A 30-month Field Evaluation of Low-Cost CO$_2$ Sensors Using a Reference Instrument

Qixiang Cai[1,2], Ning Zeng[3*], Xiaoyu Yang[4], Chi Xu[5], Zhaojun Wang[4], Pengfei Han[6,7*]

[1]State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of
5   Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

[2]Qiluzhongke Institute of Carbon Neutrality, Jinan 250100, China

[3]Department of Atmospheric and Oceanic Science, and Earth System Science Interdisciplinary Center, University of
Maryland, College Park, Maryland 20742, USA

[4]Shandong Jinan Ecological and Environmental Monitoring Center, Jinan 250102, China

10   [5]State Environmental Protection Key Laboratory of Quality Control in Environmental Monitoring, China National
Environmental Monitoring Centre, Bejing 100012, China

[6]State Key Laboratory of Atmospheric Environment and Extreme Meteorology, Institute of Atmospheric Physics, Chinese
Academy of Sciences, Beijing 100029, China

[7]Carbon Neutrality Research Center, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029,
15   China

Correspondence to: Pengfei Han (pfhan@mail.iap.ac.cn); Ning Zeng (zeng@umd.edu)

Abstract. CO$_2$ monitoring networks with low-cost and medium-precision sensors (LCSs) have become an exploratory

20   direction for CO$_2$ observation under complex emission conditions in cities. Yet the performance of such LCS after

deployment in the field faces significant challenges due to environmental impacts (e.g., temperature and humidity) and

long-term drifts due to sensor degradation (e.g., the light source). Here, we conducted 30 months of co-located

observations using LCS instruments (named SENSE-IAP) with a reference instrument (Picarro) to study the long-term

performance of the LCSs under field conditions, which is essential for the correction and validation of mid-low cost CO$_2$

25   observation networks. The environmental correction system we developed effectively corrected the impact of daily

environmental changes, which reduced the root mean square errors (RMSE) from 5.9±1.2 ppm to 1.6±0.5 ppm for

SENSE-IAP. The corrections remained robust against seasonal environmental variations, and the daily RMSE was generally

1-3 ppm over the 30 months of observation. Long-term drifts, commonly occurring in LCS, resulted in biases reaching up to

27.9 ppm over two years. Furthermore, the seasonal drift cycle contributed an RMSE of up to 25 ppm after six months of the

30   deployment. While the environmental correction system could not correct these errors, a linear interpolation method

effectively corrected the long-term drift. The long-term drift correction significantly decreased the RMSE to 2.4 ± 0.2 ppm

over the 30-month observation. To improve the accuracy of high-density CO$_2$ networks utilizing LCSs, we recommend that

the calibration frequency is better within three months and not exceed six months, with optimal calibration performed during

winter and summer to maintain daily accuracy within 5 ppm. These findings suggest that SENSE-IAP instruments can be

35   deployed for a long period without the need for taking back to re-calibrate in the laboratory or frequent standard gas

calibration in the field, thereby significantly reducing time, labor, and financial costs.

Keywords: low-cost sensors; environmental correction; long-term drift calibration; CO_2 monitoring network

1. Introduction

40   Urban CO_2 emissions demonstrate complex spatial and temporal variability (Wada et al., 2011), influenced by diverse

emission sources (Gurney et al., 2012; Kellett et al., 2013), meteorological factors (Grimmond et al., 2002; Lateb et al., 2016)

and potential misinterpretation from biogenic fluxes (Miles et al., 2021). These complexities pose significant challenges in

accurately capturing and interpreting the intricate changes in urban CO_2 concentrations, highlighting the necessity for more

advanced and comprehensive monitoring solutions.

45   Recent advancements in medium-precision carbon monitoring technologies have made establishing high-density,

low-cost, and medium-precision sensor (LCS) networks a viable and competitive strategy (Müller et al., 2020; Shusterman et

al., 2018). This approach effectively addresses the challenges associated with the variability of CO_2 in urban environments,

which are characterized by complex emission sources, vegetation carbon sinks, and dynamic meteorological conditions. In

contrast to high-precision instruments like Picarro or ABB-LGR, LCSs offer an accuracy range of 1-10 ppm and are more

50   cost-effective, with prices reduced by more than an order of magnitude (5-15 thousand USD). This cost efficiency enables

large-scale deployment, making LCS an attractive option for comprehensive urban CO_2 monitoring (Lopez-Coto et al., 2017;

Turner et al., 2016; Wu et al., 2016; Zeng et al., 2021).

