***Interactive comment* on "Characterising low-cost sensors in highly portable platforms to quantify personal exposure in diverse environments" by Lia Chatzidiakou et al.**

Response to anonymous referee #3

Overall response: We would like to thank reviewer #3 for their useful and insightful comments. The reviewer recognises the importance of the manuscript and asks for technical clarifications. However, we would like to emphasize that this manuscript is not aiming to provide a *generic* evaluation of all portable air quality sensor types but instead aims to characterise the performance of this specific configuration with currently available sensor variants for specific environmental conditions, as a necessary precursor to a series of papers on exposure and health impacts using these PAMs. While the various methodologies we used would be of value to the scientific community, as the field of air quality monitoring evolves, different sensor types, models and configurations might be expected to lead to different results.

Detailed response

Abstract 1: "*It would appear beneficial to include brief selected quantitative summary information on the precision, accuracy…."*

Response (added): ….Overall, the air pollution sensors showed high reproducibility (R^2: 0.80 -1.00, mean R^2 = 0.93) and excellent agreement with standard instrumentation (R^2: 0.56 – 0.99, mean R^2 = 0.82) in outdoor, indoor and commuting microenvironments across seasons and different geographical settings.

Introduction 2: "this paper further aims to create a roadmap for calibration and validation of portable monitors suitable for personal exposure quantification…. *wondered if this aim could be stated in a clearer way?"*

Response (re-phrased as follows pg 2 ln 35-37): As the field of novel air pollution sensing technologies expands rapidly, this paper further aims to provide methodological guidance to researchers from diverse disciplines on how to comprehensively calibrate and validate portable monitors suitable for personal exposure quantification.

Methods 3. "*it would be helpful to clarify if the monitoring systems being tested are [or will be] commercially available to wider research communities …. also there may be a substantial number of person years invested in the design and construction of the monitoring systems…"*

Response (added as follows pg 2 ln 39-41): The PAM has been developed at the Department of Chemistry, University of Cambridge in collaboration with Atmospheric Sensors Ltd. It is now commercially available (independently from the University of Cambridge) from Atmospheric Sensors Ltd (Model AS520, http://www.atmosphericsensors.com).

Methods 4 and 5. "*It would be helpful to expand on the principle of operation of the gas sensors used… accessible to a wider audience… e.g. a diagram similar to diagrams of 3-electrode systems in earlier references"* and on the same issue of 4-electrode EC sensors "*Is it ok to directly translate the finding to the newer 4-electrode sensor, or is separate quantification of the linearity and LoD of the newer sensor also required?*"

Response: (added in paper pg4 ln1-10): These EC sensors are the A4 variant from Alphasense (ref) and operate on a four electrode system. The principle of operation of the four electrode system is identical to that of the earlier variants of three-electrode system with an additional electrode, the auxiliary (or non-sensing) electrode to compensate for the temperature dependence of the cell potential. …. The linearity and LoD of the 4-electrode sensors (when integrated in the PAM) have been tested under laboratory conditions following the same methodology as described in Mead et al. 2013, yielding very similar results.

Response to referee: Rather than include inevitably incomplete sensor details (as much is proprietary to the manufacturer), we feel it more appropriate to refer readers (and the

referee) to the manufacturer's website for further details about the principles of operation of the individual sensors integrated in the PAM. The 4th electrode was implemented to correct for systematic errors and not to change the LOD. Hence, we expect the linearity and LOD to be comparable with those of the 3-electrode sensors.

Methodology 6. ***"****A strength of your model is that you use a conceptually simple model……present results of the equations fitted… in Supplementary Information".*

Response to referee: Such calibration factors are likely to be location specific. Making them public, would risk encouraging other researchers taking them as generic.

Methodology 7. *"It would appear useful to make your field calibration data available"*

Response: Data availability: https://doi.org/10.17863/CAM.41918

Methodology 8. "*Add references for Mie"*

Response (pg5 ln 7-8, references added):

We have included a suitable reference.

Mie, G. Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen. *Ann. Phys.* 1908, *330*, 377–445.
Bohren, C. F.; Huffman, D. R. Absorption and scattering of light by small particles; *Wiley*, 1988.

Methodology 9 &10. "*robust objective way of splitting the data to calibration-validation periods"*…. "*Currently it is not clear what calibration data has been used to adjust the sensor outputs in these figures*"

Response (added to the paper):

To evaluate the performance of the linear model, the datasets were split into training (i.e. calibration) and validation periods to first extract the calibration parameters and then apply them to the validation set and compare the measurements with those from reference instruments (referred to as "calibration - validation" method). The training sets ranged from 1 to 16 days, and the adjusted coefficient of determination ($\overline{R}^2$) remained stable for training periods longer than 3 days. Therefore, approximately a third of the dataset was selected as a training set. As relationships in these linear models should ideally not be extrapolated beyond the range of the observations (including meteorological conditions), the calibration periods covered the temperature and concentration ranges in which the sensors were deployed (Cross et al., 2017). Once the performance of the model was established in diverse environments, we used the full co-location periods to determine the agreement between PAM sensors and reference instruments.

Response to referee (not in the paper):

Calibration of gaseous sensors for the China deployment: The China deployment lasted about one year, and no evidence of a significant drift of the gaseous sensors was found. Using a unified **training** dataset from the outdoor co-locations **both in the heating and non-heating season** the coefficients of the linear model were determined. As the variation between seasons was greater than the variation within seasons and covered a wide range of environmental conditions and pollution levels, the choice of the training period did not affect the performance of the model. The coefficients determined in the training set were then applied to the **validation** periods of both seasons and were also used to convert raw measurements to ppb in the indoor deployment. In this way, we prove that the calibration parameters determined in outdoor co-locations are suitable across a range of environmental conditions and for indoor deployments too, and therefore for personal exposure as people spend as much as 90% of their time in indoor environments.

Calibration for the UK deployment: The PAMs were co-located outdoors next to reference instruments in 2015, and were then deployed for two years continuously to participants of the COPE project. The PAMs were again co-located next to reference instruments in 2017 after the completion of the project. The change in the sensor sensitivities

did not show a systematic error. Individual sensor drifts were quantified and were corrected by linearly interpolating sensitivities between the start and the end of the deployment.

Results 12. "*Dates and measurement times in Table 3"*
Response: Added to Table 3 first column

Colocation	Start	End	Total time
China (Dec-Jan)	28/12/2016	15/01/2017	447 hours
UK (Oct-Nov)	27/10/2017	13/11/2017	408 hours
China (June)	28/06/2017	16/07/2017	432 hours
UK (April-May)	26/03/2018	10/04/2018	342 hours

Results 13 & 14. "*Express RMSE as percentage of mean reference"* **–**
Response: We have added mean values of each pollutant during the co-location periods. We report RMSE as a percentage of the maximum (as the 95% percentile), as peak exposure events are more relevant in health studies. The reader is now also able to calculate the RMSE as percentage of the mean reference values.

Results 15. pg 11 ln 6 *"were the extreme temperatures recorded inside the monitor enclosure and within the sensors?"*
Response to referee: The temperature inside the PAM is on average 7℃ higher than ambient temperatures. The monitors were co-located in metal shelters on the roof of a research building in the Peking University campus (Figure below), and were exposed to direct sun further increasing the temperature inside the PAM.

Figure (not included in manuscript): Co-location of 60 PAMs at PKU during the heating and non-heating season for a four-week period. PAMs were placed in custom-made protective shelters in proximity to the inlets of reference instruments and were connected on mains.
Response added: In Table 3 added "internal conditions of the PAM"

Results 16. "*Why was O_3 sensor less affected by heat then NO_x and CO? - O_3 concentrations were higher than usual?"*
Response to referee: The principles of operation of the NO_2 and O_3 sensors are very similar. It is possible that the O3 sensors perform better because (a) the concentrations of O_3 were higher in the summer and (b) the sensor is further away from the internal battery of the PAM and therefore slightly cooler than the other sensors. The next generation of PAM hardware will integrate an insulation layer between battery and sensors. NO was close to the LOD (maximum= 5ppb, mean =1 ppb).

Results 17. "*Your reference to the participants on p11 could be given with further detail e.g. geographical location / time of year etc. to provide context to the comparison with the field evaluation measurements."*
Response: FIGURE A5 CAPTION Added: The co-location with the reference instruments on the roof of Peking university took place from 28/06/2017 to 16/07/2017, the field deployment was conducted in Beijing and Pinggu from 22/05/2017 to 26/06/2017.

Results18. *"In Figure 4 caption it would be helpful to specify how far the 'PKU reference site' and the 'nearby government monitoring site' were from the indoor measurement location"*

Response ADD TO FIG 4 CAPTION: PKU roof is 5.3 km and the governmental reference station (Haidianwanliu) is 6 km away from the location of the indoor experiment.

Results 19. "*On p12 you assert that your results prove the suitability of the low cost sensors for quantification of indoor air pollution. It would seem appropriate to qualify the promising indication of suitability over timescales and conditions similar to the calibration period/conditions?*"

Response: Add to PAGE 12 LINE 34:

The conversion of the raw measurements to ppb used the sensitivities extracted using outdoor co-locations both during the heating and non-heating season (subsection 3.1) with the linear model (subsection 2.1).

….proving the suitability of this monitoring platform to quantify indoor air pollution levels for these species provided they have been adequately calibrated in the local environment.

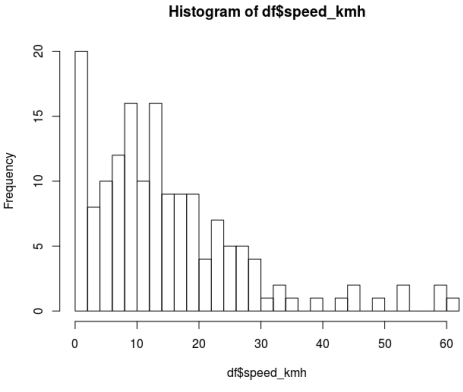
Results 20. "*Do you have information on the response times of the sensors…"*

Response to referee: Response times have been determined in the Alphasense Technical Specification sheets of each sensor (see references in Section 2.1, page 3) and ranged between t_{90} < 25s for CO (from 0 to 10ppm) and t_{90} < 60s for O_3 (from 0 to 1ppm). As the response times are smaller than the time resolution for the PAM measurements we do not expect them to affect the way of correcting the sensors' cross sensitivities.

Results 21. "*When the PAM was mounted on the roof of the vehicle how were the sensor inlets orientated in relation to airflow? Did the varying speed of airflow have any effect on agreement between sensors and reference instruments?"*

Response (added to manuscript pg 15): The PAM was mounted on the roof with the OPC inlet facing forwards and the EC sensors facing to the sides (see photo below). The reference instrument inlets were located on the car roof as well. The speed mostly varied between 5 and 20 km/h (see histogram below). There was no correlation between car speed and RMSE error in the gaseous and particulate measurements. The OPC contains an airflow measurement unit which compensates for any wind or internal flow dependence.

The histogram below shows the distribution of the speed measurements covering the same time period as Figure 6 in Section 3.2.



Results 22. *"11.30- 12.30 e.g. was vehicle static Do you know why O3 measured by sensor and reference instruments diverges during this time period?"*

Thank you for bringing this to our attention. During this period the vehicle was stationary, and we have therefore removed this data.

Response to anonymous referee #2

Overall response: Thank you for taking the time to provide useful comments. Referee #2 mentions that this manuscript is essential to underpin the validity of personal exposure measurements collected in two major health studies (APHH-Beijing, Theme3: AIRLESS https://doi.org/10.5194/acp-19-7519-2019 and COPE study https://doi:10.1136/bmjopen-2016-011330). The specific aim of the manuscript is therefore to evaluate a specific sensor package (the PAM) with a comprehensive, robust and reproducible methodology, rather than individual sensors or a generic sensor package.

The referee states that "*there are many sensor studies worldwide in recent years*", feeling this work is not novel. However, there are concerns remaining in the scientific community regarding the validity of measurements collected with miniaturised portable sensors. For example, **a recent literature review on portable sensors** (Thompson, 2016 https://doi.org/10.1016/j.teac.2016.06.001) states that "*current technology for inexpensive portable sensors is not sensitive or specific enough to meet demands*" while a commentary article in Nature 2016 (https://www.nature.com/news/validate-personal-air-pollution-sensors-1.20195) disregards novel technologies due to "*questionable air quality data*". Such opinions act as a barrier in adopting innovative methods that could revolutionise multiple disciplines including epidemiological research and the built environment and have significant societal benefits. Extending beyond the specific aim outlined above, we feel that this manuscript does also contribute significantly and positively to the wider literature of novel portable sensor technologies.

Detailed response: Temperature and relative humidity were not included explicitly in the linear model for the calibration of the electrochemical sensors. The effect of relative humidity on particulate matter estimations has been quantified in a previous publication (Di Antonio, A., et al., 2018. *Sensors*, https://doi.org/10.3390/s18092790). The cross-interference of other gases on the electrochemical sensors is covered in the manufacturers specifications and is beyond the scope of this work.

Calibration periods were selected based on campaign time periods not conflicting with deployments of the PAM to participants. The training set was about 1/3 of the total observations, an arbitrary choice. We used a combined training set from the winter and the summer co-locations. In that way, the selected training periods of each season become less important as the variation in pollutant levels between seasons is much greater providing the necessary wide range of calibration conditions.

The vehicle deployment aimed to evaluate the performance of the PAM in movement and did not aim to capture personal exposure of an individual within a vehicle. Forthcoming publications focus on the magnitude and duration of personal exposure in diverse microenvironments (including indoor locations and different commuting modes) during daily life activities.

The selected references on static outdoor co-locations are inevitably selective, and are not exhaustive of the large body of evidence on novel technologies. However, there is a lack of publications on the performance of portable platforms in diverse microenvironments, as presented in this manuscript.

Prices for individual sensors can be provided by the manufacturers. Low R^2 values especially for the NO and NO_2 sensors were noticed at temperatures above 40 C (non-heating season in Beijing), which is above operational specifications and were not recorded during the participant deployment.

Response to anonymous referee #1

Overall response: Thank you for taking the time to provide valuable feedback. The aim of this paper was to validate the performance of a specific novel personal air pollution monitor (PAM) when capturing personal exposure. While outdoor co-locations next to certified instruments have been widely adopted by researchers and governmental organisations to validate the performance of sensors in the field, this paper goes beyond those current guidelines by validating the PAM in indoor and commuting microenvironments, and thus demonstrating that novel sensing technologies can provide reliable personal exposure measurements.

