Science of the Total Environment 980 (2025) 179476

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

How electrochemical sensors measure up to reference-grade nitrogen dioxide monitors across temporal scales

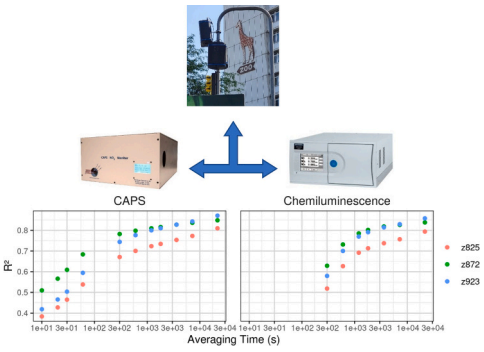
Seán Schmitz *, Alexandre Caseiro, Erika von Schneidemesser

Research Institute for Sustainability (RIFS) at GFZ, Germany

HIGHLIGHTS

- Conducted a 6-month study in Berlin using LCS, CAPS, and chemiluminescence monitors.
- Evaluated sensor performance at varying temporal resolutions (10 s to 6 h).
- LCS perform best at coarse resolutions but struggle at fine resolutions.
- CAPS-calibrated sensors outperform chemiluminescence-calibrated ones.
- Performance deteriorates during daytime and in high NO, NO_2, and O_3 conditions.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Pavlos Kassomenos

ABSTRACT

Air pollution remains a critical global health challenge, with nitrogen dioxide (NO_2) playing a significant role in adverse health outcomes. Low-cost sensors (LCS) offer promising opportunities for accessible and high-resolution air quality monitoring but face scrutiny over their accuracy and reliability. This study evaluates the performance of electrochemical LCS for NO_2 measurement in comparison to high-precision reference instruments—cavity attenuated phase shift (CAPS) and chemiluminescence NO_2 monitors—at eleven temporal resolutions (between 10-s and 6-h). Using three EarthSense Zephyrs containing electrochemical sensors, data were collected over six months at an urban-traffic air quality monitoring site in Berlin. Sensor performance was assessed based on statistical metrics, including R^2, relative error, and mean bias error (MBE). Results revealed that LCS exhibit good agreement with reference instruments at coarse time resolutions (≥1-h averages, $R^2 > 0.8$), but accuracy diminishes significantly at higher resolutions (<1-min, $R^2 < 0.5$). Overall, LCS perform better when trained against CAPS monitors than against chemiluminescence monitors. This performance is largely influenced by chemistry and emissions, with poorer performance during the daytime than at night, a pattern which is exacerbated at high time resolutions. CAPS-calibrated predictive models outperform those calibrated against chemiluminescence monitors in capturing short-term concentration peaks. These findings suggest that while LCS are suitable for coarse-resolution measurements of NO_2, their limitations in high temporal resolution dynamic environments pose significant challenges for their use in exposure studies and mobile measurements. Recommendations for

* Corresponding author.
E-mail address: sean.schmitz@rifs-potsdam.de (S. Schmitz).

https://doi.org/10.1016/j.scitotenv.2025.179476
Received 2 January 2025; Received in revised form 13 March 2025; Accepted 16 April 2025
Available online 2 May 2025

improved LCS deployment include careful calibration, strategic experimental design, and focused application on lower time-resolution monitoring.

1. Introduction

Air pollution has become a pressing global health issue, with severe consequences for human well-being (West et al., 2016). The World Health Organization (WHO) estimates that air pollution was responsible for over 8 million premature deaths in 2021, with the majority occurring in low- and middle-income countries (Health Effects Institute, 2024). In Europe, air pollution is the leading environmental health risk, causing approximately 400,000 premature deaths in 2020 (EEA, 2022). Pollutants such as fine particulate matter (PM2.5), nitrogen dioxide (NO_2), and ground-level ozone (O_3) contribute to a wide range of health issues, including respiratory and cardiovascular diseases, lung cancer, and adverse birth outcomes (Howse et al., 2021). Vulnerable populations, such as children, the elderly, and those with pre-existing conditions, are particularly at risk (Huang et al., 2023; Sharma et al., 2024).

In light of these significant health impacts, it is crucial to better understand the levels and patterns of human exposure to air pollution (Canha et al., 2021). Exposure studies aim to assess how individuals or populations come into contact with air pollutants in their everyday environments, such as their homes, places of work, or during travel (da Schio, 2020; Dias and Tchepel, 2018). These studies help to quantify the amount and duration of pollutant exposure, providing insights into how air pollution contributes to adverse health effects. By focusing on different exposure sources, such as ambient outdoor air, indoor air, or occupational environments, exposure studies can identify key risk factors and inform policies to reduce pollution-related health risks.

Exposure studies involve various methodologies, including monitoring air pollution levels in different environments and applying models to estimate exposure based on geographic and time series of pollutant concentrations. These approaches help researchers link ambient concentrations of pollutants to human exposure and ultimately to risk estimation (Anenberg et al., 2016). The findings have been crucial for designing effective interventions, refining air quality standards, and developing public health strategies that aim to mitigate the health burdens associated with air pollution, e.g., in the process of updating the WHO air quality guidelines (WHO, 2021).

The advent of low-cost sensors (LCS) has revolutionized air quality monitoring by making it more accessible and affordable (Peltier, 2020; WMO, 2018). Traditionally, air pollution data has been gathered using expensive, stationary monitoring stations that are sparsely distributed, often leaving large geographic areas and vulnerable populations without sufficient coverage. In many low-income countries or remote regions, the deployment of these conventional systems is financially and logistically prohibitive, limiting the understanding of air pollution levels in critical areas. LCS technology, however, offers a practical and cost-effective option, allowing air quality measurements to be conducted in places that previously lacked monitoring infrastructure (Morawska et al., 2018).

Low-cost sensors are compact, portable devices capable of measuring key air pollutants. Their affordability and ease of deployment have opened up new possibilities for air quality studies in dynamic and underrepresented contexts, e.g., in low- and middle-income countries. For instance, LCS can 1) be installed in dense networks across urban areas to help capture spatiotemporal variations and the identification of emission hotspots or characterize microenvironments (Kim Oanh et al., 2024; Kumar et al., 2015), 2) be deployed in rural or industrial regions with little prior data, and 3) provide a more detailed and nuanced understanding of air quality variability, both spatially and temporally (Islam et al., 2024; Lung et al., 2024).

The use of LCS in citizen science and community-based monitoring initiatives has further democratized air quality monitoring, enabling local stakeholders to engage directly with environmental health issues (English et al., 2020; Masri et al., 2022; Raheja et al., 2022). By generating high-resolution, localized data, LCS have the potential to complement traditional monitoring systems and offer valuable insights for policymakers and researchers seeking to address air pollution's complex and uneven distribution (WMO, 2024). Despite challenges with data accuracy and sensor calibration (Karagulian et al., 2019; Williams et al., 2019), LCS may be relevant in expanding global air quality monitoring capabilities and informing public health interventions.

Highly resolved data in space and time are critical for exposure studies to capture the detail of the levels to which the agents, who move and are exposed to air pollution of varying origins and concentrations, are subject (Canha et al., 2021). Such resolution can benefit the quantification of the different endpoints and eventually inform policy (Park and Kwan, 2017; Xie et al., 2017). For instance, a person on a sidewalk may be subject to a very high, very short peak in NO_2 concentration due to a passing heavy-duty vehicle. Short peaks are irrelevant for regulatory purposes but important for calculating exposure: with higher time resolution, the time during which communities or individuals are exposed to high concentrations can be more precisely estimated. By their nature, LCS, and LCS networks, have the potential to deliver data with the spatial and temporal detail to resolve such peaks. However, the available body of knowledge on the LCS performance in delivering high temporally resolved data is scarce.

In the present study, we address the research question whether calibrated LCS are suitable for delivering accurate concentrations at the high time resolution needed for exposure studies. In particular, we assess the performance of an electro-chemical sensor against high-end cavity attenuated phase shift (CAPS) and chemiluminescence NO_2 instruments as a function of averaging time. In addition, we also examine performance in the context of NO and O_3 concentrations to understand the impact of chemistry and emissions on performance.

2. Methodology

2.1. Monitoring site

The instruments used in the present study were deployed at the Hardenbergplatz station (station European code: DEBE067 or locally MC115 – https://luftdaten.berlin.de) in Berlin. The station is part of the Berlin Air Quality monitoring network (BLUME) and located in a high traffic area, with an important bus and rail interchange in the direct vicinity of the sampling point. The data analyzed was collected from 19th May 2023 until 11th November 2023. Historical NO_2 values by season can be seen in Table 1.

The Chemiluminescence and CAPS instruments were connected to the same sampling line with roughly the same length of sampling line (3–4 m) between outdoor air and instrument. The sensor systems containing LCS were placed on the roof of the station, less than one meter away from the combined inlet that passed air to the reference instruments. These sensor systems do not have sampling lines but instead have ventilators to draw in new and expel old parcels of air every 10 s. Based on calculations using the sampling line length and pump flow rate of the air flowing to the reference instruments, we determined that residence time of air in the sampling line was negligible at 10-s time resolution.