While cost-effective, low-cost non-dispersive infrared (NDIR) sensors, are sensitive to environmental changes and

often exhibit long-term drifts and abrupt jumps. Noise, environmental sensitivity, and temporal drifts result in raw

55   measurements from LCS that typically have more significant errors and uncertainties than the accuracy and resolution

required for urban CO_2 monitoring. Consequently, the accuracy of such sensors generally depends on the correction methods

that account for factors such as temperature, humidity, and pressure.

Currently, several cities have established high-density CO_2 monitoring networks utilizing LCS. For instance, the

SenseAir LP8 sensor, with a raw accuracy of ±50 ppm (SenseAir: LP8 Product Sheet, 2019), and the Vaisala CarboCap

60   GMP343 sensor, with a raw accuracy of ±3ppm+1% reading (Vaisala, 2020), are commonly employed. The Carbosense CO_2

sensor network in Switzerland, which uses the LP8 sensor (Müller et al., 2020), achieves an observation accuracy of 8-12

ppm through initial laboratory chamber correction and regular drift calibration via ambient co-location with nearby reference

instruments. Similarly, the Berkeley Environmental Air-quality and CO_2 Network (BEACO$_2$N) in California, USA, utilizes the GMP343 sensor. After correcting for bias and temporal drift using the in-situ method, the observation accuracy is

65 approximately 1-4 ppm (Shusterman et al., 2016, 2018), while the reported accuracy improved to 1.6-3.6 ppm after temperature correction (Delaria et al., 2021). However, the sensitivities of these instruments to environmental variables are inconsistent, posing significant challenges in calibrating many LCSs (Bigi et al., 2018; Delaria et al., 2021; Hagan et al., 2018; Martin et al., 2017). Additionally, compared to the stability of high-precision instruments, LCSs are more susceptible to temporal drift and fluctuations. To address the above issues, sensors are taken to the laboratory for regular correction or

70 undergo in situ field calibration using traceable standard gases. These correction processes, while essential, are labor-intensive and require substantial time investment.

  To support China's dual carbon goals and address the high spatial variability of CO_2 concentration in urban areas, the Institute of Atmospheric Physics, Chinese Academy of Sciences (IAP), established a network of 134 sites using SenseAir K30 sensor since 2017 (Han et al., 2024). This study presents the correction methods developed for the SenseAir K30

75 sensors and evaluates the accuracy achievable through the environmental dependence correction method based on laboratory simulation. At a field observation site in Beijing, environmentally corrected LCSs were deployed to co-locate with high-precision Picarro instruments for up to 30 months. By comparing the data with measurements from Picarro instruments, we gained more profound insights into the long-term performance and long-term drift characteristics of the LCS, as well as assessed the effectiveness of our long-term drift correction method. Our findings demonstrate that timely long-term drift

80 correction significantly improved the accuracy of urban CO_2 monitoring networks based on LCS and reduced time, labor, and money investment. This research provides valuable evidence for optimizing the deployment and maintenance of LCS-based monitoring networks in urban environments.

2. The application of SenseAir K30 sensor

  A multivariate linear regression analysis was used for environmental correction, which can improve the accuracy of the

85 SenseAir K30 sensor from its initial specification of ±30 ppm ±3 % of reading (SenseAir: K30 products sheets, 2022) to a range of 1.7-4.3 ppm (Martin et al., 2017). The environmentally corrected K30 sensor demonstrated reliability and consistency when compared to higher-quality instruments and standard gas under a controlled indoor environment, with a root mean square error (RMSE) ranging from 1 to 3 ppm on a monthly scale (Cai et al., 2024). Furthermore, the low-cost sensor exhibited highly consistent with Picarro system during on-road observations conducted using the same vehicle, with

90 an RMSE of 3.6 ppm (Liu et al., 2021). In a study by Bao et al., (2020), the sensor was utilized to measure the CO_2 vertical

profile in the lower troposphere in Hebei Province, China, showing good consistency with traditional gas chromatography measurements (BAO et al., 2020). Additionally, Cai et al., (2024) applied the low-cost sensor in an industrial park, revealing a CO_2 concentration enhancement of 5-28 ppm within the park compared to a reference site (Cai et al., 2025).