The importance of this work is two-fold:

1) **Addressing concerns which remain in the scientific community regarding the suitability of novel sensing technologies for policy purposes and health studies.** Such opinions act as a barrier in adopting innovative methods that could have significant societal benefits. In that sense,
2) **This paper is the first of a series of publications that, together with detailed medical outcome determinations, aim to identify underlying mechanisms of specific air pollutants on health, and is necessary to validate in the open literature the performance of the PAM.** Forthcoming publications will also focus on the modification effects of the indoor environment on personal exposure.

Detailed comments:

(1). OPC corrections: The detailed RH correction algorithm can be found in: Di Antiono, A., et al., 2018. *Sensors*, https://doi.org/10.3390/s18092790. A constant density was assumed for both, the reference instrument (Fidas Palas 200S, see Table 2, pg 7) and the portable instrument. It is true that in general, scaling to the reference compensates for density effects.

(2). Sensitivity drifts of sensors: The PAMs were co-located at the beginning and at the end of fieldwork with the reference instruments at the Department of Chemistry, UCAM. The change in sensitivities of the gaseous sensors was less than 10% and was therefore not included in the manuscript. The topic has been covered in a previous publication of the group that found similar drifts over an 11-month period (Mead, M.I., 2013, *Atmospheric Environment*, https://doi.org/10.1016/J.ATMOSENV.2012.11.060).

(3). Figure 2: This figure is an illustrative example of the methodology used to validate the performance of linear models used to convert raw units to physical measurements. The Figure presents calibrated data.

(4) (a) Chinese deployment during the non-heating season: Results are presented in italics due to the exposure of the sensors to very high temperatures. However, such temperatures were not encountered during the deployment to participants. (b) RMSE and adjusted R^2 for the combined training and testing period. (c) We used a combined training set from the winter and the summer co-locations. The training set was about 1/3 of the total observations. The sensitivities from the outdoor co-locations were then used to calibrate the indoor measurements: This proves that provided there is a diverse enough training set (both in terms of temperature and pollution levels) the linear model performs adequately in different conditions. In that way, the selected training period of each season becomes less important as the variation between seasons is much greater providing the necessary wide range of calibration conditions.

Characterising low-cost sensors in highly portable platforms to quantify personal exposure in diverse environments.

Lia Chatzidiakou[1], Anika Krause[1], Olalekan A.M. Popoola[1], Andrea Di Antonio[1], Mike Kellaway[2], Yiqun Han[3,4,5], Freya A. Squires[6], Teng Wang[4,7], Hanbin Zhang[3,5,8], Qi Wang[4,7], Yunfei Fan[4,7], Shiyi Chen[4], Min Hu[4,7], Jennifer K. Quint[9], Benjamin Barratt[3,5,8], Frank J. Kelly[3,5,8], Tong Zhu[4,7], Roderic L. Jones[1]

[1]Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW, UK
[2]Atmospheric Sensors Ltd, Bedfordshire, SG19 3SH, UK
[3] MRC-PHE Centre for Environment & Health, Imperial College London and King's College London, London, W2 1PG, UK
[4] College of Environmental Sciences and Engineering, Peking University, Beijing, 100871, China
[5] Department of Analytical, Environmental and Forensic Sciences, King's College London, London, SE1 9NH, UK
[6] Department of Chemistry, University of York, York, YO10 5DD, UK
[7]The Beijing Innovation Center for Engineering Science and Advanced Technology, Peking University, Beijing, 100871, China.
[8] NIHR Health Protection Research Unit in Health Impact of Environmental Hazards, King's College London, London, SE1 9NH, UK
[9] National Heart and Lung Institute, Imperial College London, SW3 6LR, UK

Correspondence to: Dr Lia Chatzidiakou, email: ec571@cam.ac.uk, Tel: 01223 336345

Abstract. The inaccurate quantification of personal exposure to air pollution introduces error and bias in health estimations, severely limiting causal inference in epidemiological research worldwide. Rapid advancements in affordable, miniaturised air pollution sensor technologies offer the potential to address this limitation by capturing the high variability of personal exposure during daily life in large-scale studies with unprecedented spatial and temporal resolution. However, concerns remain regarding the suitability of novel sensing technologies for scientific and policy purposes. In this paper we characterise the performance of a portable personal air quality monitor (PAM) that integrates multiple miniaturised sensors for nitrogen oxides (NO_x), carbon monoxide (CO), ozone (O_3) and particulate matter (PM) measurements along with temperature, relative humidity, acceleration, noise and GPS sensors. Overall, the air pollution sensors showed high reproducibility (mean $\overline{R}^2$=0.93, min-max: 0.80 -1.00) and excellent agreement with standard instrumentation (mean $\overline{R}^2$=0.82, min-max: 0.54-0.99) in outdoor, indoor and commuting microenvironments across seasons and different geographical settings. An important outcome of this study is that the error of the PAM is significantly smaller than the error introduced when estimating personal exposure based on sparsely distributed outdoor fixed monitoring stations. Hence, novel sensing technologies as the ones demonstrated here can revolutionise health studies by providing highly resolved reliable exposure metrics at large scale to investigate the underlying mechanisms of the effects of air pollution on health.

Commented [a1]: 0.56 in response to referee – correct value is 0.54

Keywords. personal exposure, portable air quality monitor, miniaturised sensor technologies, nitrogen oxides (NO_x), carbon monoxide (CO), particulate matter (PM)

1 Introduction

Emerging epidemiological evidence has associated exposure to air pollution with adverse effects on every major organ system (Thurston et al., 2017). Most of this evidence comes from western Europe and North America (Newell, Kartsonaki, Lam, & Kurmi, 2017) as population-scale air pollution health studies have largely relied on available outdoor air pollution measurements from fixed monitoring stations (COMEAP, 2018). Due to limitations in the availability of monitoring networks in low- and middle-income countries (LMICs), the effects of air pollution on health have been under-researched in these settings. A clear need exists for more direct epidemiological evidence in diverse geographical settings with varying air pollution sources considering the high likelihood that health effects of air pollution are not linear, and cannot be simply transcribed from the western world to LMICs (Tonne, 2017).

Secondly, the low spatial and temporal resolution of exposure metrics at postcode level or coarser which are often employed in large-scale epidemiological research cannot separate the individual health effects of pollutants which are generally highly correlated at these coarser scales. Additionally, outdoor measurements cannot capture the *total* personal exposure that results from the cumulative effects of an individual moving between different indoor and outdoor microenvironments. During daily life, peak exposure events often occur during commuting (Karanasiou, Viana, Querol, Moreno, & de Leeuw, 2014) while the indoor environment is a significant site for exposure in part because people spend as much as 90% of their time indoors (Klepeis et al., 2001). Indoor air is affected by outdoor pollutants penetrating building envelopes with additional indoor sinks, sources and emissions from building materials which cannot be detected by fixed outdoor monitoring networks. The lack of information on indoor environments at population scale is a significant factor in poorly quantified health risks. As a result, inaccurate personal exposure estimations to air pollution introduce both bias and error in health estimations ultimately preventing epidemiological research to move from general associations to the specific (Zeger et al., 2000).

Rapid advancements in novel sensing technologies of air pollution sensors now offer the potential to monitor detailed personal exposure during daily life at population scale, thanks to their significantly reduced cost, smaller size and fast-response. Instrument development is accelerating fast with a growing number of companies utilising combinations of such sensors (Cross et al., 2017) as well as auxiliary components to build different types of monitors (Morawska et al., 2018). As a special case, it is now estimated that there are currently over 30,000 sensors operating in China to monitor concentrations of air pollutants (Morawska et al., 2018). Several studies over the last 15 years have attempted to quantify personal exposure to air pollutants by employing portable sensors, but most of those studies have been restricted to small-scale surveys (Steinle, Reis, & Sabel, 2013). However, large-scale studies are necessary to assess the health effects of harmful pollutants because they are often seen in only small subgroups of the population due to varying individual susceptibility and exposure profiles. Novel sensing technologies are in fact the only method to expand the personal exposure coverage at the population level. Yet, concerns remain about the validation and quality control of those sensors (Castell et al., 2017) as few personal exposure studies have evaluated their performance in field deployment conditions (Rai et al., 2017). Typically, novel sensing platforms are exclusively evaluated in outdoor static co-locations with reference instruments and they only target small numbers of pollutants, most commonly ozone, nitrogen dioxide (Lin et al., 2015) and/or particulate matter (Holstius, Pillarisetti, Smith, & Seto, 2014)(Feinberg et al., 2018).

To address these shortcomings, a highly portable personal air pollution monitor (PAM) that measures a large number of chemical and physical parameters simultaneously has been developed. This paper aims to evaluate the performance of the PAM when capturing *total* personal exposure to air pollution in diverse environmental conditions. To do so, the PAM performance was assessed in well-characterised outdoor, indoor and commuting microenvironments across seasons and different geographical settings. The PAM has already been deployed to participants of two large cardio-pulmonary cohorts in China (Han et al., 2019) (AIRLESS- Theme 3 APHH project) (Shi et al., 2018) and the UK (COPE) (Moore et al., 2016), and in a number of smaller international pilot projects in Northern America, Europe, South and East Asia and Africa. This is the first of a series of publications that aim to capture *total* personal exposure to a large number of pollutants at unprecedented detail, and together with medical outcomes, to identify underlying mechanisms of specific air pollutants on health. As the field of novel air pollution sensing technologies expands rapidly, this paper further aims to ~~create a roadmap for calibration~~provide methodological guidance to researchers from diverse disciplines on how to comprehensively calibrate and ~~validation of~~validate portable monitors suitable for personal exposure quantification.

2 The personal air quality monitor

The PAM has been developed at the Department of Chemistry, University of Cambridge, in collaboration with Atmospheric Sensors Ltd. It is now commercially available (independently from the University of Cambridge) from Atmospheric Sensors Ltd (Model AS520, http://www.atmosphericsensors.com). The PAM (Figure 1) is an autonomous platform that incorporates multiple sensors of physical and chemical parameters (Table 1). The compact and lightweight design of the PAM (ca. 400g) makes the unit suitable for personal exposure assessment. The PAM is almost completely silent and can operate continuously. No other input is required by the user other than to place it for periodic charging (e.g. daily) and data upload in a base station. The measurements are also stored in an SD card inside the monitor and uploaded through general packet radio service (GPRS) to a secure access FTP server. Customised system software has been developed to optimise the performance of the platform. Depending on the chosen sampling interval of either 20 sec or 1 min, the battery life on a single charge lasts for 10 hours or 20 hours respectively. The combined cost of the sensors alone is less than £600 and the total cost of the PAM is less than £2,000 making it a "lower-cost" system (Cross et al., 2017).

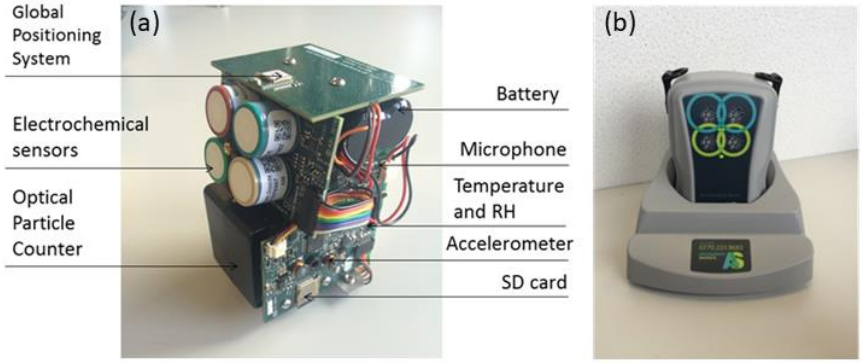


Figure 1. The personal air quality monitor. (a): design of the PAM platform internals and (b): in charging base-station. The external dimensions of the PAM are 13cm * 9 cm * 10 cm.

Table 1. Summary of monitored parameters of the PAM. PM_1, $PM_{2.5}$ and PM_{10} = the fraction of particles with an aerodynamic diameter smaller than 1 µm, 2.5 µm and 10 µm respectively; CO = carbon monoxide; NO = nitric oxide; NO_2 = nitrogen dioxide; O_3 = ozone.

Parameter	Method	Sampling Interval
Spatial coordinates	Global Positioning System (GPS)	20 sec
Background noise	Microphone	100 Hz
Physical activity	Tri-axial accelerometer	100 Hz
Temperature	Band-gap IC	4 sec
Relative Humidity (RH)	Capacitive	4 sec
PM_1, $PM_{2.5}$, PM_{10}	Optical Particle Counter (OPC)	20 sec
CO, NO, NO_2, O_3	Electrochemical sensors (EC)	100 Hz

User-friendly, bespoke software (Supplementary ~~material~~Material Figure A1) has been developed to automate the management and post-processing of the large volume of raw data collected with the PAM network. Data is held in a PostgreSQL relational database management system which has an unlimited row-storage capacity and allows the querying of large quantities of data

in a flexible manner while maintaining performance as the volume of data grows. Post-processing was performed in R software (R Development Core Team, 2008) (Supplementary ~~material~~Material Figure A1) following the methodology outlined in this paper.

2.1 Measurements of CO, NO, NO_2 and O_3

The principle of operation of all commercially available miniaturised gaseous sensors currently involves measuring changes in specific properties of a sensing material (e.g. electrical conductivity, capacitance, mass, optical absorption) when exposed to a gas species (Morawska et al., 2018). The PAM integrates small (20mm diameter) electrochemical (EC) sensors based on an amperometric principle of operation (Stetter & Li, 2008) for the quantification of carbon monoxide (CO), nitric oxide (NO), nitrogen dioxide (NO_2) and ozone (O_3). These EC sensors are the A4 variant from Alphasense (NO-A4 (Alphasense Ltd, 2016a), CO-A4 (Alphasense Ltd, 2017a), NO2-A43F (Alphasense Ltd, 2016b), Ox-A431 (Alphasense Ltd, 2017b)) and operate on a four electrode system~~.~~. The principle of operation of the four electrode system is identical to that of the earlier variants of three-electrode system (Alphasense, 2013) where the conventional setup of working electrode, counter electrode and reference electrode is supplemented with an additional electrode, the auxiliary (or non-sensing) electrode to compensate for the temperature dependence of the cell potential (Popoola, Stewart, Mead, & Jones, 2016). Earlier variants of EC sensors used in this paper have been extensively characterised in laboratory conditions and in static outdoor dense sensor networks (Mead et al., 2013). Those studies provided evidence that, after appropriate post-processing, the sensors had a linear response to the targeted pollutants and achieved excellent performance with limits of detection (LOD) < 4ppb demonstrating their suitability for atmospheric air quality measurements. The linearity and LOD of the 4-electrode sensors (when integrated in the PAM) have been tested under laboratory conditions following the same methodology as described in Mead et al. 2013, yielding very similar results.