2.2. Electrochemical NO_2 sensor

Electrochemical sensors are commonly used in low-cost setups to measure NO_2. As NO_2 interacts with the sensor, it reacts with the

working electrode embedded in an electrolyte solution, producing an electrical current proportional to the gas concentration. This simple mechanism allows electrochemical sensors to be compact, portable, and energy-efficient. In this study, EarthSense Zephyr© sensor systems were used, which are fitted with Alphasense EC sensors for the measurement of NO_2, NO, and O_3. The raw voltage output by these electrochemical cells was calibrated following the 7-step methodology (Schmitz et al., 2021). This field-calibration methodology improves both individual performance and sensor intercomparison, which is key given that LCS are known for their differing sensitivities and lifetimes (Borrego et al., 2018; Karagulian et al., 2019; Malings et al., 2019). With this approach the sensors are optimized for field performance with their current sensitivity, regardless of age, which allows for better comparison between devices.

To assess impacts of known LCS cross-sensitivity to other pollutants when measuring NO_2, an analysis of performance against various reference NO and O_3 concentrations was made, as these data were available. This could not be done for other VOCs, however, as these were not measured at the site, and is therefore a limitation of the study. LCS are known to have a response time of up to a minute to rapidly changing concentrations (Papaconstantinou et al., 2023) and are sensitive to temperature and relative humidity.

In this study, the full 6-month dataset was used to both train and test the calibration model. This choice was part of the experimental design and sought to ensure maximum possible accuracy of the linear regression model for this specific dataset. These are idealized conditions and are not intended to be generalized to other datasets, but rather to isolate impacts of time-resolution and reference instrument on LCS performance.

In the 7-step methodology, which was used in this study (Schmitz et al., 2021), the distributions of the reference (chemiluminescence and CAPS) and Zephyr data are first visually compared to gain insight whether the distributions had similar shapes and ranges. Second, point outliers are determined using a self-developed outlier detection function based on a z-score threshold and a running window. In the third step, data points outside of the bounds of the calibration set are flagged. In the fourth step, models using the output voltages and operating conditions (internal relative humidity and temperature of the Zephyr) as independent variables and the reference concentrations as dependent variable are developed. The model selection was based on four performance metrics: squared correlation coefficient (R^2), root mean square error (RMSE), mean average error (MAE) and Akaike information criterion (AIC). The selected model is validated in a fifth step by repeatedly splitting the dataset into 70/30 % training/test subsets in a continuous cross-validation. Final validated multiple linear regression models included raw data from all three EC sensors, reflecting the relative chemical importance of all three pollutants to urban NO_2 concentrations. This also mitigates known sensitivity of the NO_2 EC sensor to O_3 concentrations, as these are accounted for by another EC sensor. Averages of the error metrics calculated on each subset are then used to evaluate model performance and used in uncertainty calculation. The computation of the calibrated concentrations and their respective uncertainty constitute the sixth and seventh steps.

2.3. Chemiluminescence NO_2 monitor

Chemiluminescence is a widely used technique for measuring nitrogen oxides (NO_x) developed in the mid-20th century (Hodgeson, 1974). It relies on the chemiluminescent reaction of nitric oxide (NO) with ozone (O_3) to produce light in direct proportion to the NO concentration, which is then quantified by a photodetector. Since NO_2 itself does not directly produce chemiluminescence, it must first be converted to NO. This two-step approach allows for the detection of both NO and NO_2 by sequentially measuring NO alone and then NO + NO_2, with NO_2 derived by subtraction. The NO_2 to NO reduction is conducted at high temperatures at the surface of a molybdenum catalyst. The process can introduce interferences from other nitrogen-containing compounds, like nitrates or ammonia, which may also convert to NO on the catalyst surface. This can lead to overestimations of NO_2 levels (Dunlea et al., 2007; Matthews et al., 2002). The thermal catalytic conversion relies on diffusion and surface reactions at elevated temperatures and introduces a slight delay in response time. As an alternative, a photolytic ultraviolet convertor can be used.

The instrument used at the Hardenbergplatz station, part of the Berlin Air Quality Monitoring network (BLUME) instrumentation, is a HORIBA APNA-370 NOx Monitor which measures NO_2 concentrations in ambient air in accordance with the DIN EN 14211 norm (DIN, 2012). The instrument has a precision of <1 % and a limit of detection of 0.5 ppbv (3 δ). For this experiment the lowest available time resolution of the data was 5 min.

2.4. Cavity attenuated phase shift (CAPS) NO_2 monitor

The CAPS technique emerged around the early 2010s as an innovative approach for achieving high-precision, real-time monitoring of NO_2. CAPS monitors were designed to address limitations in existing NO_2 monitoring methods, such as chemiluminescence, which often rely on indirect methods and are sensitive to cross-interference. The CAPS instrument directly measures the light absorbed by NO_2 in a portion of the blue region of the spectrum (450 nm) which, complemented by the small, optically stable cavity and with the use of phase-shift detection, offers high selectivity.

The CAPS instrument used (Aerodyne Research CAPS NO_2 Monitor) was factory calibrated at the end of 2022 and is capable of measuring NO_2 concentrations with a precision <0.5 ppbv (2σ) and a limit of detection <1 ppbv. The time resolution was set to 1 s, suitable for capturing dynamic changes, such as those found in urban traffic environments.

2.5. Data analysis

The agreement between the instruments was investigated by interpreting the following metrics: R^2, MAE, the relative error and MBE. The squared correlation coefficient, R^2, quantifies the degree of linear association between the measurements from two instruments. The MAE represents the average absolute difference between the experimental and reference measurements, computed by averaging the pairwise absolute errors across all data points. The relative error, expressed here as a percentage, is derived by normalizing the absolute errors (by the electrochemical cell, after calibration) relative to the reference

Table 1
Historical NO_2 values by season at the Hardenbergplatz measurement site. Concentrations were measured by the same Chemiluminescence instrument used in this study.

NO_2	2015	2016	2017	2018	2019	2020	2021	2022	2023
Winter	49.7	47.9	47.3	42.6	38.7	29.2	25.8	20.0	21.7
Spring	51.9	51.3	45.9	45.9	30.5	22.4	20.5	19.7	18.0
Summer	54.6	51.5	43.5	37.9	31.5	23.6	18.8	15.9	15.3
Autumn	54.5	53.5	42.3	47.3	34.3	27.4	25.4	22.7	20.5

measurements (CAPS or chemiluminescence), offering insight into error size relative to measurement magnitude. The MBE measures the average bias in the experimental instrument's readings by averaging the differences, with signal, between the experimental and reference values; this metric indicates whether the experimental instrument tends to over- or underestimate compared to the reference.

In interpreting these metrics, each offers different insights into the LCS performance. An R^2 value closer to 1 indicates a strong linear relationship and suggests the LCS closely follows the reference instrument's trends. A lower MAE implies that, on average, the experimental readings deviate minimally from the reference values, indicating accuracy in individual readings. The relative error gives a sense of error proportionate to the measured values, where lower percentages indicate better performance, particularly important when measurements span a wide range of values. The MBE provides information on systematic bias: whether the LCS tends to overestimate or underestimate. Together, these metrics allow for a comprehensive assessment of the LCS performance.

3. Results

3.1. CAPS vs. Chemiluminescence comparison

The CAPS and the Chemiluminescence operate under very different measurement principles. As such, we compared their performance against each other across the entire measurement campaign. The highest possible time-resolution for comparison was 5-min, though the CAPS was operated at 1-s resolution so it required averaging. As can be seen in Fig. 1, their agreement varies substantially according to the averaging period considered. At 5-min resolution (Fig. 1a), the R^2 is low for reference instrument standards (< 0.8), with a high MAE (> 4 μg/m^3). The majority of this disagreement arises during short-term (< 10s) peaks in NO_2 concentrations, as evidenced by the standard deviation (SD) of averaged 1-s CAPS data. The greater the SD of the CAPS data (indicating high peak values), the greater the disagreement with the chemiluminescence. When the averaging period is increased to 1-h resolution (Fig. 1b), the agreement improves considerably ($R^2 = 0.99$; MAE = 1.15 μg/m^3). Though there is substantial disagreement between the instruments at high time-resolution peak concentrations, much of this is averaged out when considering coarser time-resolution. As such, bias between the instruments is not the source of the disagreement and both instruments capture similar base concentrations of NO_2.

Closer performance comparisons of the agreement between the two instruments were done and can be seen in Fig. 2. At 5-min resolution, differences in agreement are seen at different times of the day and in different months. When considering the time of day, R^2 decreases and relative error increases between 06:00 and 20:00. At its peak at noon, the R^2 is half of what it is in stable night-time concentrations and the relative error is roughly double. Regardless of the time of day, the MBE (calculated by subtracting the CAPS concentrations from those of the Chemiluminescence) is negative, indicating that the Chemiluminescence instrument measures less NO_2 than the CAPS, regardless of the time of day. MBE does not appear to peak alongside R^2 and relative error at noon. No notable differences in performance exist between the various days of the week. The R^2 improves in autumn months compared to the summer and relative error decreases accordingly.

At 1-h resolution, all of these trends are no longer seen. Agreement is very high ($R^2 > 0.95$) regardless of the time of day, the weekday, or the month. The relative error is also much lower (< 10 %) than at 5-min resolution and appears to decrease slightly from May to October. The MBE is relatively unchanged at 1-h resolution compared to 5-min resolution, with the Chemiluminescence tending to measure minimally lower concentrations (~ -0.8 μg/m^3) than the CAPS.

3.2. Zephyrs compared against the CAPS and Chemiluminescence instruments

The comparison of the Zephyrs to the CAPS and Chemiluminescence instruments yield varying results across different time-resolutions. As can be seen in Fig. 3, as the averaging time increases from high (10-s) to low (6-h) resolution, the performance increases substantially. This trend is seen across all three Zephyrs used in the study, though with slightly different magnitudes. At 10-s resolution, the R^2 is between 0.38 and 0.51, and the relative error is between 47 and 59 %. This performance increases steadily as the time resolution decreases to $R^2 > 0.8$ and relative error < 20 % at 6-h resolution, though the slope decreases steadily after 300 s (5-min resolution). As the time-resolution decreases, so too do the performance differences between the Zephyrs, though these are not large at any time-resolution.