95 The JJJ network, deployed with low-cost sensors, has provided valuable insights into seasonal variations, urban-rural differences, and the homology of CO_2 and $PM_{2.5}$ (Han et al., 2024). The low-cost sensors have also proven effective in detecting signals related to COVID-19. Continued CO_2 measurements in Beijing showed a 15-ppm reduction during the 2020 lockdown period compared to the before and after periods. Similarly, regular on-road CO_2 observation in Beijing before, during, and after COVID-19 lockdown showed a 40–60ppm decrease during COVID-19 lockdown period (Liu et al., 2021). These applications demonstrate the versatility and reliability of low-cost sensors in capturing both environmental and

100 anthropogenic influences on atmospheric CO_2 concentrations.

3. Instrument design and correction methods of SENSE-IAP

The SENSE-IAP instrument integrates three K30 sensors alongside a Bosch BME680 (BME) sensor (BAO et al., 2020; Liu et al., 2021) , all collected by an updated version of BeagleBone Green Wireless (BBGW). The standard version of SENSE-IAP instrument also includes a Figaro TGS 2611 sensor for CH_4 detection and a Plantower PMSA003 for $PM_{2.5}$

105 measurements. These components are compactly integrated onto a single circuit board and housed within a weatherproof enclosure, as illustrated in Fig. 1.

The BME sensor is positioned close to the K30s to simultaneously monitor the temperature (T in °C), relative humidity (RH in %), and pressure (P in hpa) of the air mass inside the instrument. This design ensures real-time correction of the sensor response values, accounting for dynamically changing external environmental conditions and enhancing measurement

110 accuracy.

Figure 1. The layout of sensors in the standard version of SENSE-IAP instrument.

The raw signals from all sensors were collected at a frequency of 2 seconds, with a background noise level of approximately ±20 ppm. Fig. 2 shows the experimental results of continuously introducing standard gas over 25 hours to
115   evaluate the instrument's noise characteristics. As shown in Fig. 2, the Allan deviation (in ppm) decreases with increasing integration time. At a 2-second measurement interval, the noise level is 4 ppm, which decreases to approximately 0.2 ppm for integration times ranging from 2 minutes to 1 hour. However, the Allan deviation increases after 1 hour of integration time, indicating the presence of drift contributions.
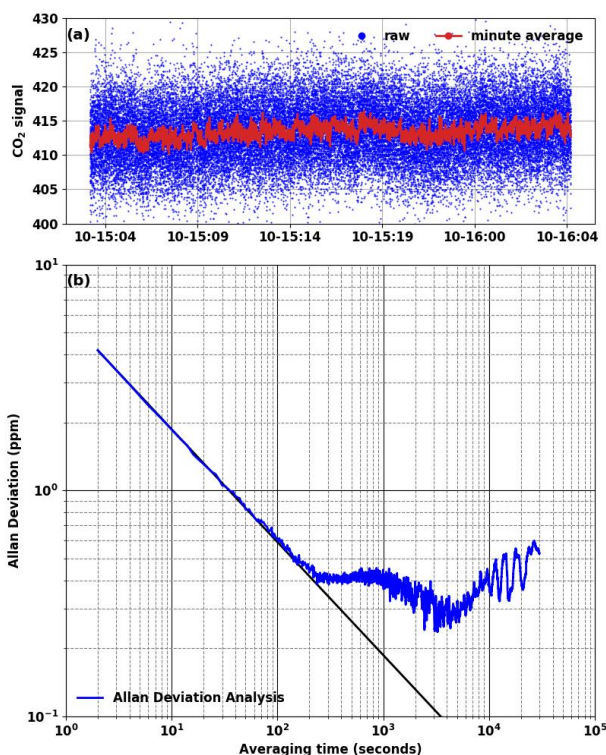
Figure 2. (a) The raw signals measured continuously over 25 hours from 03:00 on October 15th to 03:00 on October 16th, 2022. (b) Allan deviation log plots.