Currently, standards for the calibration and performance evaluation of EC sensors focus on industrial applications (British Standards Institution, 2017). Following those standards, a widely adopted approach to calibrate EC sensors are gas chamber experiments to determine offset (baseline) and sensitivity (gain). To address the lack of standards for novel sensing technologies, a number of researchers and governmental organisations are developing protocols and guidelines to evaluate sensor/monitor performance in the laboratory and in the field, such as the European Metrology Research Programme of EURAMET (Spinelle, Aleixandre, Gerboles, & European Commission. Joint Research Centre. Institute for Environment and Sustainability., 2013), the European Standardisation Committee (CEN/TC 264/WG 42, 2018) and US-based groups (Long, Beaver, & Williams, 2014) (AQ-Spec, 2017).

Building on those protocols, the EC sensors were calibrated by co-location with certified reference instruments in similar environmental conditions and same geographical area where the monitors had been or were to be deployed. The considerable advantage of this approach over laboratory calibration includes the exposure of the sensor to the actual air pollution and temperature/relative humidity conditions under which it is expected to operate, as well as the assessment of any site-specific potential cross-interferences. A linear regression model (Equation 1) was applied to the co-location data to determine the calibration parameters used to convert raw sensor signals (mV) to mixing ratios (ppb). Temperature effects were corrected through the auxiliary electrode AE which might have a different sensitivity to the working electrode WE ($a \neq b$). The cross-sensitivities between the NO_2 and O_3 measurements were corrected via parameter c (the cross sensitive gas Y is NO_2 for O_3 measurements and vice versa). As the CO and NO sensors were found to be sufficiently selective, c was set to zero for the calibration of those sensors.

$$[\mathbf{X}]_{\mathbf{ref}} = \mathbf{a}\ \mathbf{WE_X} + \mathbf{b}\ \mathbf{AE_X} + \mathbf{c}\ \mathbf{WE_Y} + \mathbf{d} \qquad \textbf{Equation 1}$$

where

$[X]_{ref}$	reference measurement of pollutant X [ppb]
a	sensitivity of the working electrode [ppb / mV]
WE_X , AE_X	raw signal of the working and auxiliary electrode [mV]
b	sensitivity of the auxiliary electrode [ppb / mV] (accounts for temperature)
c	cross sensitivity with gas Y [ppb / mV], $c = 0$ for CO and NO
WE_Y	raw signal of the working electrode of the cross sensitive gas Y [mV]
d	intercept [ppb]

To evaluate the performance of the linear model, the datasets were split into training (i.e. calibration) and validation periods to first extract the calibration parameters and then apply them to the validation set and compare the measurements with those from reference instruments (referred to as "calibration - validation" method). ~~As relationships in these~~The training sets ranged from 1 to 16 days, and the adjusted coefficient of determination ($\overline{R}^2$) remained stable for training periods longer than 3 days. Therefore, approximately a third of the dataset was selected as a training set. As relationships in these linear models should ideally not be extrapolated beyond the range of the observations (including meteorological conditions), the calibration periods ~~were made sufficiently long to cover~~covered the temperature and concentration ranges in which the sensors were deployed (Cross et al., 2017). Once the performance of the model was established in diverse environments, we used the full co-location periods to determine the agreement between PAM sensors and reference instruments.

2.2 Particulate mass measurements

The operation of virtually all miniaturised particulate matter (PM) sensors that are currently commercially available is based on the light scattering principle, either volume scattering devices or optical particle counters (OPCs) (Morawska et al., 2018). The PAM integrates a commercially available miniaturised OPC (Alphasense OPC-N2) (Alphasense Ltd, 2018) which uses Mie scattering for real-time aerosol characterisation~~.~~ (Mie, 1908). Particles pass through a sampling volume illuminated by a light source (in this case a laser) and scatter light into a photo detector~~.~~ (Bohren & Huffman, 1983). The amplitudes of the detected scattering signals pulses are then related to particle size. The OPC counts these pulses and typically sorts them into different particle size bins (Walser, Sauer, Spanu, Gasteiger, & Weinzierl, 2017). The OPC-N2 classifies particles in 16 sizes (bins) in the range 0.38 -17 μm. The performance of this OPC in the laboratory (AQ-Spec, 2017) (Sousan, Koehler, Hallett, & Peters, 2016) showed a high degree of linearity. Similarly studies evaluating the OPC performance in outdoor static deployments (Di Antonio et al., 2018) (Crilley et al., 2018) showed that once site- and seasonally-specific calibrations were applied, the miniaturised sensor could be used to quantify number and mass concentrations of particles with a precision similar to other standard commercial reference optical PM instruments.

The complexity of evaluating PM sensor performance is much greater than that of gas sensors. Compared with standard instrumentation, optical PM instruments face four inherent limitations which introduce potential differences in mass estimations compared with reference gravimetric methods:

(a) Exposure of the particles to relative humidity (RH) results in hygroscopic growth of particles and leads to mass overestimation (Di Antonio et al., 2018).

(b) Small variations in the sensitivities of the photodetector and the intensity/ angle of the laser may result in a systematic error specific to each OPC sensor. Additionally, as particles enter the optical chamber, they may deposit on internal surfaces and optics of the sensor leading to a reduction in the measured scattered light, and thus instrument sensitivity.

(c) A further limitation of all optical methods is their inability to detect particles with a diameters below a certain size, typically 200-400 nm (Morawska & Salthammer, 2003).

(d) Finally, optical methods cannot distinguish the physical and chemical parameters of the aerosol (e.g. density, hygroscopicity, volatility) which might vary significantly as people move between different microenvironments with diverse emission sources further increasing the uncertainty of mass estimation.

(a) To compensate for these limitations, this work firstly corrected for the effect of RH by applying an algorithm based on the particle size distribution which was developed for aerosols in urban environments (Di Antonio et al., 2018) ~~(a.~~ (b). In the second step, a scaling factor for each OPC was determined to account for sensor-sensor variability ~~(b).~~ This scaling factor was determined from a linear fit between the RH-corrected mass and the reference measurements for each season independently and to compensate for instrument sensitivity that may change over time. (c) As the reference instruments (e.g. TEOM) include particles below the size range of the OPC in their mass estimations, the scaling factor addresses partly the under-prediction of mass due to undetected smaller particles ~~(c).~~ which may vary between seasons. The varying aerosol composition (d) remains a challenge, and therefore a constant density of 1.65 g/cm^3 was assumed. Although the OPC is able to measure PM_1, $PM_{2.5}$ and PM_{10}, this paper focusses on the performance of the $PM_{2.5}$ measurements because of the availability of reference instruments.

3 Performance of the PAM under well-characterised conditions in the field

In the following three sections the performance of the PAM is assessed when measuring air pollution concentrations in different environments that are relevant for the quantification of *total* personal exposure (outdoor, indoor and in movement) in the UK and China. Sensor performance may vary significantly with season (e.g. temperature and RH artefacts) while meteorological conditions may affect the variation of outdoor air pollution levels directly (e.g. stability of the atmosphere) and indirectly by socio-economic patterns (e.g. increased energy demand for heating). Similarly, indoor air may be directly affected by outdoor air pollution levels and indirectly through occupants' behavioural patterns (e.g. windows' adjustment to achieve thermal comfort). Taking into account the strong seasonal variation of air pollution levels, the performance of the PAM was evaluated by co-locating one or multiple PAMs with reference instruments both during the "heating" (when the majority of householders heat their home on a regular basis) and "non-heating" season. The residential central heating season in Beijing is from 15th November – 15th March (Beijing Municipal Government), while in the UK the equivalent heating season is 5.6 months (October – March/April) (BRE, 2013).

The description of the sites, principle of operation and models of certified reference instrumentation used can be found in Table 2. The co-locations in China involved 60 PAMs which had been previously deployed to 250 participants of a cardio-pulmonary cohort for one month during the heating and one month during the non-heating season (Han et al., 2019). The co-location in the UK involved 60 PAMs that have been previously deployed to 150 participants of a COPD cohort for two years continuously (Moore et al., 2016). The reproducibility between co-located sensors was very high even when the ambient concentrations were close to the LOD (mean $R^2 \geq 0.80$ for EC sensors and $R^2 \geq 0.91$ for the OPC, see Figure 2 and Supplementary ~~material~~Material Table T1). Hence, the performance of the selected PAMs in static deployments as described in this section is representative and can be extrapolated to the entire sensor network.

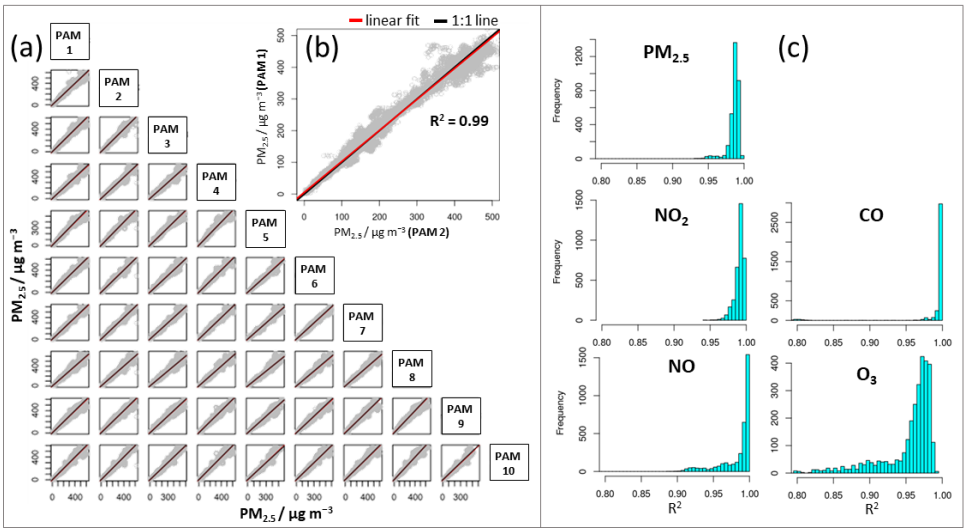


Figure 2. Reproducibility of a PAM network (in that case 60) co-located outdoors in Beijing during the heating season after one month of field deployment. (a): Scatterplot of the $PM_{2.5}$ measurements between 10 sensor pairs. The 1:1 line in black; and linear fit line in red. (b): "Close-up" of a scatterplot from (a) of one representative sensor pair. (c): histogram of the coefficient of determination (R^2) between all sensor pairs. R^2 values during this deployment were higher than 0.90 for all pollutants indicating the high reproducibility of the sensors' readings (see Table T1 for all co-locations). O_3 sensors R^2 > 0.80 due to very low ambient levels close to the LOD of the sensors.

Table 2. Details of the reference instruments used in this study. Time resolution of all measurements was 1 minute. * Due to malfunctioning of the TEOM in PKU during non-heating season, measurements from a TEOM in a nearby governmental site (Haidianwanliu, time resolution 1 hour) were used.

Deployment	Site description	NO, NO_2	CO	PM	O_3
Outdoor China	Urban background in Peking University (PKU) campus, Beijing	Chemiluminesence, Thermo-Science Model 42i	Nondispersive Infrared, Thermo-Science Model 48i	$PM_{2.5}$ * TEOM (Tapered Element Oscillating Microbalance)	UV absorption Thermo Model 49i
Outdoor UK	Urban background at the Department of Chemistry, Cambridge	Chemiluminesence, Thermo-Science Model 42i	Nondispersive Infrared, Thermo-Science Model 48i	aerosol spectrometer FIDAS PALAS 200S	UV absorption Thermo-science Model 49i
Indoor residential China	Indoor deployment in an urban high-rise Beijing flat	NO_2 cavity attenuated phase shift spectroscopy (CAPS) Teledyne API T500U	NA	aerosol spectrometer GRIMM 1.108	NA
Commuting environment UK	Monitoring ~~Vehicle~~vehicle equipped with commercial instruments driving in central London	NO_2 CAPS Teledyne API T500U	NA	Nephelometer (scattering) Met One ES642	UV absorption Teledyne API T400

3.1 Outdoor performance of sensors in diverse urban environments with varying pollution profiles and meteorological parameters

In total, four outdoor co-location deployments have been evaluated to comprehensively characterise the performance of the sensors (in the UK and China during the heating and non-heating season, see Table 3). The PAMs were placed in protective shelters close to the inlets of the certified air pollution monitoring stations. The sensor measurements were converted to physical units following the methodology described in subsections 2.1 and 2.2.

As an illustrative example, the outdoor co-location in Beijing, China (19 days, Dec 2016 to Jan 2017) is presented in Figure 3 to demonstrate the previously mentioned "calibration - validation" method (subsection 2.1). The time series of the pollutants measured by the PAM (blue) follow closely the reference instruments (red) in both the

calibration (Figure 3a) and validation (Figure 3b) periods. Similarly, the time series and scatterplots of the other three co-locations (UK in the heating season, China and UK in the non-heating season) can be found in the Supplementary ~~material~~Material (Figures A2, A3, A4).

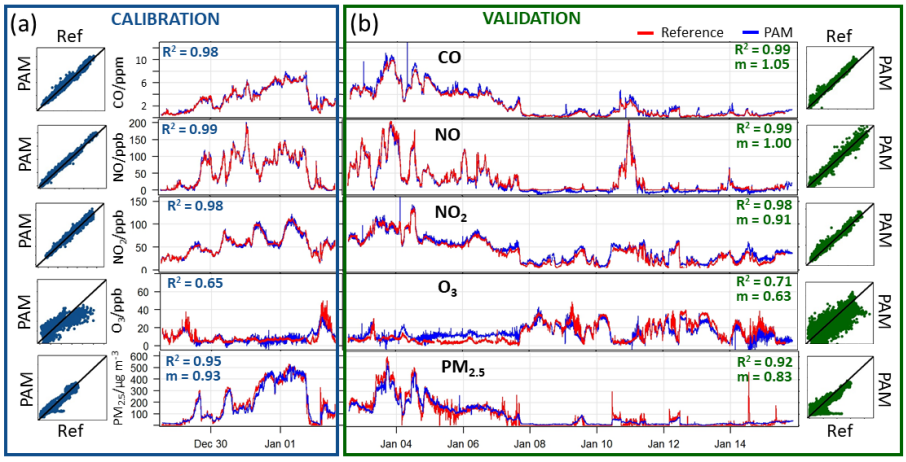


Figure 3. Outdoor co-location of one representative PAM with calibrated reference instruments in China (winter 2016/17) at 1-min time resolution demonstrating the "calibration-validation" methodology to evaluate the performance of the linear model. The first five days (a) were used to calibrate the EC sensors. The remaining co-location data (14 days, b) were used to validate the extracted calibration parameters. The scatterplots on each side show the correlations between reference and PAM measurements with the 1:1 line in black. $\overline{R}^2$ and gradients (m) are shown on each side in the corresponding colour.