A notable difference in performance is seen at 5-min resolution when the Zephyrs are trained with CAPS data ($R^2 = 0.67$–0.78; relative error = 27–36 %) versus when trained with Chemiluminescence ($R^2 = 0.52$–0.63; relative error = 39–49 %) data. The Zephyrs are more accurate at lower time resolutions (⟨1−h) when trained with the CAPS than with the Chemiluminescence instrument, though at the 6-h time-resolution, this difference is marginal (CAPS: $R^2 = 0.81$–0.87; Chemiluminescence: $R^2 = 0.79$–0.86).

The MBE is in both instruments and for all Zephyrs very close to 0. This is likely due to the calibration approach taken in this paper, where all available data were used both as the training and test sets. As such, the error is distributed around 0 and averages out to either 0 or very

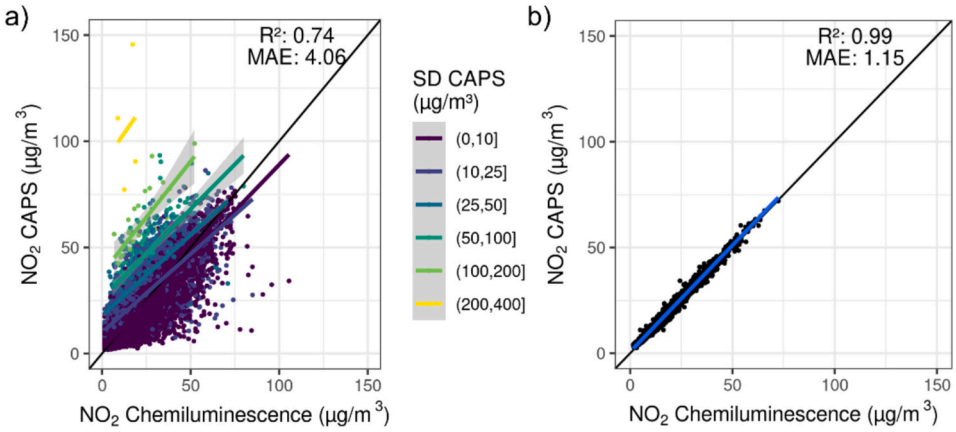


Fig. 1. Scatter plots of the performance comparison between the CAPS and Chemiluminescence instruments at a) 5-min and b) 1-h time resolutions using data from the entire campaign. The line of best-fit, the R^2, and mean average error (MAE) of the relationship are presented. At 5-min time resolution, points are colored by the standard deviation of the original resolution CAPS concentrations (1-s) and lines of best-fit are drawn accordingly for comparison to the 1:1 line (in black).

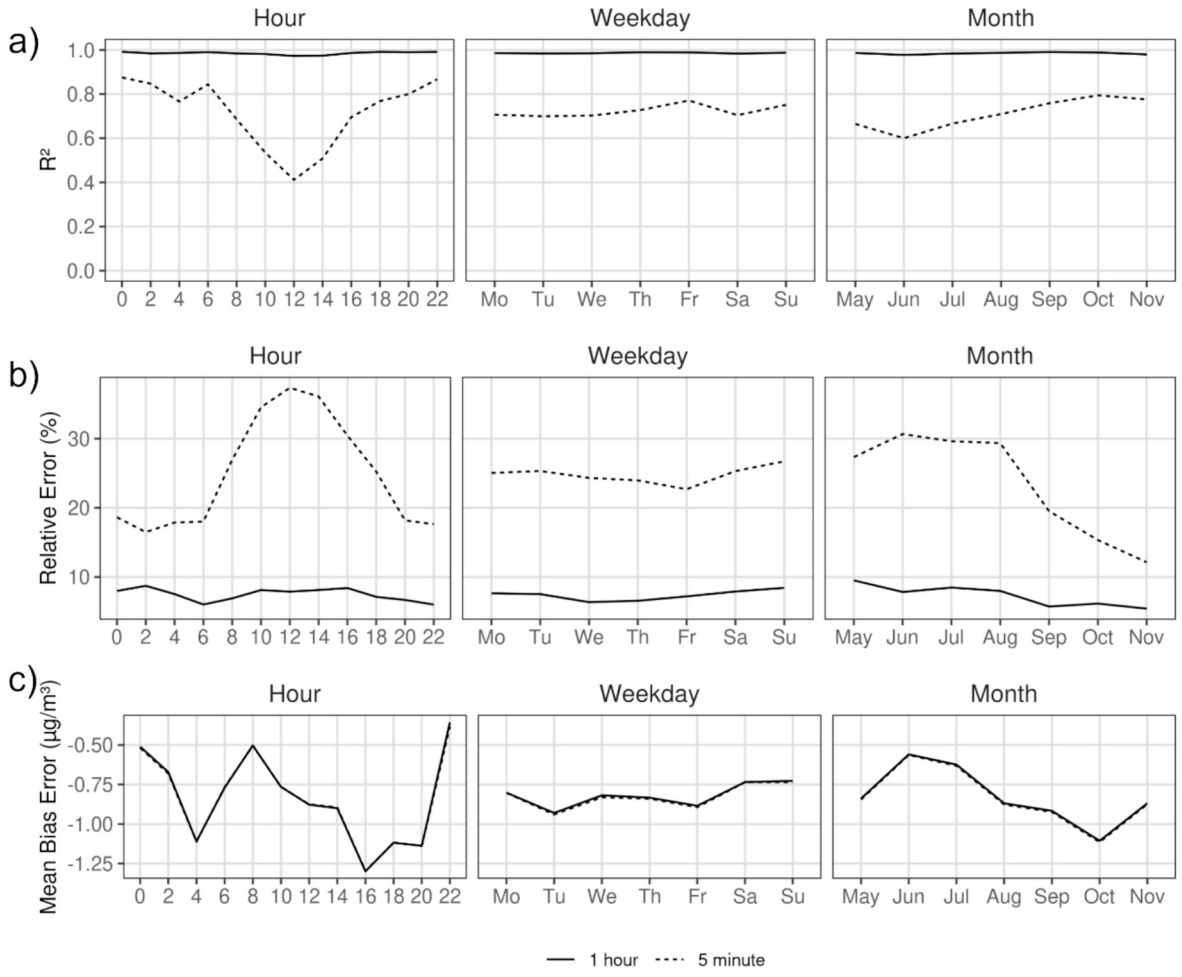


Fig. 2. Performance comparisons of the CAPS and Chemiluminescence instruments using various metrics (R^2, relative error, and MBE) by a) hour of the day, b) weekday, and c) month using 5-min and 1-h resolution data from the entire campaign.

close to it.

As with the comparison between the two reference instruments, a closer look was taken at the performance comparison between the Zephyrs and each of the reference instruments. In Figs. 4 and 5, the Zephyrs were compared to the CAPS and Chemiluminescence instruments, respectively, using the same three metrics, but broken down by time of day and day of the week. Several key trends are visible in these comparisons. When compared against either reference instrument, a diurnal pattern of Zephyr performance emerges (Fig. 4a-c and Fig. 5a-c). At night-time (22:00–04:00), the agreement (R^2) is good, while there is a small peak in relative error between 00:00 and 04:00 and a tendency towards overprediction (positive MBE). During the day, however, starting around 08:00 and ending around 15:00, R^2 steadily decreases and relative error increases, with a tendency towards underprediction (negative MBE). From 15:00 until 22:00, the performance then continuously increases until the nighttime stable period, with a shift back towards overprediction. Differences in hourly concentrations are not attributable to differences across weekdays and weekends. This trend is evident in all three Zephyrs. When compared to the CAPS instrument, the trend is visible at all time-resolutions higher than 5-min averages, after which the trend is reduced or negligible. In general, the relative change in the statistics is greater for the highest time resolutions and smallest for the lowest time resolution. For comparisons with the Chemiluminescence instrument, the trend is notable up until 20-min averages. This is further evidence that the poorer performance of the Zephyr against the Chemiluminescence (as compared to the CAPS) is resultant from the failure of the Chemiluminescence instrument to capture short-term peak concentrations.

When looking at performance metrics subset by days of the week (Fig. 4d-f and Fig. 5d-f), another trend emerges. On weekdays (Mo—Th), general agreement (R^2) is lower than on Fridays and weekends (Sa-Su), whereas the relative error is higher on weekends. There is a tendency towards underprediction on weekdays and overprediction on weekends. Performance increases substantially between the highest and lowest time-resolutions, but for the CAPS these stepwise increases are marginal for lower time-resolutions after 5-min averages. For the Chemiluminescence, performance increases more marginally at time-resolutions lower than 20-min averages.

