE) in assessing average model performance, Clim. Res., 30, 79–82,
- 343 https://doi.org/10.3354/cr030079, 2005.
- 344 Zauli-Sajani, S., Marchesi, S., Boselli, G., Broglia, E., Angella, A., Maestri, E., Marmiroli, N., and Colacci, A.:
- 345 Effectiveness of a Protocol to Reduce Children's Exposure to Particulate Matter and NO2 in Schools
- during Alert Days, Int. J. Environ. Res. Public. Health, 19, 11019,
- 347 https://doi.org/10.3390/ijerph191711019, 2022.