Table 3 gives a quantitative overview of the agreement between the PAM measurements and the reference instruments in outdoor co-locations during the heating and non-heating season. Ambient temperature and RH (median, range: 5% -95%) as well as the maximum pollutant concentration measured are presented to describe the ambient conditions of each co-location. Because the PAM internal temperature is on average 7°C higher than the ambient temperature due to heat generated by the internal battery, the internal conditions the sensors were exposed to are also presented. The sensor performance against the reference instruments was evaluated using (1) $\overline{R}^2$ of the linear regression between PAM and reference and (2) the root mean square error (RMSE) using both the validation and calibration period (Table 3). $\overline{R}^2$ may be a misleading indicator of sensor performance when measurements are taken close to the LOD of the instruments. The RMSE can be a complementary parameter of $\overline{R}^2$ for the evaluation of performance, as it summarises the mean difference between measurements from the sensor and certified instruments The average values of $\overline{R}^2$ and RMSE of all N sensors during all co-locations are given in Table 3.

Table 3. Overview of sensors' performance during outdoor co-locations in China and the UK (7 to 19 days). Median values (range: 5th - 95th percentile) of the ambient temperature and relative humidity (RH), internal temperature and RH of the platform are presented. The 95th percentile of the concentration measurements of the reference over the entire co-location period is given as maximum concentration for each pollutant. The mean adjusted coefficients ($\overline{R}^2$) and root mean square errors (RMSE) indicate the agreement between the measurements of the sensors and reference instruments. The average values of all N sensors for each variable are given. Co-location in China in June shown in Italics as sensors were regularly exposed to temperatures higher than 40°C where sensors do not show linear temperature responses. The sensor reproducibility for these co-locations is presented in Table T1 in the Supplementary ~~material~~Material.

		Heating season		Non-heating season	
Location Start date – end date (Total hours of co-location deployment)		**China ~~(Dec- Jan)~~ 28/12/2016 – 15/01/2017 ~~(447h)~~**	**UK ~~(Oct- Nov)~~ 27/10/2017 – 13/11/2017 (408h)**	***China ~~(June~~ 28/06/2017 - 16/07/2017 ~~(432h)~~***	**UK ~~(April-May~~ 26/03/2018 – 10/04/2018) ~~(342h)~~**
Illustrative graphical example		Figure 3	Figure A2	*Figure A3*	Figure A4
Ambient conditions	Ambient Temp (°C)	1.1 (-3.6 – 6.1)	9.3 (4.3-14.4)	*29.9 (22.8-36.3)*	8.3 (4.7-18.1)
	Ambient RH (%)	40 (15-79)	81 (61-93)	*68 (43-96)*	83 (48-93)
Internal conditions of the PAM	Internal Temp (°C)	10.5 (5.3-18.0)	15.9 (11.0–20.8)	*40.2 (32.7–45.8)*	17.7 (12.2–26.8)
	Internal RH (%)	27 (14-44)	52 (39 -59)	*38 (23 – 55)*	52 (34 – 60)
Number of sensors (N)	[-]	N =59	N=3	*N=59*	N= 3
CO	Maximum (mean) mixing ratio (ppb)	6845 (2561)	357 (237)	*916 (575)*	276 (192)
	$\overline{R}^2$	0.98	0.74	*0.71*	0.67
	RMSE in ppb (percentage of max)	31 (0.5%)	31.6 (8.9%)	*212 (23%)*	33.3 (12.1%)
NO	Maximum (mean) mixing ratio (ppb)	132 (38)	19 (5)	*5 (1)*	6 (2)
	$\overline{R}^2$	0.94	0.89	*0.20*	0.58
	RMSE in ppb (percentage of max)	11.7 (8.9%)	3.0 (15.8%)	*13.0 (260%)*	2.2 (36.6%)
NO_2	Maximum (mean) mixing ratio (ppb)	98 (42)	35 (15)	*42 (22)*	19 (10)
	$\overline{R}^2$	0.84	0.90	*0.20*	0.84
	RMSE in ppb (percentage of max)	11.8 (12.0%)	3.0 (8.6%)	*13.3 (31.7%)*	2.6 (13.7%)
O_3	Maximum (mean) mixing ratio (ppb)	33 (13)	30 (16)	*109 (49)*	44 (28)
	$\overline{R}^2$	0.87	0.92	*0.80*	0.89
	RMSE in ppb (percentage of max)	3.6 (10.9%)	2.7 (9%)	*14.9 (13.7%)*	4.2 (9.5%)
$PM_{2.5}$	Maximum (mean) conc. ($\mu g\ m^{-3}$)	432 (114)	32 (12)	*110 (55)*	37 (3)
	$\overline{R}^2$	0.93	0.57 *	*0.65 **	0.80
	RMSE in $\mu g\ m^{-3}$ (percentage of max)	37 (8.6%)	9 (28%) *	*25 (22.7%)**	2 (5.4%)

*** due to unavailable data, PM mass measurements are not corrected for RH effects**
****comparison with governmental station ~ 3km away**

3.1.1 Outdoor performance of the PAM during the heating season co-locations

During the heating season outdoor co-locations of a number of PAMs next to certified reference instruments, ambient temperatures ranged from –4°Cto 6°C in China and between 4°C and 14°C in the UK. Air pollution in China was characterised by elevated levels of CO and $PM_{2.5}$ (Table 3) for extended time periods ("*haze*" events) partially driven by stagnant winds or a weak southerly wind circulation (Shi et al., 2018). Compared with pollutant levels in the UK, the concentrations of these CO and $PM_{2.5}$ were approximately ten times higher while the contrast in ambient NO_2 levels was less marked with levels in China only approximately three-fold higher.

The O_3, NO and NO_2 sensors exhibited an excellent performance ($\overline{R}^2 \geq 0.84$) in both geographical settings (Table 3). The median RMSE values were close to the LOD of the sensors (< 3 ppb) in the UK and slightly higher in China (< 12 ppb) (Figure 3, Table 3). In both deployments, the RMSE values of these gaseous sensors were negligible compared to the ambient concentration ranges of the targeted pollutants (less than 16% of the maximum mixing ratio recorded by the reference instruments).While the median $\overline{R}^2$ between the CO sensor and the corresponding reference was reasonably high in both outdoor deployments (≥ 0.74), the median RMSE values were also quite large (< 32 ppb). In fact, this is due to the known high intrinsic noise and LOD of the reference instrumentation (> 40 ppb) which is much higher compared to that of the electrochemical sensors (LOD < 4 ppb, see subsection 2.1).

Following the correction of the size segregated particle measurements for the effect of RH (subsection 2.2) the PM mass quantification with the miniaturised OPC agrees with the TEOM reference instrument with an adjusted $\overline{R}^2$ of 0.93. The low RMSE values (> 8.6% of the maximum concentration) demonstrate that the scaling factor addresses adequately the under-prediction of mass due to undetected smaller particles when derived from field calibration in the local environment. Due to unavailable measurements, the PM measurements in the UK could not be corrected for RH effects which resulted in only a moderate correlation with the reference instrument ($\overline{R}^2$ = 0.57, Figure A2).

3.1.2 Outdoor performance of the PAM during the non-heating season co-locations

One outdoor co-location in China (Figure A3) and one in the UK (Figure A4) were performed during the non-heating season, both over periods of two weeks (Table 3). In the UK, seasonal variation of ambient temperatures, RH and pollution levels was relatively small. In contrast, in China, seasonal variation was large with ambient temperatures reaching up to 36.3°C (median: 29.9°C) and generally lower pollution levels compared to the heating season. However, in both geographical settings, O_3 was significantly elevated. The performance of the O_3 sensor remained reliable in all deployments with median $\overline{R}^2$ = 0.80 and RMSE values <15 ppb, which might provide valuable insights on the health effects of this pollutant because (a) ozone is a strong oxidant with a high potential to affect the body and (b) has highest concentrations during the non-heating season compared to other pollutants which usually peak during the heating season.

Due to a malfunction of the PM reference (TEOM) instrument during the non-heating season at PKU, the PAM PM measurements had to be compared with a TEOM installed in a nearby governmental site ("*Haidianwanliu*").

Although not closely co-located (~ 3km), the gradient between the PAMs and reference measurements was close to unity (average m = 0.96, see example Figure A3) and there was still a notable correlation ($\overline{R}^2$ = 0.65) with a median RMSE of 25 μg m^{-3} indicating that away from direct sources, PM concentrations are essentially homogenous over relatively large urban areas. Compared with the heating season, PM concentrations in China were significantly lower, whereas PM levels in the UK varied little with season. After correcting for the effects of RH on PM, the PAM performance in the UK during the non-heating season significantly improved compared with the heating season (RMSE = 2 μg m^{-3} within the particle size range 0.38 – 17 μm).

While the performance of the O_3 and OPC sensor remained reliable across seasons and geographical settings, the performance of the CO, NO and NO_2 sensors decreased significantly ($\overline{R}^2 \geq 0.20$) during the hottest parts of the non-heating season in China due to extreme temperatures (internal median temperatures of the PAM: 40.2°C, 5-95%: 32.7°C– 45.8°C, Table 3). It should be noted that NO levels were close to the LOD of the sensor which also affects the $\overline{R}^2$ values. We conclude that the measurements of the CO, NO and NO_2 sensors should be interpreted with caution when the ~~PAM is~~sensors are exposed to temperatures above 40°C. However, during the field deployment to participants, the sensors were exposed to lower temperatures (see Figure A5) that did not impact on their performance (see subsection 3.2).

3.2 Indoor performance of the NO_2 and PM sensors

Low-cost air pollution sensors have generally been characterised outdoors next to reference instruments as described in the previous section. However, little is known about the performance of these sensors in indoor environments, where the population spends most of their time (Klepeis et al., 2001), where environmental conditions (e.g. temperature, RH) and emission sources may be significantly different compared with nearby outdoor environments.

To evaluate the indoor performance of the NO_2 and the OPC sensors, an experiment in an urban flat in central Beijing was performed during the non-heating season (May 2017). One PAM was deployed in the living area next to two commercial instruments that were used to provide reference measurements: (1) a cavity attenuated phase shift spectroscopy instrument (CAPS Teledyne T500U) for NO_2 and (2) a portable commercial spectrometer (GRIMM 1.108) for particulate matter measurements (Table 2). During the experiment the occupants relied on natural ventilation adjusting the windows freely to achieve thermal comfort. Median indoor temperatures were 26.0°C (5%-95% range: 17.1°C – 28.8°C), and the median internal PAM temperature was 33.0 °C (5%-95% range: 24.3°C-36.2°C), which is comparable with the temperature range during the non-heating season field deployment to participants (internal median temperature: 35.0°C, 5%-95% range: 28.5°C- 39.9°C, Figure A5).

The conversion of the raw measurements to ppb used the sensitivities extracted using outdoor co-locations both during the heating and non-heating season (subsection 3.1) with the linear model (subsection 2.1).

Commented [a2]: moved from graph caption for more clarity

The performance of the low-cost sensors in the indoor environment (Figure 4 and Figure A6) was comparable to the outdoor performance demonstrated in the previous section ($\overline{R}^2$ = 0.91, gradient m = 1.1, RMSE = 3 ppb for NO_2 (Figure 4c) and $\overline{R}^2$ = 0.86, gradient m = 0.86, RMSE = 7 μg m^{-3} for $PM_{2.5}$ (Figure 4d)) proving their suitability to quantify indoor air pollution levels for these species provided they have been adequately calibrated in the local environment.

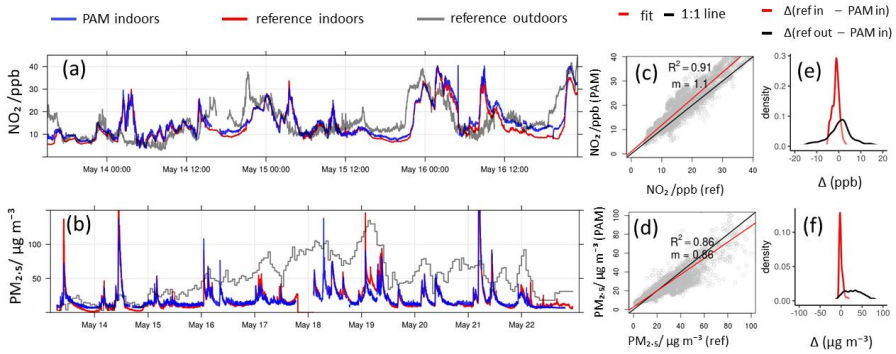


Figure 4. Indoor co-location of a PAM with portable commercial instrumentation (Table 2) in an urban flat in China during the non-heating season. (a): Time-series of NO_2 from the PAM (blue) and a cavity attenuated phase shift spectroscopy (CAPS) instrument (red). ~~The conversion of the raw measurements to ppb used the sensitivities extracted using outdoor co-locations both during the heating and non-heating season (subsection 3.1) with the linear model (subsection 2.1).~~ Outdoor NO_2 measurements (grey) were collected at PKU reference site (Table 2) which was located 5.3km away. Time resolution of measurements is 1-minute. (b): Time series of $PM_{2.5}$ mass measured with the PAM (blue) next to a commercial portable spectrometer (GRIMM 1.108, red). Mass concentrations were calculated from particle counts within the size range 0.38 -17 µm and same aerosol density for both instruments. Outdoor $PM_{2.5}$ mass measurements (grey) were collected at ~~a nearby~~the closest governmental station (Table 2, 1-hour time resolution) which was located 6km away. (c) and (d): Scatterplots show an excellent agreement between commercial instruments and miniaturised sensors making them suitable for the quantification of indoor pollution levels. The 1:1 line is in black and gradient m in red. (e) and (f): Density plots of the difference between measurements from the PAM and the indoor reference (red) are compared with the difference between the PAM and the outdoor reference (black).