To understand potential effects of chemistry and emissions on Zephyr performance versus each reference instrument, the metrics were broken down by concentration bins of NO and O_3 (Figs. 6–9 (MBE), and Figs. S1-S8 (R^2 and relative error)). The highest possible time-resolution for the NO and O_3 reference data was 5-min data, so higher time-resolutions available with the CAPS were not considered. Zephyr performance against the CAPS is best at mid-range concentrations of NO_2 (10–40 µg/m^3) and low O_3 concentrations (0–50 µg/m^3), as defined by R^2 (Fig. S2), relative error (Fig. S1), and MBE (Fig. 6). The highest relative error occurs at low NO_2 concentrations (0–5 µg/m^3) and high O_3 concentrations (>100 µg/m^3). For mid-range concentrations (the majority of the measurements of NO_2 and O_3), the Zephyrs tend to slightly overpredict NO_2 (positive MBE). At high concentrations of NO_2 (>60 µg/m^3) and high O_3 concentrations (>150 µg/m^3), the Zephyrs tend to underpredict NO_2 (negative MBE). When compared against the Chemiluminescence instrument (Fig. 8, Fig. S3 and S4), the trends are the

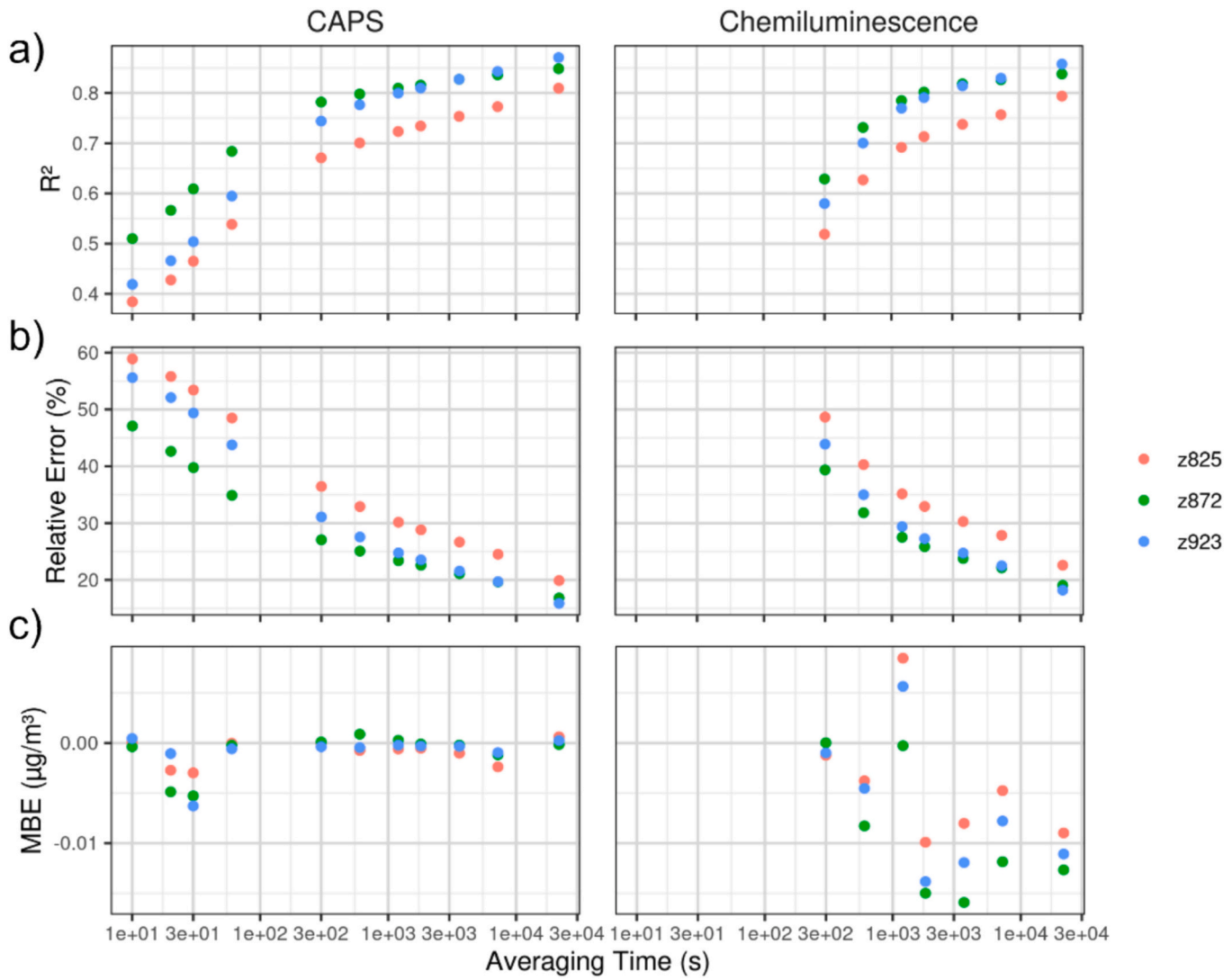


Fig. 3. a) R^2, b) relative error, and c) MBE of calibrated NO_2 concentrations of all three Zephyrs vs. NO_2 concentrations from each reference instrument across the entire campaign. The metrics were calculated at eleven different averaging times from ten seconds to six hours.

same, except that at low NO_2 the relative error is even higher and the R^2 is poorer at time resolutions above 20 min. In general, as time resolution decreases these trends subside, with R^2 increasing, relative error decreasing, and underprediction at high NO_2 and high O_3 concentrations subsiding.

When Zephyr performance is compared against increments of NO concentrations, the trends are different. Performance against the CAPS is best at medium concentrations of both NO_2 and NO (10–50 µg/m^3), as defined by R^2 (Fig. S6), relative error (Fig. S5), and MBE (Fig. 7), Relative error is highest at low NO_2 and NO concentrations (0–5 µg/m^3). The tendency is towards overprediction at most concentrations, which becomes more extreme at high NO concentrations (> 50 µg/m^3), except at high NO_2, where the tendency is to underpredict. These trends are similar when compared against the Chemiluminescence instrument, with poorer performance (relative error higher and R^2 lower) at low NO_2 and NO concentrations (0–10 µg/m^3). When compared to the CAPS, there is little difference in Zephyr performance across different time-resolutions, whereas there are notable differences between 5-min and 20-min averaging times when compared against the Chemiluminescence instrument.

4. Discussion

4.1. LCS performance at different time-resolutions

Our study is one of the first to comprehensively assess the performance of LCS at different time-resolutions. Our results show that as time-resolution decreases, the LCS performance increases. Given the similarity in sensor technology within different LCS systems, this is likely the case for any LCS with electrochemical sensors for NO_2. This suggests that LCS are well-suited for measuring ambient air pollution at lower time-resolutions, e.g., 1-h averaging intervals or coarser. Good performance at low time-resolution appears to be unaffected by the measurement principle (e.g., CAPS or Chemiluminescence) of the reference instrument used calibrate the LCS.

At medium time-resolution (e.g., 5-min intervals), however, there are notable differences in LCS performance when calibrated with the CAPS or the Chemiluminescence instruments. This appears to be due to the inability of the Chemiluminescence instrument to pick up high time-resolution peak NO_2 concentrations. As the SD of the averaged original resolution CAPS data increase, the worse the comparison is to the Chemiluminescence instrument, indicating that the Chemiluminescence fails to accurately capture peak concentrations. This translates to LCS performance, as at 5-min intervals, the LCS perform better against the

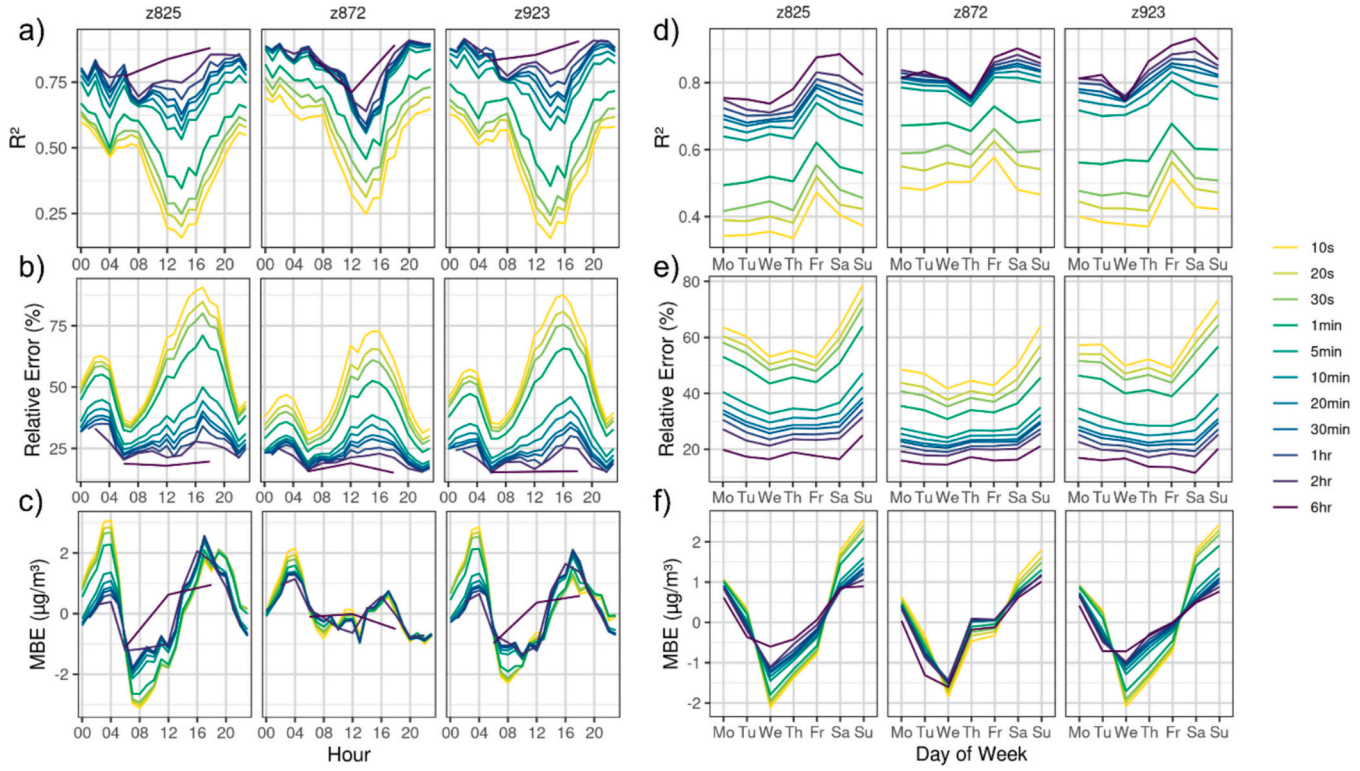


Fig. 4. Performance metrics of the three Zephyrs vs. the CAPS reference instrument by (a-c) hour of the day and (d-f) weekday across the entire campaign.