In addition to white noise, the raw signals often contain outliers that must be removed through quality control. According to formula 1, following the 3-sigma principle, the original data points xi collected during the sampling period are treated as samples. The average value $\overline{x}$ of non-missing x_i is calculated, along with the standard deviation (SD) of the xi is σ. If | xi-$\overline{x}$|>4 σ, the data point is identified as an outlier and removed. We adopted the 4σ threshold to strike a balance between effectively removing outliers and preserving the natural ranges of variability in the data.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}} \quad (1)$$

Subsequently, the raw signals are averaged from a 2-second interval to a 1-minute interval (resample) to reduce background noise. A 1-minute integration time was chosen as an optimal trade-off between noise reduction and maintaining sufficient time resolution to track natural variations in CO_2 concentrations accurately. As formula 2, the resampled value Y_i for each minute is calculated as the average of all non-outlier values x_j within that minute.

$$Y_i = \frac{\sum_{j=1}^{m} x_j}{m} \quad (2)$$
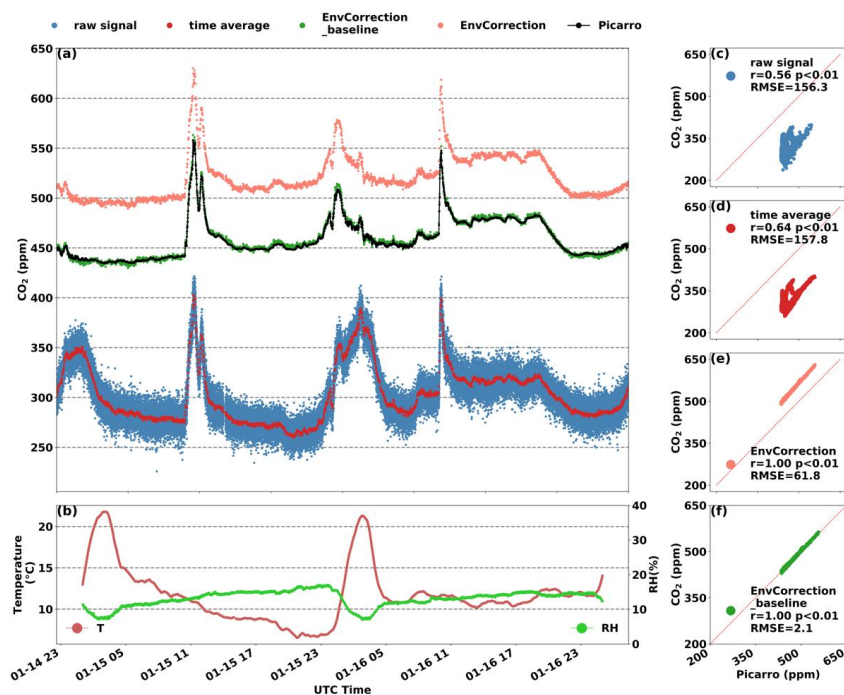
Figure 3. (a) Compared the CO_2 at each processing step, (b) Synchronized monitoring of T and RH. (c-f) The scatterplots of

CO₂ measurements form LCS and Picarro instrument at different processing steps, including (c) the raw signal in 2-second

resolution, (d) the values after noise reduction at a 1-minute resolution, (e) the CO_2 concentrations after environment

corrections and (f) systematic bias correction.

The correction system we developed substantially improves the accuracy of CO_2 measurements through a

comprehensive process that includes outlier removal and noise reduction. Fig. 3 shows the main steps of the correction

system, with the data cleaning method described earlier constituting the initial two steps. As shown in Fig.3 (a), the raw

signal (blue) undergoes de-specking and denoising (red). However, a noticeable deviation remains between the LCS

measurements and the true value, with a correlation coefficient (r-value) of approximately 0.6 (Fig.3(c-d)). The differences

mainly come from environmental sensitivity and baseline deviations in concentration.