Although this short experiment is only a "snap-shot" of indoor exposure, it shows that the measurement error of the PAM relative to established commercial instruments is negligible compared with the error in indoor exposure estimates introduced from using inadequate exposure metrics, in that case, outdoor measurements from the closest monitoring reference site. For example, using outdoor measurements from the closest monitoring station would have resulted in an over-prediction of indoor $PM_{2.5}$ concentrations (moderated by attenuation effects of the building envelope) with an average difference of 30 µg m^{-3} (standard deviation: 29 µg m^{-3}) which is significantly higher than the 7 µg m^{-3} RMSE value of the PAM (Figure 4f). While indoor NO_2 levels broadly followed outdoor levels, the range of the error in under-predicting and over-predicting exposure events is much broader (min-max range: -18 to 18 ppb; Figure 4e) compared with the error introduced from measurement uncertainties (-7 to 5 ppb). Such peak exposure events might be important triggers for acute health responses.

3.3 Performance of the PAM in non-static configurations

The aim of this section is to evaluate the PAM reproducibility and accuracy while in movement, with pedestrian and in-vehicle deployments.

3.3.1 Reproducibility of the PAM when not static

Multiple (in this case nine) PAMs were carried by a pedestrian while keeping an activity diary and walking between two indoor environments via a highly trafficked road in Cambridge, UK (weekday in January). Using NO measurements (the main traceable component from combustion engines) as an illustrative example, Figure 5a shows the simultaneous measurements of all PAMs as a time series and the scatterplots between the measurements of two of those PAMs separated into indoor (Figure 5b) and outdoor data (Figure 5c).

Significant changes of the pollution levels were observed when moving between the different environments illustrating the high granularity of personal exposure in daily life. Compared with the indoor environments, walking in traffic resulted in elevated pollution exposure events. As illustrated in the time series of Figure 5, the difference in pollution levels between the three micro-environments was significantly higher than the variability between PAM measurements.

Table 4 gives an overview of the correlations within the co-located moving network: In indoor environments an excellent agreement between all sensors (median $R^2 > 0.96$) was found, indicating a high sensor reproducibility. An exception was the O_3 sensor which showed poor between-sensor reproducibility due to very low indoor and outdoor concentrations (< 5 ppb) near the LOD of the sensor. The between-sensor correlations in the road environment were lower than indoors (median $R^2 > 0.85$) due to highly heterogeneous air pollution concentrations driven by complex factors (e.g. canyon air mixing, moving vehicle sources, topology). This signifies that in such environments air pollution concentrations might differ on such a short spatial and temporal scale that even sensors that are less than one meter apart from each other capture a slightly different exposure profile.

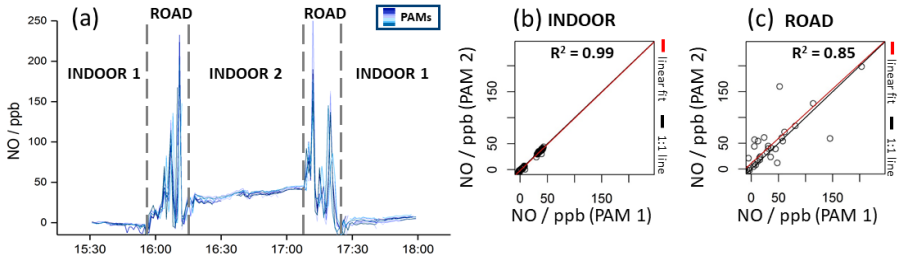


Figure 5. Short-term deployment of nine PAMs carried simultaneously by a pedestrian moving between two indoor environments (laboratory, café) in Cambridge, UK, in January 2018. (a): Time-series of NO measurements from the PAM sensors (blue lines). (b) and (c): Scatterplots between two of those PAMs, whereby indoor data was separated from outdoor data. The 1:1 line in black; and linear fit line in red.

Table 4. Correlations between PAM sensors: Adjusted $\overline{R}^2$ values of each sensor pair of the simultaneously carried PAMs were determined. Median R^2 value of all combinations are presented in the table below. Very low O_3 levels (< 5 ppb) resulted in poor between-sensor correlations.

	Indoor	Outdoor
	median $\overline{R}^2$	median $\overline{R}^2$
NO	0.99	0.87
NO_2	0.96	0.94
O_3	*0.16*	*0.46*
CO	0.99	0.95
$PM_{2.5}$	0.99	0.85

When moving rapidly between different environments with different temperatures (i.e. from outdoors to a warmer indoor microenvironment) false peaks were observed in the EC sensor measurements (Figure A7) (Alphasense Ltd, 2013). The response and recovery time following rapid temperature transitions was found to vary for different sensor types. To account for the false sensor responses, firstly an algorithm to identify those events was developed and then a 15-min window for CO and a 5-min window for NO, NO_2 and O_3 measurements was removed from the data (Figure A7 and Figure A8). Though it potentially excludes peak exposure events as rapid temperature changes often occur when people leave heated buildings and enter (colder) traffic environments to commute, this correction method removes typically less than 0.1% of the exposure data set under daily life conditions. The PM measurements are not affected by these temperature transitions.

3.2 Accuracy of the PAM when not static

A PAM was mounted on the roof of a battery-powered vehicle equipped with multiple commercial instruments (Table 2) mapping air pollution levels in London at speeds ~~of~~ up to 60 km/h for one day during the non-heating season (Figure 6). The PAM was mounted on the roof with the OPC inlet facing forwards and the EC sensors facing to the sides. The reference instrument inlets were located on the car roof as well. There was no correlation between car speed and RMSE error in the gaseous and particulate measurements. The OPC contains an airflow measurement unit which compensates for any wind or internal flow dependence.

Considering the high spatial variability of air pollution in traffic environments (see subsection 3.3.1), the accuracy of the PAM in a mobile configuration was high for all targeted pollutants ($\overline{R}^2 \geq 0.54$). To illustrate the large degree of variability of air pollution concentrations over time, the investigated area was mapped throughout the day multiple times with highest concentrations of $PM_{2.5}$ and NO_2 recorded during the morning rush hour.

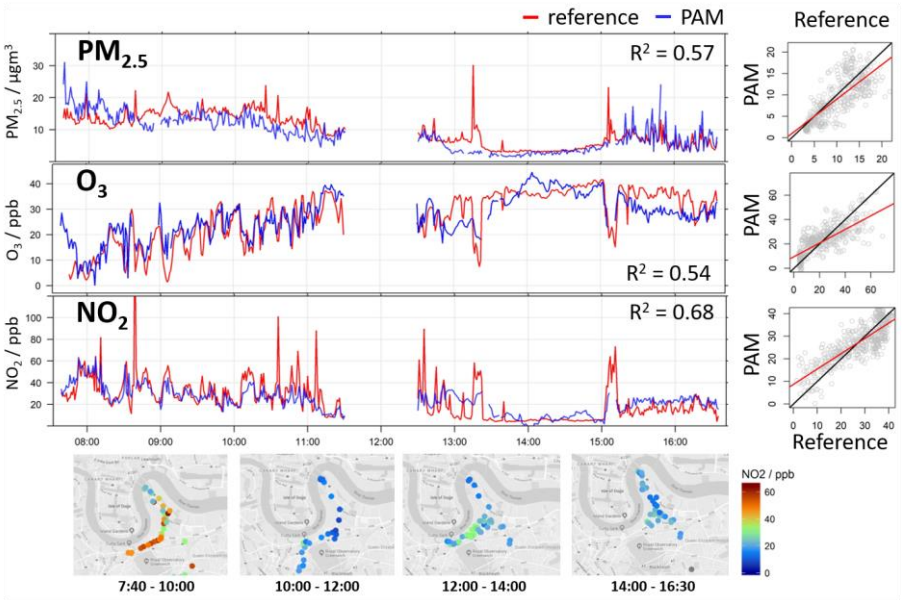


Figure 6. The vehicle deployment in London, UK: a PAM was attached to a car equipped with multiple commercial instruments (Table 2) for four days. (a): Time-series of one-day measurements of the PAM (blue) and commercial instruments (red). (b): Corresponding scatterplots between measurements from commercial instruments and the PAM in motion in an urban environment. The 1:1 line in black; and linear fit line in red. (c): Maps (map data 2019 Google) of the mobile deployment over 2-hour windows illustrating the large temporal variability of NO_2.

4 Discussion and conclusions

Mounting evidence points towards a causal link between exposure to air pollution and health outcomes. However, due to current limitations in cost, maintenance and availability of instrumentation, most large-scale health studies have focused on developed countries and have relied on low spatial and temporal resolution (generally outdoor) air quality data as metrics of exposure, severely limiting causal inferences in epidemiological research worldwide. Emerging low-cost sensing technologies can offer a potential paradigm shift in capturing personal exposure of the population during daily life in addressing this critical shortcoming.

In this paper we demonstrated that, with suitable calibration and post-processing, the performance of currently available low-cost air quality sensors, in this case incorporated into a highly portable personal monitor (the PAM) is comparable with the performance of reference instrumentation across a wide range of conditions:

- in diverse outdoor environments (urban background and traffic);
- across seasons (over a wide temperature and RH range);
- in two geographical settings with differing air pollution levels and meteorological profiles (UK and China);
- in indoor environments (residential, laboratory, café) with varying emission sources, and
- in static and in non-static deployments.

A critical important outcome of this study is that the performance of the sensors substantially exceeds that needed to quantify the differences between indoor and outdoor pollution levels, and thus to quantify exposure levels in a reliable manner.

There are certain performance caveats with the low-cost sensors used in this study, which once identified are likely to be addressed in future generations of sensors:

- The performances of the CO, NO and NO_2 sensors were found to degrade at temperatures above 40 ^{0}C. In fact, such extreme environmental conditions were not encountered during the actual personal exposure sample periods for which the PAMs were used, and the performance criteria discussed above were met.
- A limitation of all optical PM sensors, low-cost or reference, is that they cannot measure small particles below a critical size threshold (typically 200 – 400 nm). In this work we show that by appropriate local calibration, this shortcoming can be largely accounted for.

The toxicity of particles is also likely to depend on their chemical composition. Most national networks measure total mass only, and measuring particle chemical composition is currently largely the domain of the research community. A major challenge will be to develop techniques to allow routine PM composition measurements, both for the regulatory networks and for applications such as personal monitoring.

The key conclusion is that when suitably operated, highly portable air pollution personal monitors can deliver traceable high quality exposure metrics which can address scientific, health and policy questions for the indoor and outdoor environment in a way that has not been possible before. Mobile and static PAM networks have now been deployed in a range of health studies, and these will be the focus of future papers.

Acknowledgments

The authors would like to thank the following: Qiang Zhang and Kebin He (Tsinghua University) for PM outdoor data during the non-heating season, Paul Williams (University of Manchester) for providing the GRIMM instrument, Envirotechnology Services for reference mobile measurements. The authors would also like to thank MRC, Newton Fund, NERC and National Natural Science Foundation of China (NSFC Grant 81571130100) for funding for the AIRLESS project and MRC for funding of the COPE project.

Author contribution

L.C. and A.K. have contributed equally to this manuscript. Conceptualisation: L.C., A.K., B.B. and R.L.J.; **Sensor platform development:** L.C. and M.K.; **Sensor platform deployment:** L.C., A.K., Y.H., T.W., H.Z., Q.W. and S.F.; **Data curation:** L.C., A.K., O.A.M.P., A.D.A, M.K., Y.H., F.A.S., S.C. and B.B.; **Formal data analysis:** L.C., A.K., O.A.M.P., A.D.A. and R.L.J.; **Resources:** M.H., J.K.Q., B.B., F.J.K., T.Z. and R.L.J.; **Software:** L.C., A.K. and M.K.; **Visualisation:** L.C. and A.K; **Writing original draft:** L.C. and A.K.; **Review and editing:** O.A.M.P., Y.H. and R.L.J.

Competing interests. The authors declare that they have no conflict of interest.

Data availability. ~~All data can be provided by the corresponding author upon request.~~: https://doi.org/10.17863/CAM.41918

References

Alphasense. (2013). *Alphasense Application Note AAN 104 How Electrochemical Gas Sensors Work*. Retrieved from www.alphasense.com

Alphasense Ltd. (2013). *Alphasense Application Note AAN 110 ~~Emnvironmental~~Environmental changes: temperature, pressure, humidity*. Retrieved from www.alphasense.com

Alphasense Ltd. (2016a). *Technical specification. NO-A4 Nitric Oxide Sensor 4-Electrode*. Retrieved from http://www.alphasense.com/WEB1213/wp-content/uploads/2016/03/NO-A4.pdf

Alphasense Ltd. (2016b). *Technical specification. NO2-A43F Nitrogen Dioxide Sensor 4-Electrode*. Retrieved from http://www.alphasense.com/WEB1213/wp-content/uploads/2016/04/NO2-A43F.pdf

Alphasense Ltd. (2017a). *Technical specification. CO-A4 Carbon Monoxide Sensor 4-Electrode*. Retrieved from http://www.alphasense.com/WEB1213/wp-content/uploads/2017/01/COA4.pdf

Alphasense Ltd. (2017b). *Technical specification. OX-A431 Oxidising Gas Sensor Ozone and Nitrogen Dioxide 4-Electrode*. Retrieved from http://www.alphasense.com/WEB1213/wp-content/uploads/2017/03/OX-A431.pdf

Alphasense Ltd. (2018). Technical Specification. OPC-N2. Retrieved November 19, 2018, from http://www.alphasense.com/WEB1213/wp-content/uploads/2018/02/OPC-N2-1.pdf

AQ-Spec. (2017). Air Quality Sensor Performance Evaluation Center (AQ-SPEC): Field evaluation of low-cost air quality sensors. Retrieved from http://www.aqmd.gov/home

Bohren, C. F., & Huffman, D. R. (1983). *Absorption and scattering of light by small particles*. Wiley.