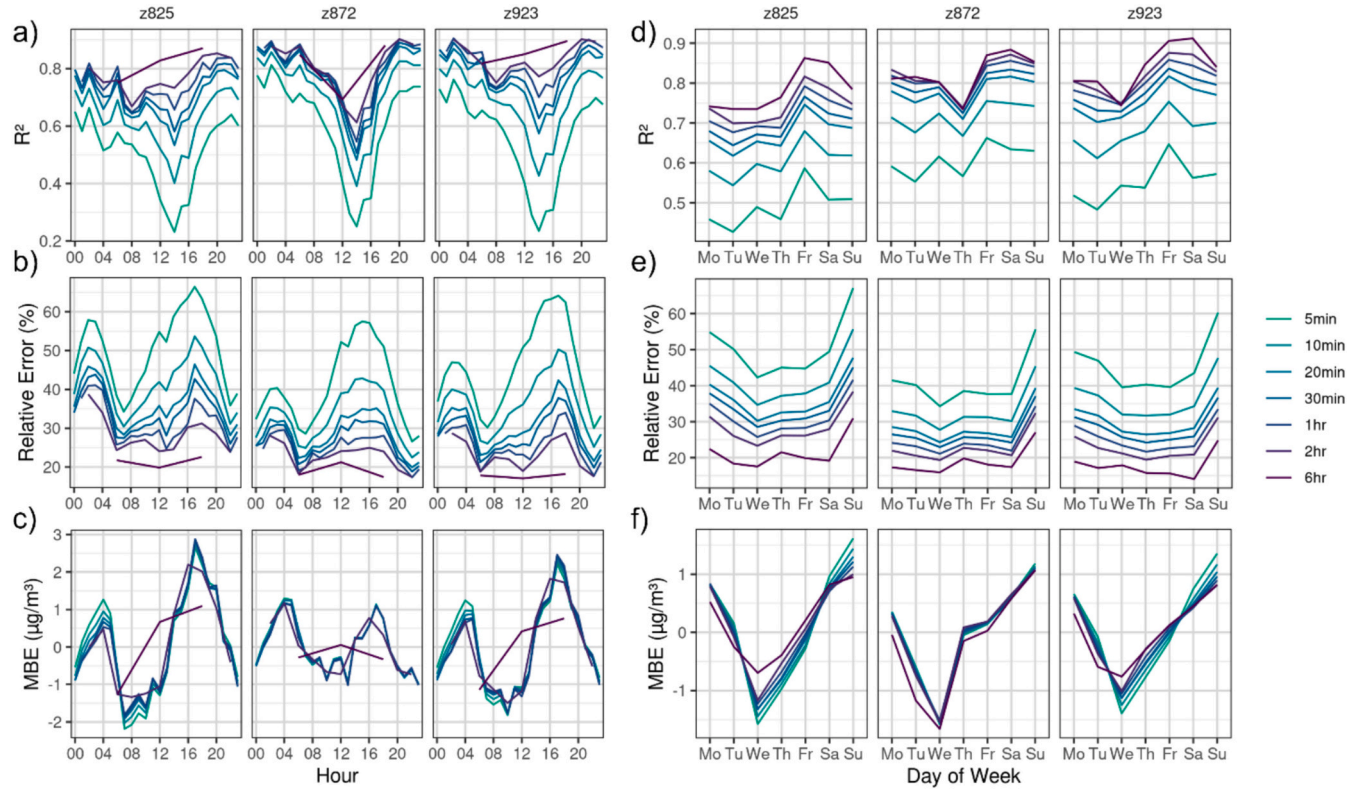


Fig. 5. Performance metrics of the three Zephyrs vs. the Chemiluminescence reference instrument by a-c) hour of the day and d-f) weekday across the entire campaign.

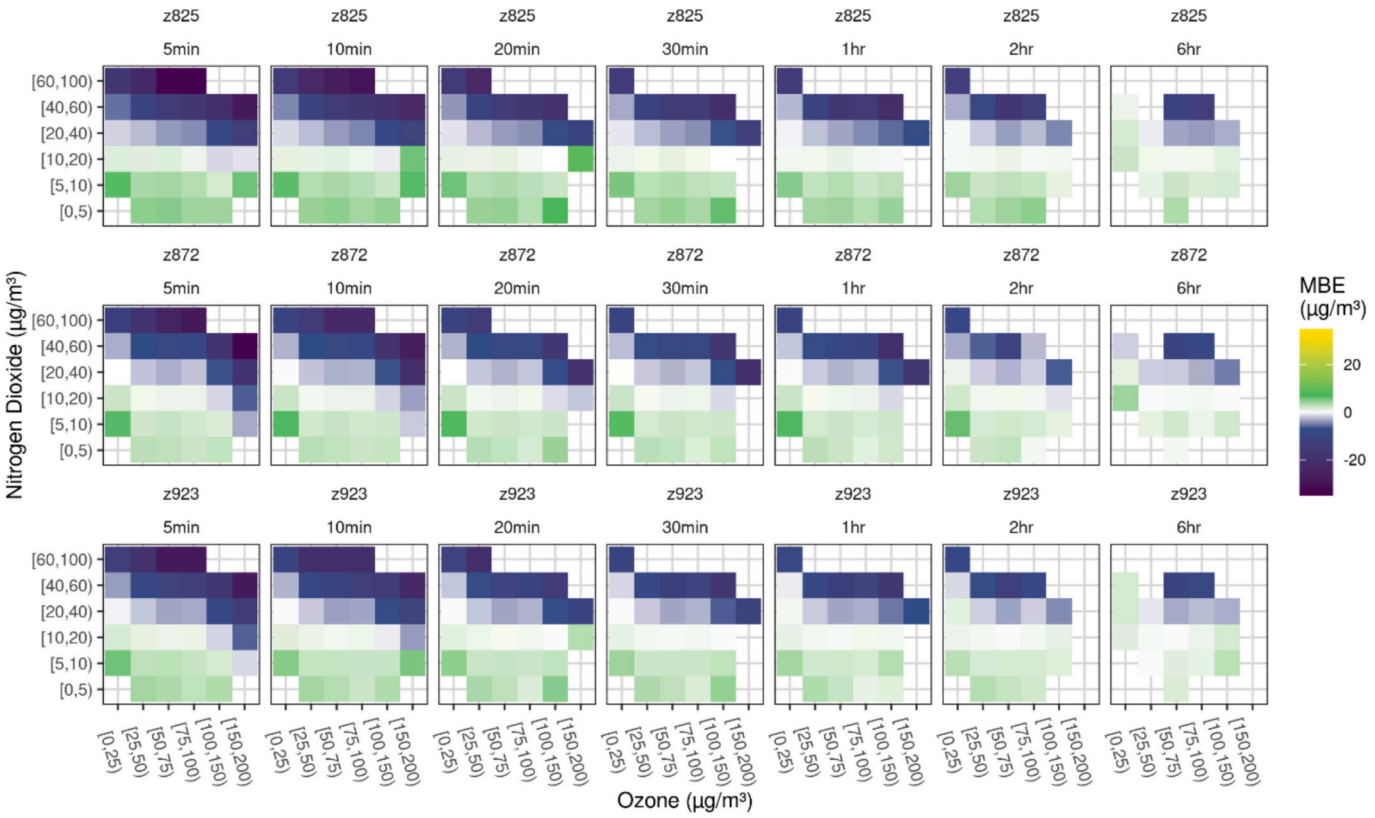


Fig. 6. MBE of the three Zephyrs vs. the CAPS reference instrument by increments of NO_2 and O_3 concentrations across the entire campaign. The highest available time resolution is 5-min data, as reference NO_2 and O_3 data was only available at that resolution.