Similar to the correction standards used in high-precision systems, our LCS units are calibrated using standard gas

traceable to the WMO X2007 scale. This calibration adjusts the span and calibrates system bias before deployment. For the

typical CO_2 concentration range (400-700ppm), a concentration-dependent offset (ΔC) exists between the time-averaged

LCS measurements and the standard gas concentration. Since this concentration dependence varies for each K30 sensor,

laboratory corrections are essential for accuracy improvement. Our concentration correction process includes multiple

concentration gradients, ensuring applicability to real-world monitoring scenarios. The fitting parameters of ΔC against the

measured values are determined through regression analysis, enabling precise correction across the operational range.

$$\Delta C = Y_m - y_0 \quad (3)$$

where, y_0 represents the concentration of standard gas or the high-precision reference instrument (Picarro G2301); Y_m is

the minute-averaged values from LCSs.

The environmental sensitivity correction of LCS includes T (10-50 °C), RH (10% -90%), and P compensation

correction. This correction is based on sensitivity testing conducted in the laboratory (Martin et al., 2017), with comparisons

made against Picarro. Each sensor is assigned unique sensitivity parameters through multivariate regression and iteration

analysis.

$$\Delta C = f(a_T, Y_T) + f(a_H, Y_H) + f(a_P, Y_P) + f(a_C, Y_C) + \varepsilon \quad (4)$$

where, the baseline correction coefficient is ε; the Y_T, Y_H, Y_P, and Y_C represent the compensation values for the T, RH, P,

and concentration sensitivity, respectively, applied to the minute-averaged CO_2 measurement; the a_T, a_H, and a_P are the

regression coefficients against T, RH and P respectively, while a_C is the span correction coefficient against concentration.

Thus, the corrected CO_2 can be expressed as:

$$C = Y_m - (f(a_T, Y_T) + f(a_P, Y_P) + f(a_H, Y_H) + f(a_C, Y_C) + \varepsilon) \quad (5)$$

The r values between the CO_2 corrected in the final two steps and the Picarro measurements are close to 1 (Fig. 3(e-f)).

Additionally, the difference between the environmental correction (pink) and baseline calibration (green) in Fig. 3(a)

represents the coefficient ε. After applying these correction steps, the accuracy of the LCS measurements improved to 1-4

ppm compared to Picarro (Liu et al., 2021).

Before deployment, the span and system bias of the LCSs were calibrated. However, once deployed to field stations,

LCSs tended to drift on a weekly to monthly scale, necessitating time-dependent drift calibration. We defined S_{cor} as the

starting time of drift and E_{cor} as the time when the drift slope stabilized or when calibration was required, with $\Delta = E_{cor} - S_{cor}$.

ΔC_{drift} represents the bias between the concentration C measured by the instrument and the standard concentration C_0 at E_{cor}.

Using the formula (6), the drift rate over time (ppm/min) at E_{cor} is calculated as s_t. The b_{cal} is a constant deviation,

representing the difference between the baseline and the standard value before long-term drift occurs (at S_{cor}). This value is

generally considered zero, since system bias has been calibrated before departure. The error at any time between S_{cor} to E_{cor}

can be calibrated to C_{Cor}^{drift} using the following formulas:

$$\Delta C_{drift} = s_t \Delta t + b_{cal} \quad (6)$$

$$C_{Cor}^{drift} = C - \Delta C_{drift} \quad (7)$$

This integrated instrument with environmental correction and drift correction is named as SENSE-IAP.

3. Co-located observation system

180   Our experiment has been conducted since July 2022 at IAP (Beijing-IAP site). Located in a central urban area with high population density, the Beijing-IAP site is significantly influenced by traffic emissions.

     To evaluate the performance of LCS, we developed a synchronous observation system that compares LCS with high-precision instruments. This system includes two SENSE-IAP units (numbered pi688 and pi736), each equipped with three K30 sensors. A cavity ring-down spectrometer (Picarro G2301) was used as the high-precision instrument for CO_2

185 measurements (Picarro G2301 Analyzer Datasheet, 2023). The precision and accuracy of the Picarro instrument are better than 0.1 ppm (Yang et al., 2021). At the Beijing-IAP, the Picarro analyzer was calibrated monthly using high-pressure standard gases provided by the Meteorological Observation Center of the China Meteorological Administration (MOC/CMA), which are traceable to the World Meteorological Organization (WMO) X2007 scale.