BRE. (2013). *Main heating systems (BRE report number 286733a)*. Retrieved from https://assets.publishing.service.gov.uk/

British Standards Institution. BS EN 50402:2017 Electrical apparatus for the detection and measurement of combustible or toxic gases or vapours or of oxygen. Requirements on the functional safety of gas detection systems (2017). Retrieved from https://standardsdevelopment.bsigroup.com/projects/2015-02652

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

CEN/TC 264/WG 42. (2018). *Ambient air - Air quality sensors.* Vienna. Retrieved from https://www.din.de/

COMEAP. (2018). *Committee on the Medical efeects of air pollutants. Long-term exposure to air pollution : effect on mortality*. Health Protection Agency. Retrieved from https://www.gov.uk/government/publications/comeap-long-term-exposure-to-air-pollution-effect-on-mortality

Crilley, L. R., Shaw, M., Pound, R., Kramer, L. J., Price, R., Young, S., … Pope, F. D. (2018). Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring. *Atmospheric Measurement Techniques*, *11*(2), 709–720. https://doi.org/10.5194/amt-11-709-2018

Cross, E. S., Williams, L. R., Lewis, D. K., Magoon, G. R., Onasch, T. B., Kaminsky, M. L., … Jayne, J. T. (2017). Use of electrochemical sensors for measurement of air pollution: correcting interference response and validating measurements. *Atmospheric Measurement Techniques*, *10*(9), 3575–3588. https://doi.org/10.5194/amt-10-3575-2017

Di Antonio, A., Popoola, O., Ouyang, B., Saffell, J., Jones, R., Di Antonio, A., … Jones, R. L. (2018). Developing

a Relative Humidity Correction for Low-Cost Sensors Measuring Ambient Particulate Matter. *Sensors*, *18*(9), 2790. https://doi.org/10.3390/s18092790

Feinberg, S., Williams, R., Hagler, G. S. W., Rickard, J., Brown, R., Garver, D., … Garvey, S. (2018). Long-term evaluation of air sensor technology under ambient conditions in Denver, Colorado. *Atmospheric Measurement Techniques*, *11*(8), 4605–4615. https://doi.org/10.5194/amt-11-4605-2018

Han, Y., Chen, W., Chatzidiakou, L., Yan, L., Wang, Y., Cai, S., … Zhu, T. (2019). Effects of AIR pollution on cardiopuLmonary disEaSe in urban and peri-urban reSidents in Beijing: protocol for the AIRLESS study. *BMC Public Health*.

Holstius, D. M., Pillarisetti, A., Smith, K. R., & Seto, E. (2014). ~~Atmospheric Measurement Techniques~~ Field calibrations of a low-cost aerosol sensor at a regulatory monitoring site in California. ~~*Atmos. Meas. Tech*~~*Atmospheric Measurement Techniques*, *7*~~;~~(4), 1121–1131. https://doi.org/10.5194/amt-7-1121-2014

Karanasiou, A., Viana, M., Querol, X., Moreno, T., & de Leeuw, F. (2014). Assessment of personal exposure to particulate air pollution during commuting in European cities—Recommendations and policy implications. *Science of The Total Environment*, *490*, 785–797. https://doi.org/10.1016/J.SCITOTENV.2014.05.036

Klepeis, N. E., Nelson, W. C., Ott, W. R., Robinson, J. P., Tsang, A. M., Switzer, P., … Engelmann, W. H. (2001). The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *Journal of Exposure Science & Environmental Epidemiology*, *11*(3), 231–252. https://doi.org/10.1038/sj.jea.7500165

Lin, C., Gillespie, J., Schuder, M. D., Duberstein, W., Beverland, I. J., & Heal, M. R. (2015). Evaluation and calibration of Aeroqual series 500 portable gas sensors for accurate measurement of ambient ozone and nitrogen dioxide. *Atmospheric Environment*, *100*, 111–116. https://doi.org/10.1016/J.ATMOSENV.2014.11.002

Long, R., Beaver, M., & Williams, R. (2014). *Procedures and Concepts of EPA's Ongoing Sensor Evaluation Efforts*. *AIR AND WASTE MANAGEMENT ASSOCIATION'S MAGAZINE FOR ENVIRONMENTAL MANAGERS*. Pittsburgh, PA. Retrieved from https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NERL&dirEntryId=290592

Mead, M. I., Popoola, O. A. M., Stewart, G. B., Landshoff, P., Calleja, M., Hayes, M., … Jones, R. L. (2013). The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmospheric Environment*, *70*, 186–203. https://doi.org/10.1016/J.ATMOSENV.2012.11.060

Mie, G. (1908). Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen. *Annalen Der Physik*, *330*(3), 377–445. https://doi.org/10.1002/ANDP.19083300302

Moore, E., Chatzidiakou, L., Jones, R. L., Smeeth, L., Beevers, S., Kelly, F. J., … Barratt, B. (2016). Linking e-health records, patient-reported symptoms and environmental exposure data to characterise and model COPD exacerbations: protocol for the COPE study. *BMJ Open*, *6*(7). https://doi.org/10.1136/bmjopen-2016-011330

Morawska, L., & Salthammer, T. (2003). *Indoor environment : airborne particles and settled dust*. Wiley-VCH.

Morawska, L., Thai, P. K., Liu, X., Asumadu-Sakyi, A., Ayoko, G., Bartonova, A., … Williams, R. (2018). Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone? *Environment International*, *116*, 286–299. https://doi.org/10.1016/J.ENVINT.2018.04.018

Newell, K., Kartsonaki, C., Lam, K. B. H., & Kurmi, O. P. (2017). Cardiorespiratory health effects of particulate

ambient air pollution exposure in low-income and middle-income countries: a systematic review and meta-analysis. *The Lancet. Planetary Health*, *1*(9), e368–e380. https://doi.org/10.1016/S2542-5196(17)30166-3

Popoola, O. A. M., Stewart, G. B., Mead, M. I., & Jones, R. L. (2016). Development of a baseline-temperature correction methodology for electrochemical sensors and its implications for long-term stability. *Atmospheric Environment*, *147*, 330–343. https://doi.org/10.1016/J.ATMOSENV.2016.10.024

R Development Core Team. (2008). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org.

Rai, A. C., Kumar, P., Pilla, F., Skouloudis, A. N., Di Sabatino, S., Ratti, C., … Rickerby, D. (2017). End-user perspective of low-cost sensors for outdoor air pollution monitoring. *Science of The Total Environment*, *607–608*, 691–705. https://doi.org/10.1016/j.scitotenv.2017.06.266

Shi, Z., Vu, T., Kotthaus, S., Grimmond, S., Harrison, R. M., Yue, S., … Zhang, Y. (2018). Introduction to Special Issue: In-depth study of air pollution sources and processes within Beijing and its surrounding region (APHH-Beijing). *Atmospheric Chemistry and Physics Discussions*, *19*, 1–62. https://doi.org/10.5194/acp-2018-922

Sousan, S., Koehler, K., Hallett, L., & Peters, T. M. (2016). Evaluation of the Alphasense optical particle counter (OPC-N2) and the Grimm portable aerosol spectrometer (PAS-1.108). *Aerosol Science and Technology*, *50*(12), 1352–1365. https://doi.org/10.1080/02786826.2016.1232859

Spinelle, L., Aleixandre, M., Gerboles, M., & European Commission. Joint Research Centre. Institute for Environment and Sustainability. (2013). *Protocol of evaluation and calibration of low-cost gas sensors for the monitoring of air pollution.* Publications Office. Retrieved from https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/protocol-evaluation-and-calibration-low-cost-gas-sensors-monitoring-air-pollution

Steinle, S., Reis, S., & Sabel, C. E. (2013). Quantifying human exposure to air pollution—Moving from static monitoring to spatio-temporally resolved personal exposure assessment. *Science of The Total Environment*, *443*, 184–193. https://doi.org/10.1016/J.SCITOTENV.2012.10.098

Stetter, J. R., & Li, J. (2008). Amperometric Gas SensorsA Review. https://doi.org/10.1021/CR0681039

Thurston, G. D., Kipen, H., Annesi-Maesano, I., Balmes, J., Brook, R. D., Cromar, K., … Brunekreef, B. (2017). A joint ERS/ATS policy statement: what constitutes an adverse health effect of air pollution? An analytical framework. *European Respiratory Journal*, *49*(1), 1600419. https://doi.org/10.1183/13993003.00419-2016

Tonne, C. (2017). A call for epidemiology where the air pollution is. *Lancet*, *1*(9). https://doi.org/10.1016/S2542-5196(17)30163-8

Walser, A., Sauer, D., Spanu, A., Gasteiger, J., & Weinzierl, B. (2017). On the parametrization of optical particle counter response including instrument-induced broadening of size spectra and a self-consistent evaluation of calibration measurements. *Atmos. Meas. Tech*, *10*, 4341–4361. https://doi.org/10.5194/amt-10-4341-2017

Zeger, S. L., Thomas, D., Dominici, F., Samet, J. M., Schwartz, J., Dockery, D., & Cohen, A. (2000). Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environmental Health Perspectives*, *108*(5), 419–426. https://doi.org/10.1289/ehp.00108419

Formatted: Font: Not Italic

Characterising low-cost sensors in highly portable platforms to quantify personal exposure in diverse environments.

Lia Chatzidiakou[1], Anika Krause[1], Olalekan A.M. Popoola[1], Andrea Di Antonio[1], Mike Kellaway[2], Yiqun Han[3,4,5], Freya A. Squires[6], Teng Wang[4,7], Hanbin Zhang[3,5,8], Qi Wang[4,7], Yunfei Fan[4,7], Shiyi Chen[4], Min Hu[4,7], Jennifer K. Quint[9], Benjamin Barratt[3,5,8], Frank J. Kelly[3,5,8], Tong Zhu[4,7], Roderic L. Jones[1]

[1]Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW, UK
[2]Atmospheric Sensors Ltd, Bedfordshire, SG19 3SH, UK
[3] MRC-PHE Centre for Environment & Health, Imperial College London and King's College London, London, W2 1PG, UK
[4] College of Environmental Sciences and Engineering, Peking University, Beijing, 100871, China
[5] Department of Analytical, Environmental and Forensic Sciences, King's College London, London, SE1 9NH, UK
[6] Department of Chemistry, University of York, York, YO10 5DD, UK
[7]The Beijing Innovation Center for Engineering Science and Advanced Technology, Peking University, Beijing, 100871, China.
[8] NIHR Health Protection Research Unit in Health Impact of Environmental Hazards, King's College London, London, SE1 9NH, UK
[9] National Heart and Lung Institute, Imperial College London, SW3 6LR, UK

Correspondence to: Dr Lia Chatzidiakou, email: ec571@cam.ac.uk, Tel: 01223 336345

Supplementary material

A1. Data processing with a bespoke automated R package

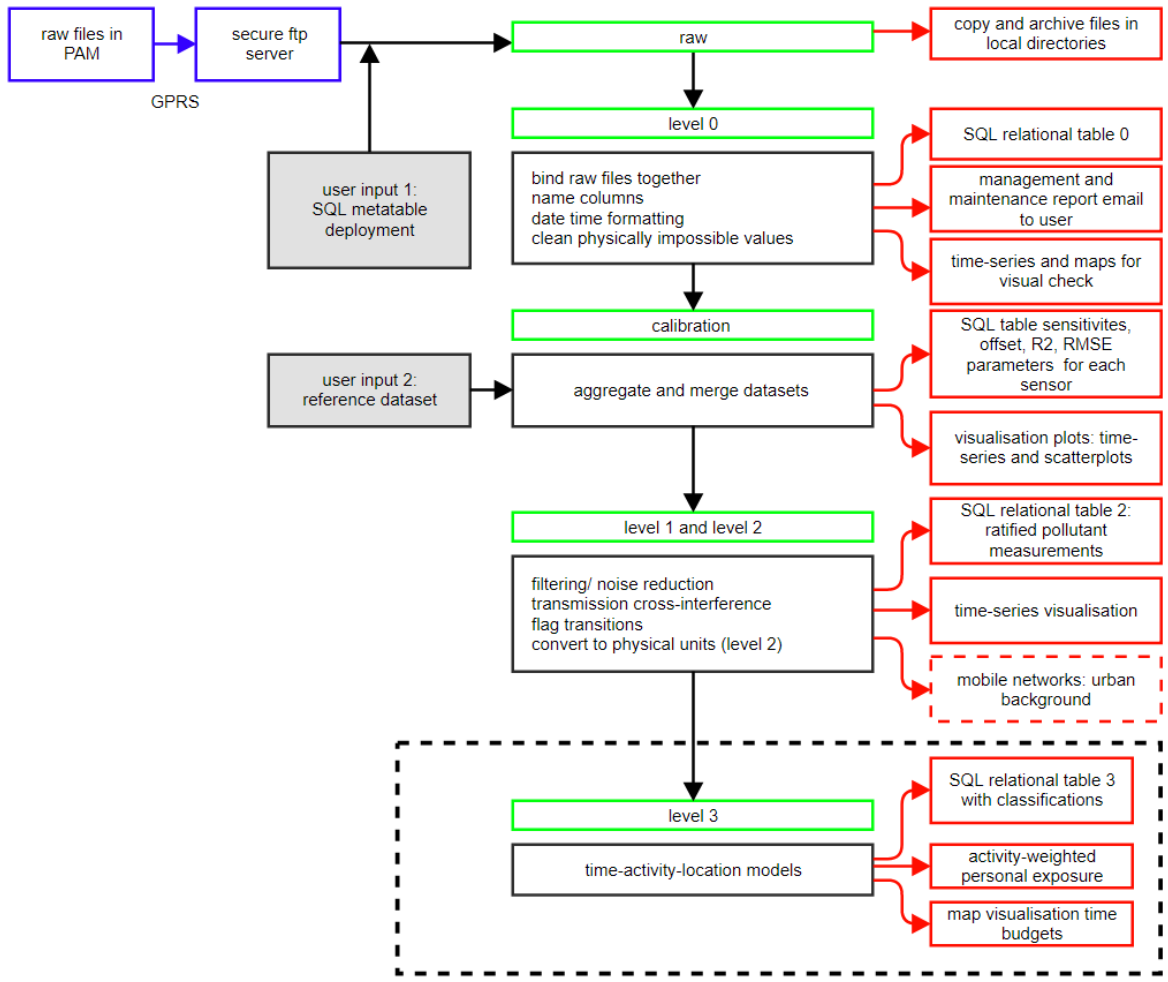


Figure A1. Data architecture of the bespoke software for the transparent and reproducible management and post-processing of data transmitted from the PAM network. Data is held in a relational database management system (RDBMS) in PostgreSQL to provide complex query capabilities and to create relationships between different components of the data itself. Data querying and post-processing is performed in R software and was selected as an established, open-source system where several independent but interoperable packages were used to address a wide range of statistical analysis and visualisation techniques. The user is required to input the metadata (grey) in the relational database containing information on deployment periods, reference instrumentation, hardware settings, maintenance schedule etc. The code can be executed with automated scheduled tasks (green) to output ratified databases (red). Level 0 produces time-series graphs and maps to facilitate fast visual inspection of the data and sends reports by email to improve efficiency of the maintenance of the PAM network. Levels 1 and 2 focus on the conversion of raw measurements to physical units and the quality assurance/ quality control of the personal exposure measurements. Dashed lines indicate work that will be presented in upcoming publications: 1) the creation of automated algorithms to classify time-activity-location profiles of individuals to estimate activity-weighted exposure; and, 2) the disaggregation of outdoor and indoor sources to overall personal exposure using information extracted from the wireless sensor network.