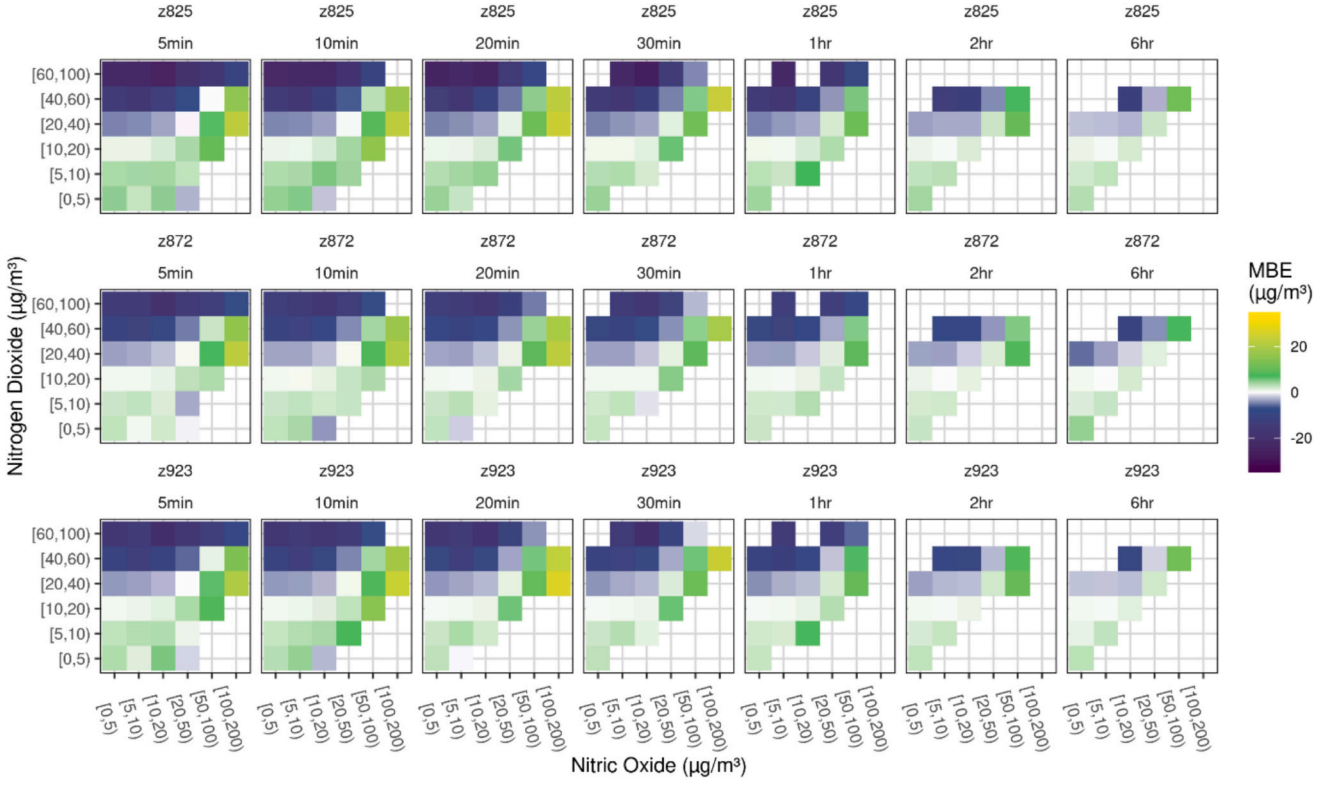


Fig. 7. MBE of the three Zephyrs vs. the CAPS reference instrument by increments of NO_2 and NO concentrations across the entire campaign.

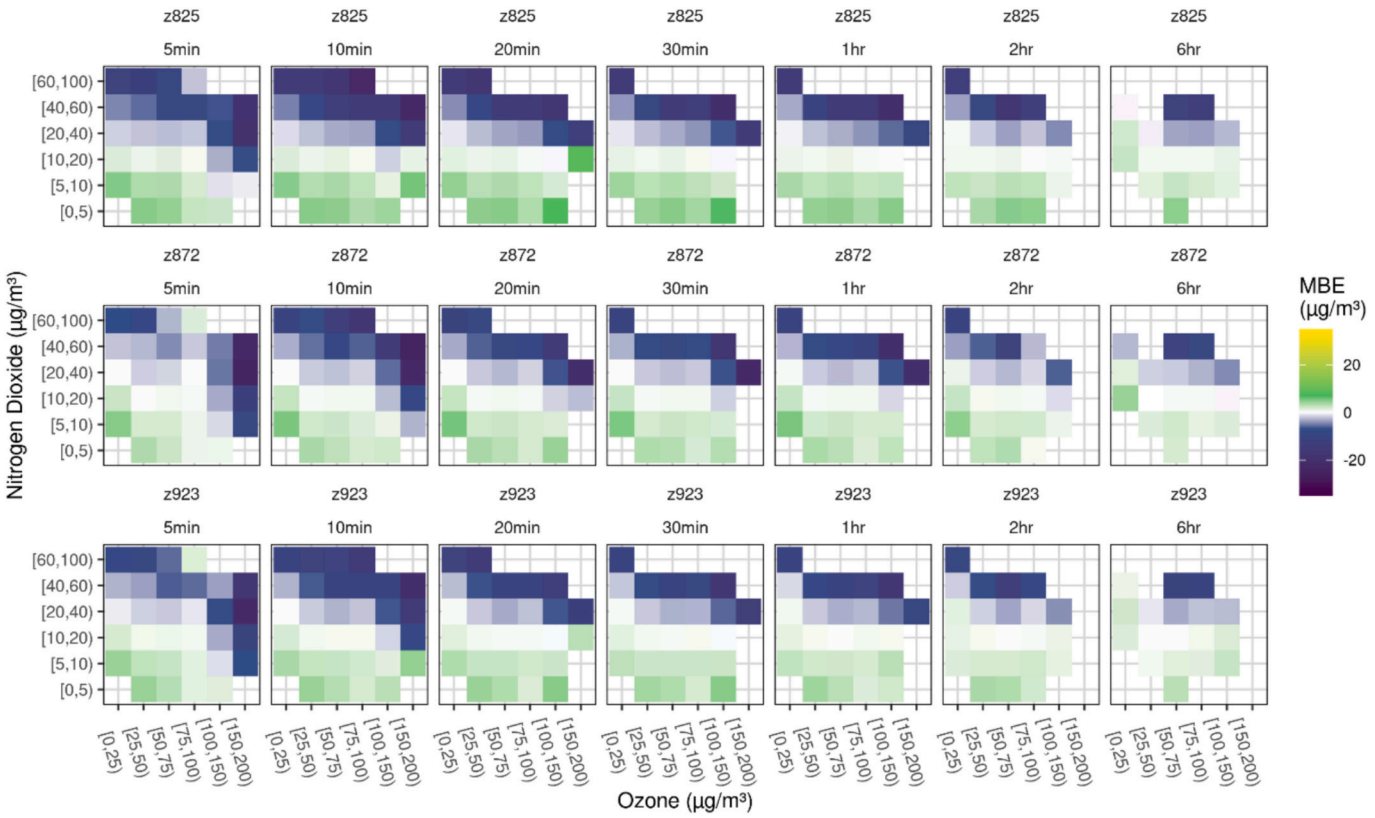


Fig. 8. MBE of the three Zephyrs vs. the Chemiluminescence reference instrument by increments of NO_2 and O_3 concentrations across the entire campaign.

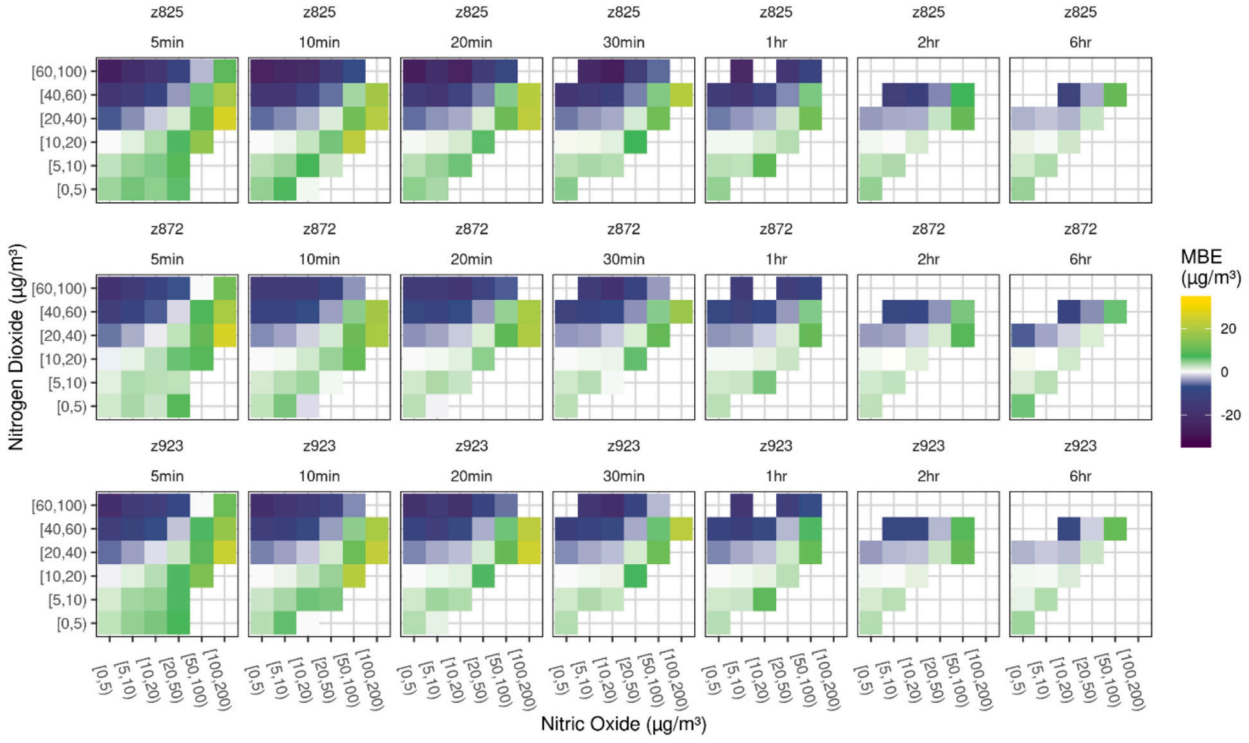


Fig. 9. MBE of the three Zephyrs vs. the Chemiluminescence reference instrument by increments of NO_2 and NO concentrations across the entire campaign.