     To ensure long-term synchronous observation between the LCSs and Picarro, deployment enables two sets of

190 instruments measure the same gas mass. This ensures that any differences in observed values only come from the effects of T, RH and P as well as the concentration span, all of which can be adjusted through correction methods. The deployment setup is shown in Fig. 4(b-c).

     Ambient air is drawn from outside the window using a pump, the intake connected through a pipe equipped with a particulate matter filter and a water dryer. A four-way valve splits the gas stream, directing it to both the LCSs and the

195 Picarro. The instruments were hung on an open window to directly measure the outdoor air and environmental changes, which are the same as field-deployed conditions.
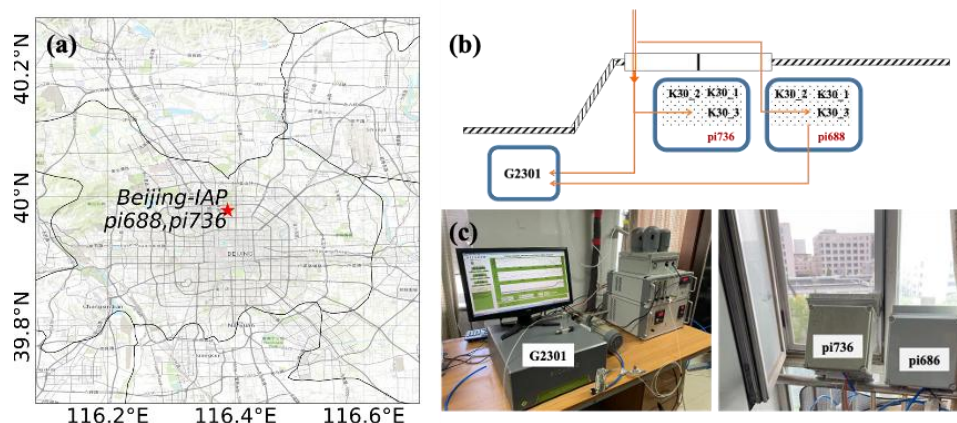
Figure 4. (a) Map of the location of Beijing-IAP, (b) the diagram of gas flow design for the synchronous observation system, (c) the photographs of the instrument installation setup. The source for the basemap used in subplot (a) was from ESRI

200 (https://server.arcgisonline.com/arcgis/rest/services/World_Topo_Map/MapServer).

4. Environmental corrections for field measurements

Fig. 5 shows the results from environment-corrected SENSE-IAP at the Beijing- IAP site, compared with those from the Picarro system. After approximately two weeks of data collection during both summer and winter, the SENSE_IAP showed highly consistent results with Picarro, with RMSEs of 1.6 ppm in summer and 1.8 ppm in winter. In contrast, the raw