A2. Sensor reproducibility

To evaluate the reproducibility of the sensors, the coefficient of determination between the raw PM measurements and working electrode readings of all EC sensor pairs during the four co-locations in the UK and China in both seasons are presented in Table T1. All EC sensors show excellent correlations across the network with mean R^2 > 0.80 with the exception of the NO_2 sensors during the non-heating season in China, possibly affected by the high temperatures that the sensors were exposed to (see subsection). The coefficient of determination between the OPC-N2 sensors was high across seasons and settings (R^2 > 0.91); however, the gradient was lowest in the UK co-location during the heating season.

Table T1. Summary of reproducibility between sensors in outdoor co-locations in China and the UK that ranged from 7 to 19 days during the heating and non-heating season. Linear fitting equations between sensors generated the coefficient of determination (R^2) and the gradient [m] matrices. The average values of all N sensors for each variable are given, the standard deviations (σ) indicate the variation within the sensor network. The corresponding summary of the linear models with the reference instruments is presented in Table 3.

		Heating season		**Non-heating season**	
		China (Dec- Jan)	UK (Oct- Nov)	*China (June)*	UK (April)
	Internal PAM temperature (C)	10.5 (5.3-18.0)	15.9 (11.0–20.8)	*40.2 (32.7–45.8)*	17.7 (12.2–26.8)
	Internal RH (%)	27 (14-44)	52 (39 -59)	*38 (23 – 55)*	52 (34 – 60)
N sensors		N =59	N=3	*N=59*	N= 3
CO	Maximum mixing ratio (ppb)	6845	357	*916*	276
	R^2 (σ)	1.00 (0.01)	0.98 (0.00)	*0.94 (0.03)*	0.94 (0.00)
NO	Maximum mixing ratio (ppb)	132	19	*5*	6
	R^2 (σ)	0.89 (0.17)	0.80 (0.01)	*0.81 (0.07)*	0.94 (0.005)
NO_2	Maximum mixing ratio (ppb)	98	35	*42*	19
	R^2 (σ)	0.97 (0.04)	0.98 (0.00)	*0.60 (0.12)*	0.89 (0.03)
O_3	Maximum mixing ratio (ppb)	33	30	*109*	44
	R^2 (σ)	0.91 (0.05)	0.80 (0.00)	*0.90 (0.03)*	0.96 (0.01)
$PM_{2.5}$	Maximum conc. (μg m^{-3})	432	32	*110*	37
	R^2 (σ)	0.95 (0.17)	0.91 (0.05)	*0.96 (0.01)*	0.98 (0.01)

A3. Outdoor co-locations

A.3.1 Outdoor co-locations during the heating season

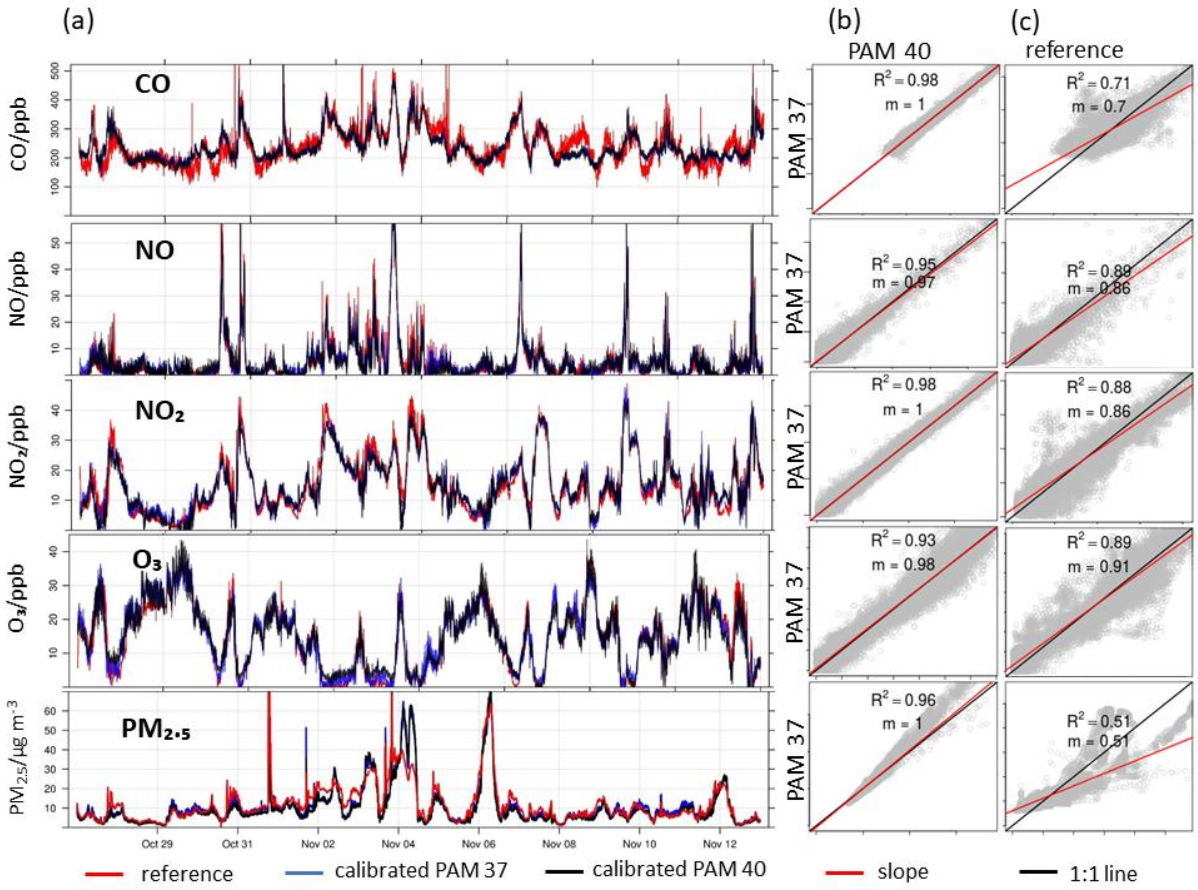


Figure A2. Outdoor co-location of PAMs with reference instruments in Cambridge for three weeks during the heating season. The PAMs had been previously deployed for two years to participants of a London cohort with chronic pulmonary disease. (a): Time series of two PAMs (blue, black) and the reference instruments (red). Time resolution is 20 sec averaged over 1 minute for comparison with the reference instruments. Due to unavailable data, the PM measurements of the PAM are presented without the RH correction resulting in an over-prediction of PM mass during certain periods. (b): correlation between PAMs and (c) correlation between PAM and reference. The 1:1 line in black; and gradients m in red.

A.3.2 Outdoor co-locations during the non-heating season

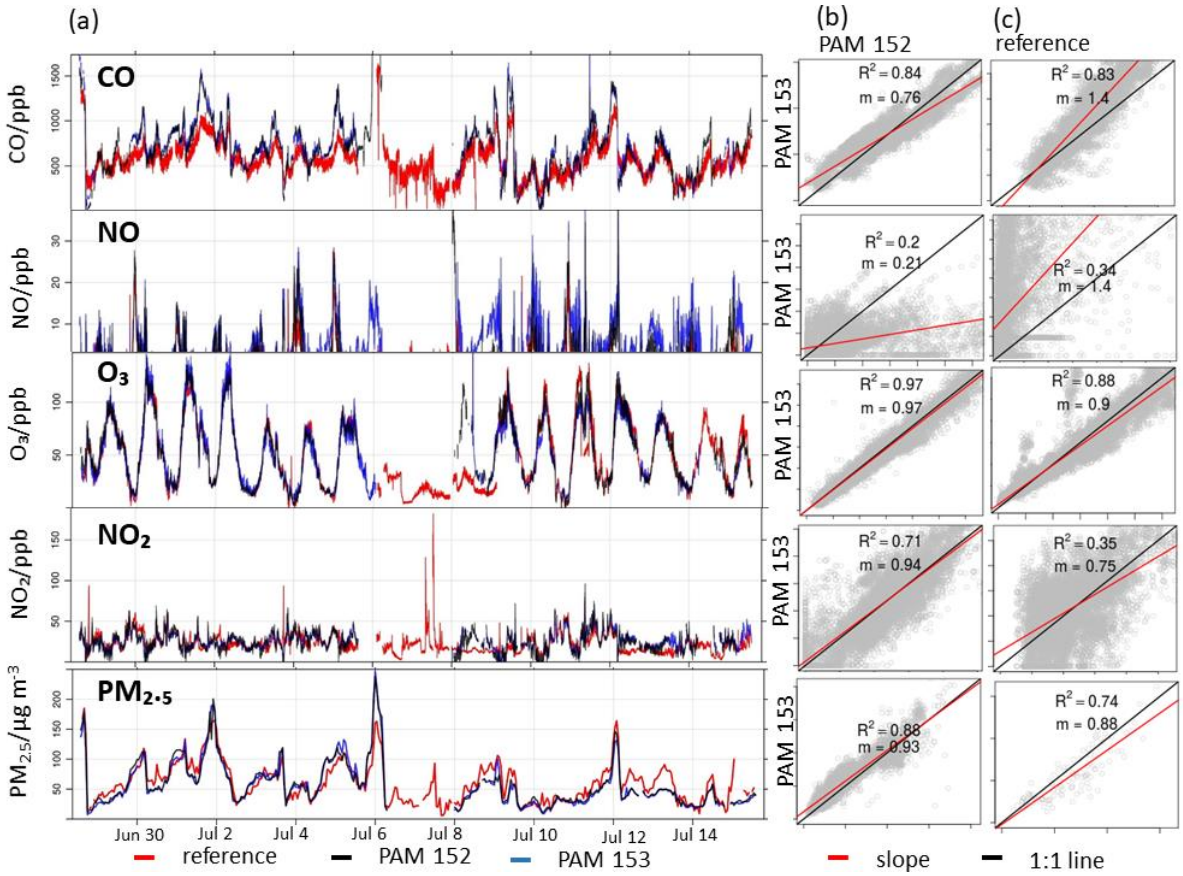


Figure A3. Outdoor co-location of PAMs with reference instruments in Beijing China during the non-heating season (June). The PAMs have been previously deployed in China for two months to participants of a cardio-pulmonary cohort. Due to malfunctioning of the TEOM in PKU during the non-heating season, measurements from a TEOM in a nearby governmental site (Haidianwanliu, time resolution 1 hour) were used for comparison with the PAM. (a): Time series of two PAMs (blue, black) and the reference instruments (red) in 1-minute time resolution. (b): correlation between PAMs and (c): correlation between PAM and reference. The 1:1 line (x=y) in black; and gradients m in red.

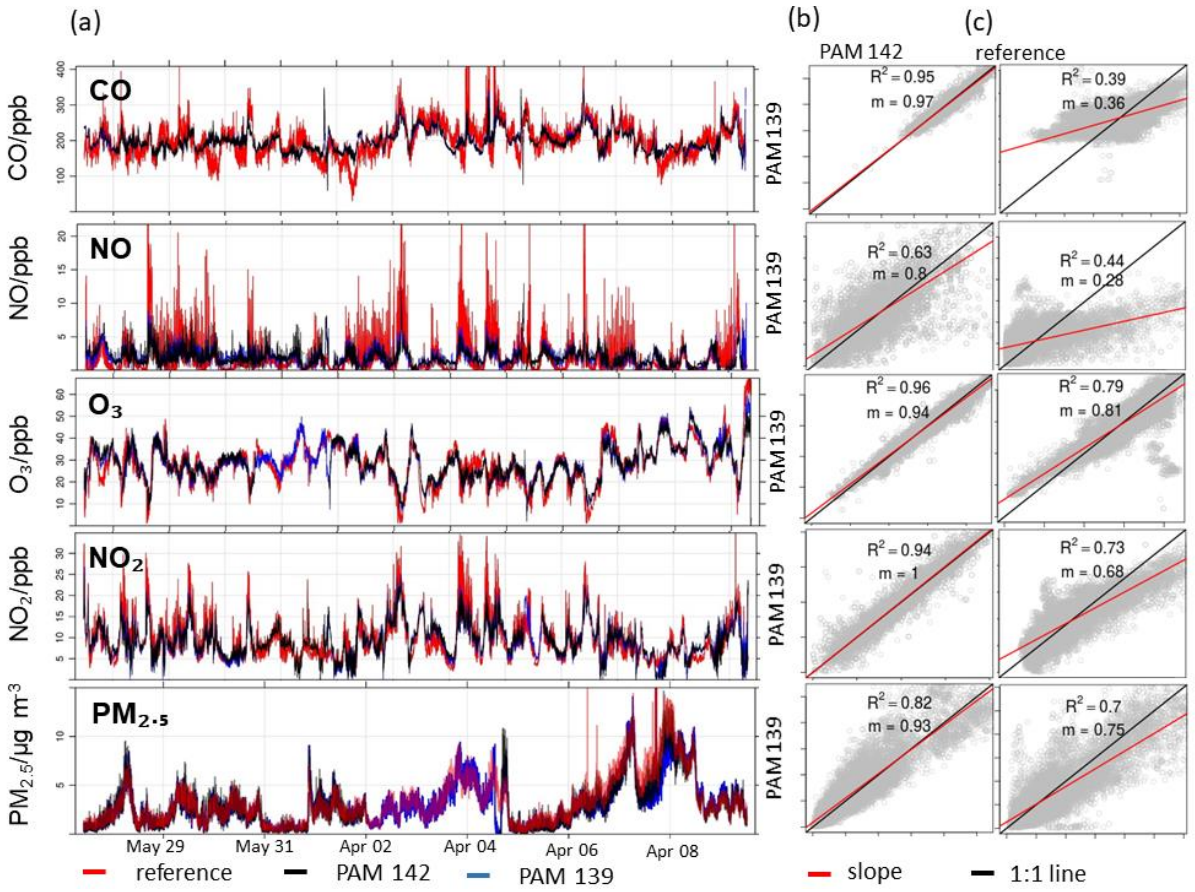


Figure A4. Outdoor co-location of PAMs with reference instruments in the UK during the non-heating season (April-May). (a): Time series of two PAMs (blue, black) and the reference instruments (red) in 1-minute time resolution. (b): correlation between PAMs and (c): correlation between PAM and reference. The 1:1 line (x=y) in black and gradients m in red.