CAPS than the Chemiluminescence instrument (see Section 3.2 and Fig. 3). We take this as evidence that the LCS are able to capture at least some of the peak concentrations more than the Chemiluminescence does, which is why there is a substantial performance gap at 5-min averages. The CAPS, however, does capture these high time-resolution peak concentrations, so the comparison with the LCS is then improved.

However, at even higher time-resolution (e.g., 10−/20-s intervals), when compared against the CAPS, LCS performance is generally poor

(Fig. 3, $R^2 < 0.5$, relative error > 50 %). This is driven mostly by poor performance during the daytime (Fig. 3), which is related to high NO_2, O_3, and NO concentrations (Figs. 6–7). Given that during more stable nighttime conditions LCS performance is very similar between high and low time-resolutions, we believe the fundamental driver of poor performance to be the physical limitations of the electrochemical sensors themselves, such as the known slow sensor recovery time following reaction (Mirsky, 2020). In dynamic environments, where for example, daytime chemistry or proximity to emissions sources (e.g., diesel vehicles) causes NO_2 concentrations to rapidly change, the electrochemical sensors likely do not respond quickly enough. Hence, the LCS capture some of these peaks, but not all of them, and typically underpredict their intensity, especially in high O_3 and NO environments because of competition among the oxidizing species.

4.2. Implications for exposure studies

LCS have become a popularly cited solution for improving measurements used in exposure studies, largely due to their lower cost and size, allowing them to measure in a variety of mobile and stationary positions. However, the results of our experiment presented in this study lead us to conclude that electrochemical sensors are not implicitly appropriate for use in exposure studies for the measurement of NO_2. They may yet have potential in such research, but there are key limitations to these technologies that must be properly considered.

First, as we have shown, electrochemical sensors are not good at capturing short-term peak NO_2 concentrations, but can be better when trained against a CAPS monitor. It is these peak concentrations in particular that are important to capture to measure real-world exposure to NO_2, as they have been postulated to account for the majority of the health effects (Gulliver and Briggs, 2004). In a variety of real-world situations (e.g., sitting in traffic, cycling alongside vehicles, walking on the sidewalk near a busy street), individuals are repeatedly exposed to elevated levels of NO_2. To properly quantify exposure to NO_2 pollution, LCS must be able to capture short-term peak concentrations, which we have shown they do not do well. If LCS are to be used in exposure studies, this limitation must be properly considered and mitigated through e.g., calibration against a CAPS monitor.

Second, exposure studies typically use LCS by providing them to individuals (pedestrians, cyclists, car-drivers, etc.) (Lin et al., 2017; Yixuan et al., 2021) and encouraging them to wear or hold them during their daily routines. While this approach seeks to understand direct exposures during each daily activity, it fails to consider the key limitations of the LCS, namely: 1) they do not respond quickly enough to rapid concentration changes, as this study has shown, 2) they typically do not have controlled air flow, and 3) they are susceptible to unknown levels of turbulence at the inlet. By adding another uncontrollable factor to this (movement of the device itself), it becomes even more challenging for peak concentrations to be properly captured. Further research should design experiments that explicitly test the response time of LCS in rapidly changing environments.

Last, if LCS are to be used in exposure studies, they need to be calibrated against a CAPS rather than a Chemiluminescence instrument. The CAPS has an appropriate time-resolution (1-s) and measurement principle (direct absorption) for capturing short-term peak NO_2 events. As such, it will capture the peak concentrations vital for exposure studies, thereby providing a more representative training dataset for calibrating the LCS.

4.3. Recommendations for measurements with LCS

In the context of this study, there are several recommendations that arise for use of LCS.

1) Researchers and individuals seeking to measure peak concentrations for e.g., exposure studies using LCS should design their experiments so as to account for the limitations of LCS. This includes, but is not limited to: i) using appropriate reference instruments in LCS calibration (e.g., CAPS), ii) minimizing uncertainties in air flow to the LCS through e.g., stationary deployment, iii) focusing on long-term measurements at specific exposure-relevant sites, rather than short-term coverage of multiple exposure-relevant environments. With proper experimental design, researchers can ensure best-possible coverage of peak concentration events.
2) Appropriate averaging times should be used for LCS measurement campaigns. For example, if longer-term stationary measurements of ambient NO_2 are envisioned, then calibrating LCS and reporting concentrations at 1-h intervals is likely sufficient (pending the size of the training set) and will show good performance. If shorter-term peak concentrations are important to capture for e.g., exposure studies, then LCS should be calibrated at the highest possible time resolution and averaged to coarser resolutions only as needed.
3) When performing measurements in mobile environments, careful consideration should be given to the repeated sampling of environments, because, as we have shown, electrochemical sensors can miss high resolution peak concentrations and are known to have response times of up to a minute (Papaconstantinou et al., 2023), which can lead to distortion of the sensor signal (Peltier, 2020). For example, if mobile measurements are performed in an urban environment, multiple loops of the same route should be conducted so that more data are available at all locations, thereby mitigating potential loss of individual short-term peak concentrations. Not accounting for these limitations in subsequent analysis could lead to the presentation of misleading results.

5. Conclusions

In this study we compared low-cost electrochemical sensors against high-end reference CAPS and chemiluminescence instruments for the measurement of NO_2 at various time resolutions. Our results suggest that electrochemical sensors are not suited for the measurement of NO_2 at the highest time-resolutions (e.g., 10-s), perform better at medium time-resolutions (e.g., 5-min) when calibrated against a CAPS reference instrument, and perform well at low time-resolutions (e.g., 1-h) regardless of which instrument they are trained with. This has implications for their use in the measurement of NO_2 pollution, as it suggests that due to their design, they require special treatment when used for research in which the high time-resolution capture of peak NO_2 concentrations (e.g., exposure studies, mobile measurements). Our evidence shows that LCS have diminished sensitivity to high time-resolution peaks, but that their performance at such resolutions can be improved if calibrated against a high time-resolution reference instrument such as the CAPS. Conversely, this study shows that these sensors perform well ($R^2 > 0.8$) at coarse time-resolutions such as is typically done in longer-term, stationary monitoring studies.

We recommend that researchers seeking to measure air pollution with electrochemical sensors take these physical limitations into account when designing their measurement campaigns. Depending on what the aim of the study is, researchers should 1) calibrate and report concentrations at an appropriate time-resolution, 2) use appropriate reference instruments for calibrating the LCS with, and 3) avoid measurement at high time-resolutions where possible, or design studies appropriately to account for the underprediction or ignorance of peak concentration events.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2025.179476.

CRediT authorship contribution statement

Seán Schmitz: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Alexandre Caseiro:** Writing – review

& editing, Writing – original draft, Software, Methodology, Investigation, Data curation, Conceptualization. **Erika von Schneidemesser:** Writing – review & editing, Supervision, Resources, Project administration, Conceptualization.

Funding

The research of EvS, SS, and AC is supported by the Research Institute for Sustainability (RIFS) at GFZ, with financial support provided by the Federal Ministry of Education and Research of Germany (BMBF) and the Ministry for Science, Research and Culture of the State of Brandenburg (MWFK).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Katja Grunow and Philipp Guse (Berlin Senate Department for Urban Mobility, Transport, Climate Action and the Environment), and the entire EarthSense team for their support of this work.

Data availability

Data will be made available on request.

References

Anenberg, S.C., Belova, A., Brandt, J., Fann, N., Greco, S., Guttikunda, S., Heroux, M.E., Hurley, F., Krzyzanowski, M., Medina, S., Miller, B., Pandey, K., Roos, J., Van Dingenen, R., 2016. Survey of ambient air pollution health risk assessment tools. Risk Anal. 36 (9), 1718–1736. https://doi.org/10.1111/risa.12540.

Borrego, C., Ginja, J., Coutinho, M., Ribeiro, C., Karatzas, K., Sioumis, T., Katsifarakis, N., Konstantinidis, K., De Vito, S., Esposito, E., Salvato, M., Smith, P., André, N., Gérard, P., Francis, L.A., Castell, N., Schneider, P., Viana, M., Minguillón, M.C., Penza, M., 2018. Assessment of air quality microsensors versus reference methods: the EuNetAir joint exercise – part II. Atmos. Environ. 193, 127–142. https://doi.org/10.1016/j.atmosenv.2018.08.028.

Canha, N., Diapouli, E., Almeida, S.M., 2021. Integrated human exposure to air pollution. Int. J. Environ. Res. Public Health 18 (5). https://doi.org/10.3390/ijerph18052233.

Dias, D., Tchepel, O., 2018. Spatial and temporal dynamics in air pollution exposure assessment. Int. J. Environ. Res. Public Health 15 (3). https://doi.org/10.3390/ijerph15030558.

DIN, 2012. Außenluft - Messverfahren zur Bestimmung der Konzentration von Stickstoffdioxid und Stickstoffmonoxid mit Chemilumineszenz. Deutsche Fassung EN 14211, 2012. https://doi.org/10.31030/1852617.

Dunlea, E.J., Herndon, S.C., Nelson, D.D., Volkamer, R.M., San Martini, F., Sheehy, P.M., Zahniser, M.S., Shorter, J.H., Wormhoudt, J.C., Lamb, B.K., Allwine, E.J., Gaffney, J. S., Marley, N.A., Grutter, M., Marquez, C., Blanco, S., Cardenas, B., Retama, A., Ramos Villegas, C.R., Molina, M.J., 2007. Evaluation of nitrogen dioxide chemiluminescence monitors in a polluted urban environment. Atmos. Chem. Phys. 7 (10), 2691–2704. https://doi.org/10.5194/acp-7-2691-2007.

EEA, 2022. Air Quality in Europe 2022.