205   CO_2 concentration data from the SenseAir showed a higher RMSE of 6.2 ppm in summer and 7.0 ppm in winter. Furthermore, the effectiveness of environmental correction is evident across different seasons.
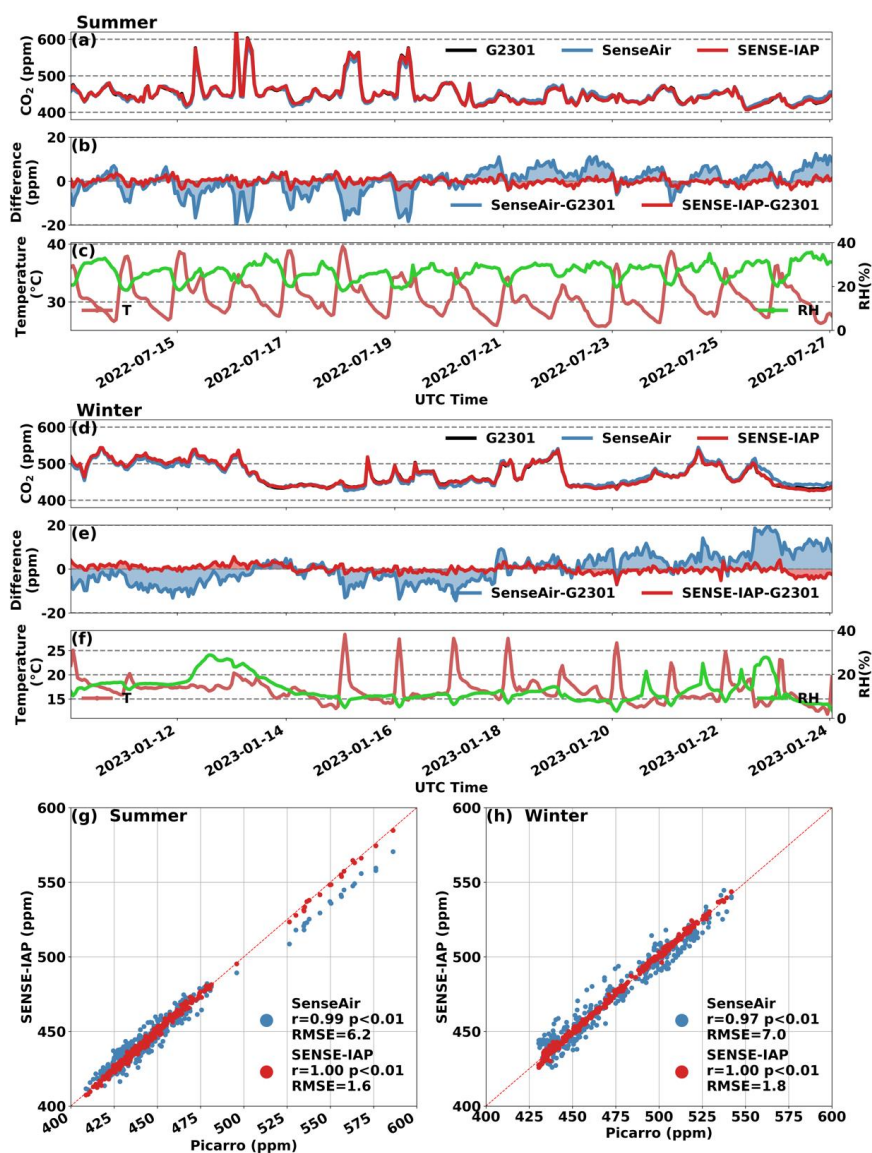
Figure 5: Comparison between hourly CO_2 concentrations measured by the SENSE-IAP and Picarro systems at the Beijing site from July 13th to 27th in 2022 (a-c, g) and January 10th to 24th in 2023 (d-f, h). We compared both the raw CO_2 data from SenseAir (blue) and the environmentally corrected data from SENSE-IAP (red) with that from Picarro (black). The T and RH were measured for the environment inside the instrument.

The correction system effectively adjusted the CO_2 concentration within the 400-700 ppm measurement range. As shown in Fig. 5b, during the period from July 16th to 19th, even when the ambient CO_2 concentration experienced significant fluctuations, the instrument showed high consistency with Picarro. The T and RH detected by the BME sensor were used to monitor the instrument's internal environment. Notably, in winter mornings, sunrise caused a significant temperature increase due to the presence of metal components on the circuit board (Fig. 5f). Our environmental correction successfully corrected for the temperature dependence, as the deviation between the SenseAir and Picarro showed a strongly correlated with temperature in both seasons (Fig. S1a). Additionally, the deviation of SenseAir relative to Picarro was significantly associated with RH in summer (Fig. S1b), and our correction system incorporated humidity, which was related to ambient temperature. Compared to the raw SenseAir data, the consistency of all six sensors improved markedly, with the RMSE decreasing from 5.0±1.0 ppm to 1.3±0.2 ppm in summer and from 6.8±0.8 ppm to 2.0±0.4 ppm in winter (Fig. S2).