A4. Deployment temperatures

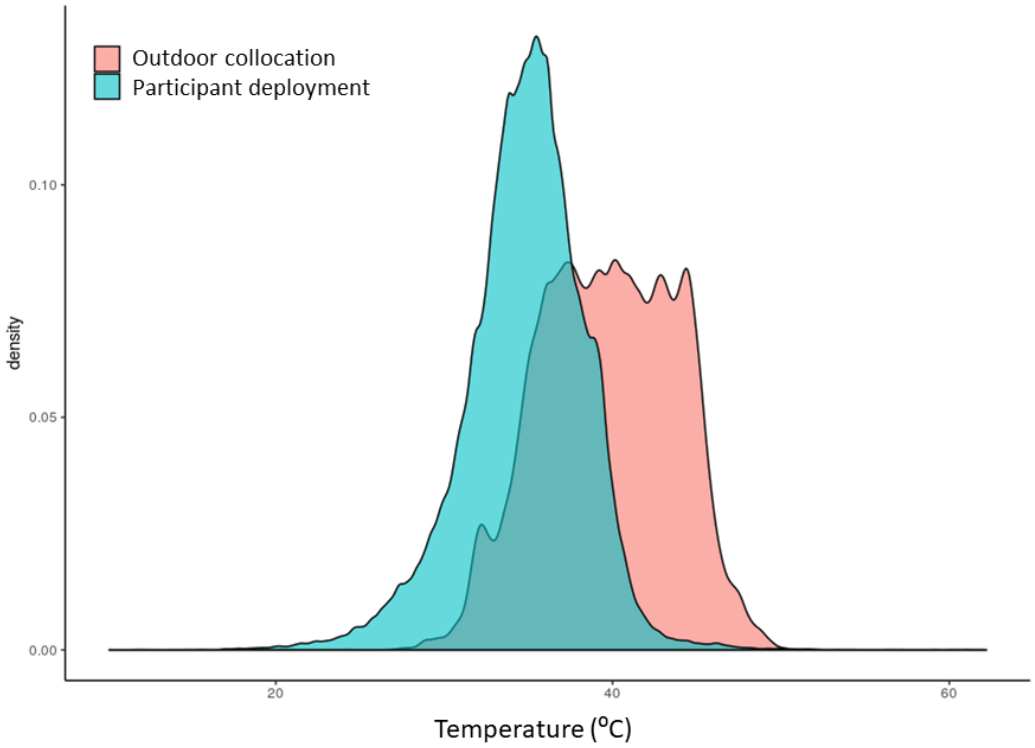


Figure A5. Histogram of internal temperatures of 60 PAMs recorded during the co-location period with reference instruments (pink) and deployment to participants (cyan) during the non-heating season in China. The co-location with the reference instruments on the roof of Peking university took place from 28/06/2017 to 16/07/2017, the field deployment was conducted in Beijing and Pinggu from 22/05/2017 to 26/06/2017. Internal temperatures of the PAMs are on average 7 °C higher than ambient temperatures.

A5. Indoor co-location

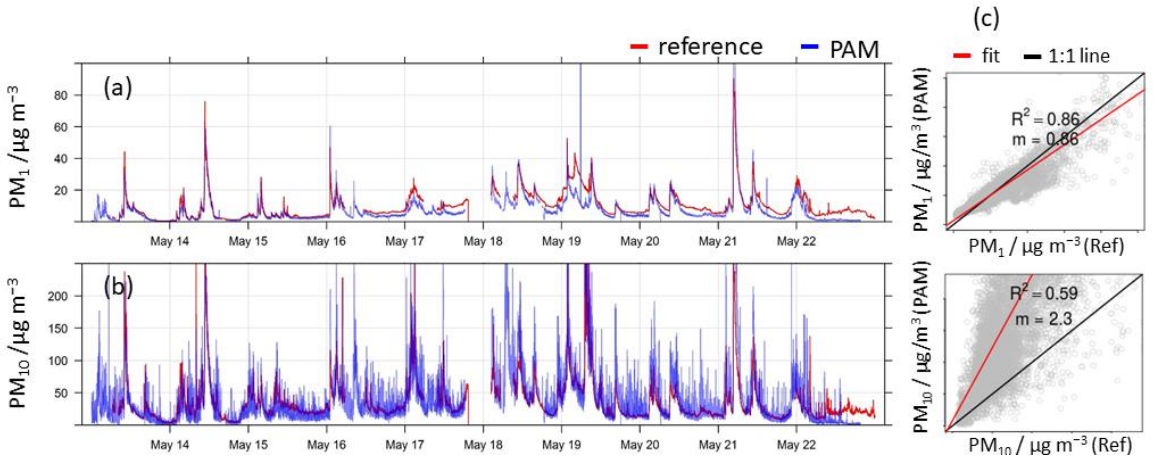


Figure A6. Indoor co-location of a PAM with portable commercial instrumentation (Table 2) in an urban flat in China during the non-heating season. Time resolution of the measurements is 1-minute. Time series of PM_1 (a) and PM_{10} (b) mass measured with the PAM (blue) next to a commercial portable spectrometer (GRIMM 1.108, red). Mass concentrations for PM were calculated from particle counts within the size range 0.38 -17 µm and same aerosol density for both instruments. (c): Corresponding scatterplots show an excellent agreement between commercial instrument and the PAM making them suitable for the quantification of indoor pollution levels. The 1:1 line is in black and gradient (m) in red.

A6 Effects of fast environmental changes on EC sensor performance ~~(shall we skip this section)~~

To investigate the effect of fast environmental changes on sensor performance, a controlled experiment was set up at a residential house in an urban background area in Cambridge. One PAM was deployed outdoor and a second PAM was deployed indoors. A third PAM was moved rapidly between an indoor and an outdoor location while a detailed time record ("diary") of the transition movements was kept (Figure A7, top). Commonly encountered residential indoor emission sources such as cooking (e.g. 11:30) and cigarette smoking (e.g. 17:30, PM measurements) were then introduced. The temperature difference between the indoor and the outdoor microenvironment was about 15°C (from 25°C indoors to 10°C outdoors) resulting in an RH difference of about 20% (30% indoors vs 50% outdoors). While the temperature sensor inside the PAM showed a relatively slow response when moved between indoors and outdoors, a rapid response of the RH sensor was evident (Figure A7, derivative of RH). The environmental changes were identified based on the time-derivative of the RH and were clearly distinguished from emission events (cooking, smoking). PM measurements were not affected by the rapid temperature changes; however, short-term transient changes were noticed in the measurements of the EC sensors. Figure A8 demonstrates that the recovery time of the NO, NO_2 and O_3 sensors was shorter (< 5 min) than that of the CO sensor (15 min).

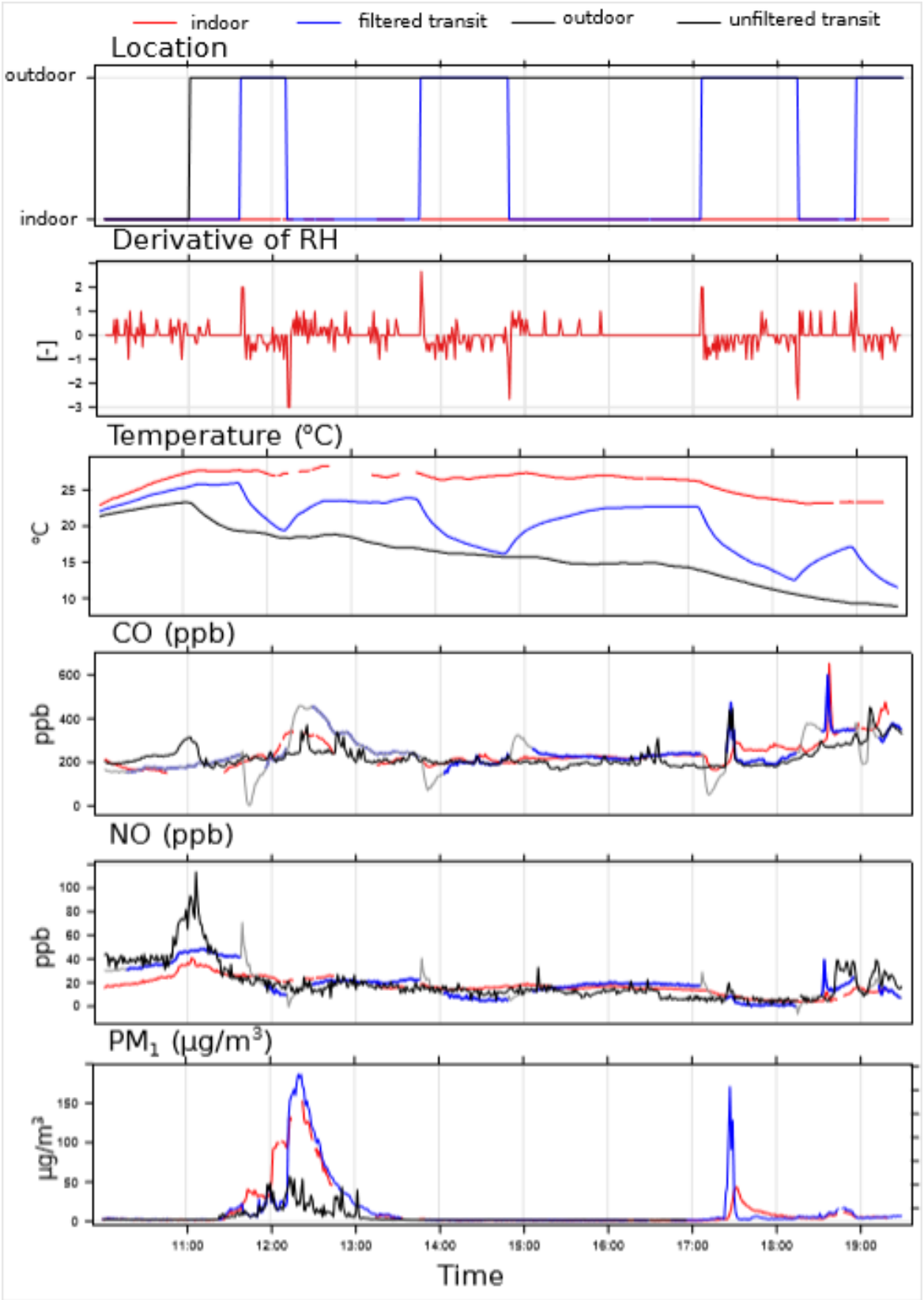


Figure A7. An experiment to characterise sensor responses to rapid environmental changes between different microenvironments. Time-series of one PAM deployed outdoors (black), a second PAM was deployed indoors (red) and a third PAM (grey, original signal) was moved rapidly between the indoor and the outdoor environment in the premises of a residence in an urban background area in Cambridge, UK. The rapid temperature changes resulted in sharp responses of the EC sensors whereas the PM measurements stay unaffected. A selective algorithm based on the on the time-derivative of the RH (dRH) was applied on the data of the moving PAM to flag and remove the false responses (blue line, filtered data) without filtering out short-term pollution events.

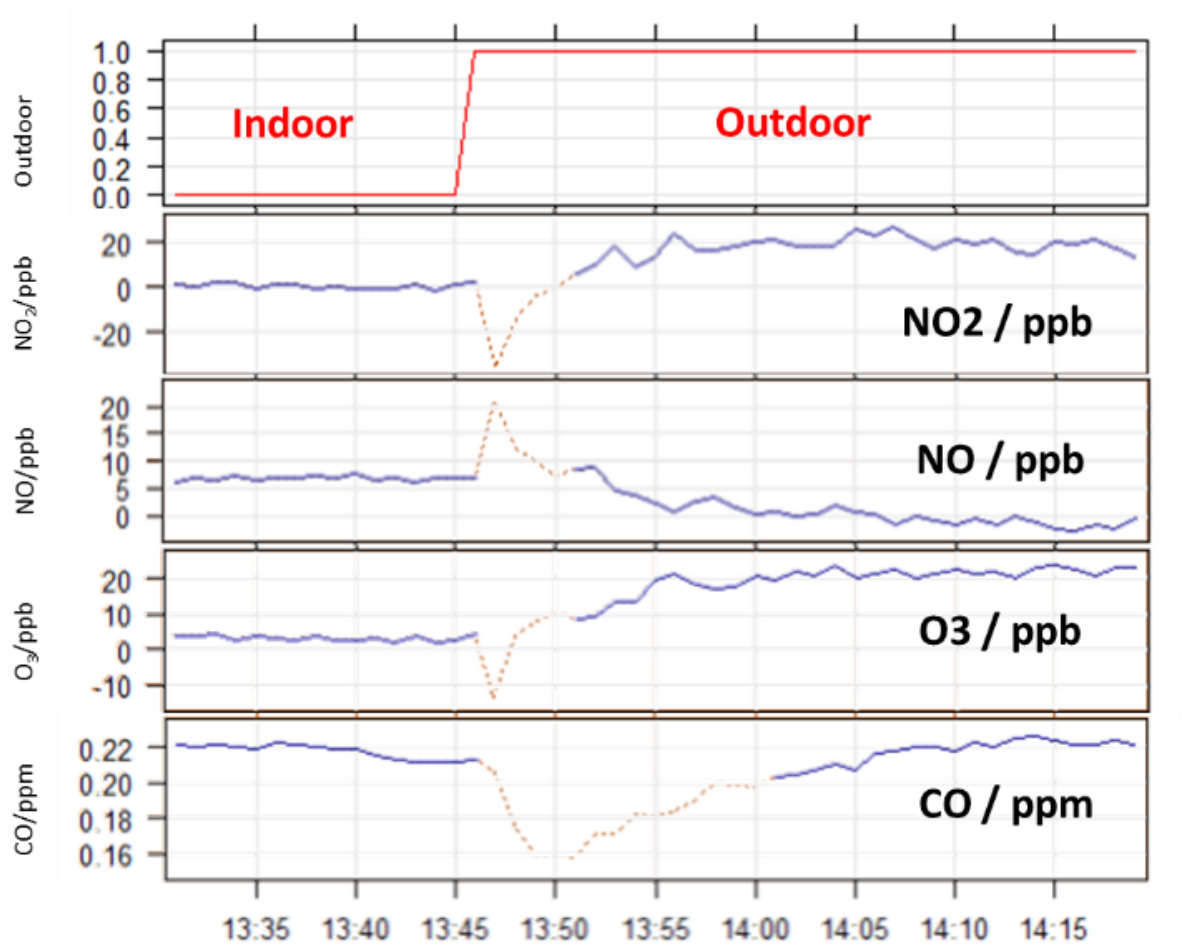


Figure A8. Close-up of one indoor-outdoor transition. The dotted orange line shows the original signal before removing the false temperature response. The blue line shows the remaining measurement after the removal. The stabilisation time after the transition depends on the sensor type with higher times for CO (15 min) and faster recovery times for the rest of the sensors (5 min).