English, P., Amato, H., Bejarano, E., Carvlin, G., Lugo, H., Jerrett, M., King, G., Madrigal, D., Meltzer, D., Northcross, A., Olmedo, L., Seto, E., Torres, C., Wilkie, A., Wong, M., 2020. Performance of a low-cost sensor community air monitoring network in Imperial County, CA. Sensors (Basel) 20 (11). https://doi.org/10.3390/s20113031.

Gulliver, J., Briggs, D.J., 2004. Personal exposure to particulate air pollution in transport microenvironments. Atmos. Environ. 38 (1), 1–8. https://doi.org/10.1016/j.atmosenv.2003.09.036.

Health Effects Institute, 2024. State of Global Air.

Hodgeson, J.A., 1974. A review of chemiluminescent techniques for air pollution monitoring. Toxicol. Environ. Chem. Rev. 2 (2), 81–97. https://doi.org/10.1080/02772247409356921.

Howse, E., Crane, M., Hanigan, I., Gunn, L., Crosland, P., Ding, D., Hensher, M., Rychetnik, L., 2021. Air pollution and the noncommunicable disease prevention agenda: opportunities for public health and environmental science. Environ. Res. Lett. 16 (6). https://doi.org/10.1088/1748-9326/abfba0.

Huang, H.C., Zou, M.L., Chen, Y.H., Jiang, C.B., Wu, C.D., Lung, S.C., Chien, L.C., Lo, Y. C., Chao, H.J., 2023. Effects of indoor air quality and home environmental characteristics on allergic diseases among preschool children in the Greater Taipei Area. Sci. Total Environ. 897, 165392. https://doi.org/10.1016/j.scitotenv.2023.165392.

Islam, M.S., Zaman, S.U., Islam, M.S., Roy, S., Yesmin, M., Lung, S.-C.C., Salam, A., 2024. Particulate matter exposure to street vendors in Dhaka city and associated health risk study. Air Quality, Atmosphere & Health. https://doi.org/10.1007/s11869-024-01666-4.

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

Kim Oanh, N.T., Huy, L.N., Maneepatra, W., Winijkul, E., Giandomenico, A., Tantrakarnapa, K., Co, H.X., Cuong, D.M., Tsou, M.-C.M., Hien, T.T., Chi, N.D.T., Ngan, T.A., Lung, S.C.C., 2024. Comparative analysis of PM2.5 levels in various microenvironments associated with common cooking practices in selected Asian countries. Air Qual. Atmos. Health. https://doi.org/10.1007/s11869-024-01615-1.

Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., Bell, M., Norford, L., Britter, R., 2015. The rise of low-cost sensing for managing air pollution in cities. Environ. Int. 75, 199–205. https://doi.org/10.1016/j.envint.2014.11.019.

Lin, C., Masey, N., Wu, H., Jackson, M., Carruthers, D.J., Reis, S., Doherty, R.M., Beverland, I.J., Heal, M.R., 2017. Practical field calibration of portable monitors for mobile measurements of multiple air pollutants. Atmosphere 8 (12). https://doi.org/10.3390/atmos8120231.

Lung, S.C., Tsou, M.M., Cheng, C.C., Setyawati, W., 2024. Peaks, sources, and immediate health impacts of PM(2.5) and PM(1) exposure in Indonesia and Taiwan with microsensors. J. Expo. Sci. Environ. Epidemiol. https://doi.org/10.1038/s41370-024-00689-4.

Malings, C., Tanzer, R., Hauryliuk, A., Kumar, S.P.N., Zimmerman, N., Kara, L.B., Presto, A.A., Subramanian, R., 2019. Development of a general calibration model and long-term performance evaluation of low-cost sensors for air pollutant gas monitoring. Atmos. Meas. Tech. 12 (2), 903–920. https://doi.org/10.5194/amt-12-903-2019.

Masri, S., Cox, K., Flores, L., Rea, J., Wu, J., 2022. Community-engaged use of low-cost sensors to assess the spatial distribution of PM2.5 concentrations across disadvantaged communities: results from a pilot study in Santa Ana, CA. Atmosphere 13 (2). https://doi.org/10.3390/atmos13020304.

Matthews, R.D., Sawyer, R.F., Schefer, R.W., 2002. Interferences in chemiluminescent measurement of nitric oxide and nitrogen dioxide emissions from combustion systems. Environ. Sci. Technol. 11 (12), 1092–1096. https://doi.org/10.1021/es60135a005.

Mirsky, V.M., 2020. Electrochemical sensors between the academic world and harsh reality: a few thoughts on the past, present, and future. J. Solid State Electrochem. 24 (9), 2147–2149. https://doi.org/10.1007/s10008-020-04711-1.

Morawska, L., Thai, P.K., Liu, X., Asumadu-Sakyi, A., Ayoko, G., Bartonova, A., Bedini, A., Chai, F., Christensen, B., Dunbabin, M., Gao, J., Hagler, G.S.W., Jayaratne, R., Kumar, P., Lau, A.K.H., Louie, P.K.K., Mazaheri, M., Ning, Z., Motta, N., Williams, R., 2018. Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: how far have they gone? Environ. Int. 116, 286–299. https://doi.org/10.1016/j.envint.2018.04.018.

Papaconstantinou, R., Demosthenous, M., Bezantakos, S., Hadjigeorgiou, N., Costi, M., Stylianou, M., Symeou, E., Savvides, C., Biskos, G., 2023. Field evaluation of low-cost electrochemical air quality gas sensors under extreme temperature and relative humidity conditions. Atmos. Meas. Tech. 16 (12), 3313–3329. https://doi.org/10.5194/amt-16-3313-2023.

Park, Y.M., Kwan, M.P., 2017. Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. Health Place 43, 85–94. https://doi.org/10.1016/j.healthplace.2016.10.002.

Peltier, R., 2020. An Update on Low-Cost Sensors for the Measurement of Atmospheric Composition.

Raheja, G., Harper, L., Hoffman, A., Gorby, Y., Freese, L., O'Leary, B., Deron, N., Smith, S., Auch, T., Goodwin, M., Westervelt, D.M., 2022. Community-based participatory research for low-cost air pollution monitoring in the wake of unconventional oil and gas development in the Ohio River Valley: empowering impacted residents through community science. Environ. Res. Lett. 17 (6). https://doi.org/10.1088/1748-9326/ac6ad6.

da Schio, N., 2020. Self-portraits of personal exposure to air pollution: on where and when people are exposed, and on why it is difficult to avoid. Hum. Ecol. 48 (4), 465–479. https://doi.org/10.1007/s10745-020-00176-y.

Schmitz, S., Towers, S., Villena, G., Caseiro, A., Wegener, R., Klemp, D., Langer, I., Meier, F., von Schneidemesser, E., 2021. Unravelling a black box: an open-source methodology for the field calibration of small air quality sensors. Atmos. Meas. Tech. 14 (11), 7221–7241. https://doi.org/10.5194/amt-14-7221-2021.

Sharma, A., Hsiao, H.-Y., Liu, J.-Y., Lung, S.-C.C., Su, H.-J., Shen, C.-F., Chen, N.-T., Wu, P.-C., Lin, C.-Y., Liang, S.-F., Cheng, T.-J., Chan, T.-C., Tsay, Y.-S., Chung, H.-Y., Wang, Y.-C., 2024. Assessment of PM2.5-associated mortality burden among the elderly: insights into demographic, socio-economic, and geographic factors. Air Qual. Atmos. Health 17 (7), 1535–1545. https://doi.org/10.1007/s11869-024-01526-1.

West, J.J., Cohen, A., Dentener, F., Brunekreef, B., Zhu, T., Armstrong, B., Bell, M.L., Brauer, M., Carmichael, G., Costa, D.L., Dockery, D.W., Kleeman, M., Krzyzanowski, M., Kunzli, N., Liousse, C., Lung, S.C., Martin, R.V., Poschl, U., Pope 3rd, C.A., Wiedinmyer, C., 2016. What we breathe impacts our health: improving understanding of the link between air pollution and health. Environ. Sci. Technol. 50 (10), 4895–4904. https://doi.org/10.1021/acs.est.5b03827.

WHO, 2021. WHO Global Air Quality Guidelines. Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide. World Health Organization.

Williams, R., Duvall, R., Kilaru, V., Hagler, G., Hassinger, L., Benedict, K., Rice, J., Kaufman, A., Judge, R., Pierce, G., Allen, G., Bergin, M., Cohen, R.C., Fransioli, P., Gerboles, M., Habre, R., Hannigan, M., Jack, D., Louie, P., Ning, Z., 2019. Deliberating performance targets workshop: potential paths for emerging PM2.5 and O3 air sensor progress. Atmos. Environ. X 2, 100031. https://doi.org/10.1016/j.aeaoa.2019.100031.

WMO, 2018. Low-cost sensors for the measurement of atmospheric composition: overview of topic and future applications.

WMO, 2024. Integrating Low-Cost Sensor Systems and Networks to Enhance Air Quality Applications.

Xie, X., Semanjski, I., Gautama, S., Tsiligianni, E., Deligiannis, N., Rajan, R., Pasveer, F., Philips, W., 2017. A review of urban air pollution monitoring and exposure assessment methods. ISPRS Int. J. Geo Inf. 6 (12). https://doi.org/10.3390/ijgi6120389.

Yixuan, J., Xinlei, Z., Chen, C., Yihui, G., Weidong, W., Zhuohui, Z., Jing, C., Haidong, K., 2021. On-field test and data calibration of a low-cost sensor for fine particles exposure assessment. Ecotoxicol. Environ. Saf. 211, 111958. https://doi.org/10.1016/j.ecoenv.2021.111958.