Our correction system can preform environmental sensitivity analysis and correction on individual sensors, with the correction efficacy remaining robust across seasonal environmental changes. We further analyzed the daily RMSE of the SENSE-IAP relative to Picarro during 30 months of co-located observation. As shown in Fig. 6a, the daily RMSE of one sensor (pi688-K30) ranged from 1.5 to 4.0 ppm throughout the observation period, with the light blue shadow representing the monthly mean ± standard deviation. The average and median of daily RMSEs for this sensor were less than 2.0 ppm (Fig. 6b), except for an increase to higher than 2 ppm in summer (dark green). Compared to Picarro, the consistency of SENSE-IAP was better in winter, with an RMSE below 2.0 ppm on most days. Except for a few sensors exhibiting slightly higher RMSEs (approximately 3.0 ppm) in spring and autumn (pi732-K30), the daily RMSEs of the six sensors showed no significant seasonal variation (Fig. 6b).
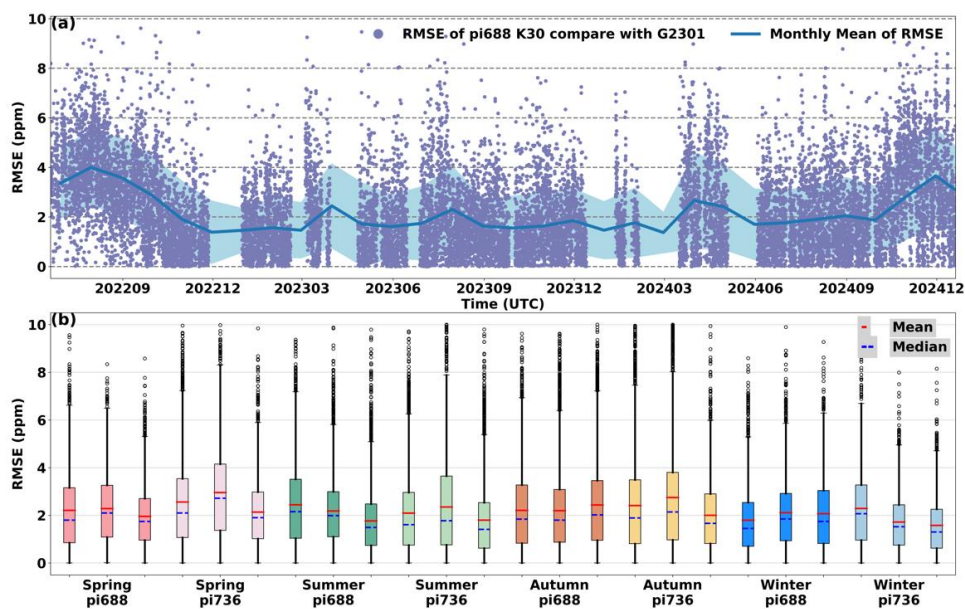
Figure 6. (a) The time series of daily RMSEs for hourly CO_2 concentration relative to Picarro (purple points) from June 2022 to Dec 2024, with a monthly rolling mean (blue line and shadow). (b) Box plot of the daily RMSEs of all six sensors across different seasons. Sensors from the same instrument are represented in the same colors (spring: dark pink/pink; summer: dark green/green; autumn: orange/yellow; and winter: dark blue/blue). Within each box, the red line indicates the mean value, while the blue dashed line represents the median values.

5. Performance of typical long-term drift and correction method

After environmental correction, the six SENSE-IAP sensors were co-located with Picarro synchronously for over 30 months. As shown in Fig. 7, two types of long-term drift were identified: 1) a downward drift trend and 2) a seasonal drift cycle. While the environmental correction system effectively corrects the impact of diurnal environmental changes, significant errors occurred in sensors due to the two types of long-term drift. Without the long-term drift calibration algorithm, the bias of SENSE-IAP could reach 27.9 ppm, with an RMSE of approximately 28.1 ppm (Fig. 7).

During the observation period, all six sensors exhibited long-term downward drifts ranging from 0.1 to 1.2 ppm per month (ppm/mo) (Fig. 7 and Table 1). Among the six sensors deployed in this study, only pi736 K30_3 showed a drift trend of less than 0.1 ppm/mo (Table 1). For sensors such as pi688-K30 and pi688-K30_2, the ΔCO_2 displayed a continuous downward trend over the 30 months, with slopes of 1.2 and 1.0 ppm/mo, respectively (Table 1). Notably, the drift trends for these sensors did not show significantly stabilization over time.

In addition to the downward drift trend, sensors like pi688-K30_3 and pi736-K30 exhibited varying seasonal cycle trends. After six months of deployment, these sensors showed RMSEs of 25.3 and 24.8, respectively (Fig. 7). However, after more than one year of observation, the impact of seasonal drift decreased, and the errors caused by long-term drift were highlighted, resulting in reduced RMSEs as 17.3 and 10 ppm, respectively.
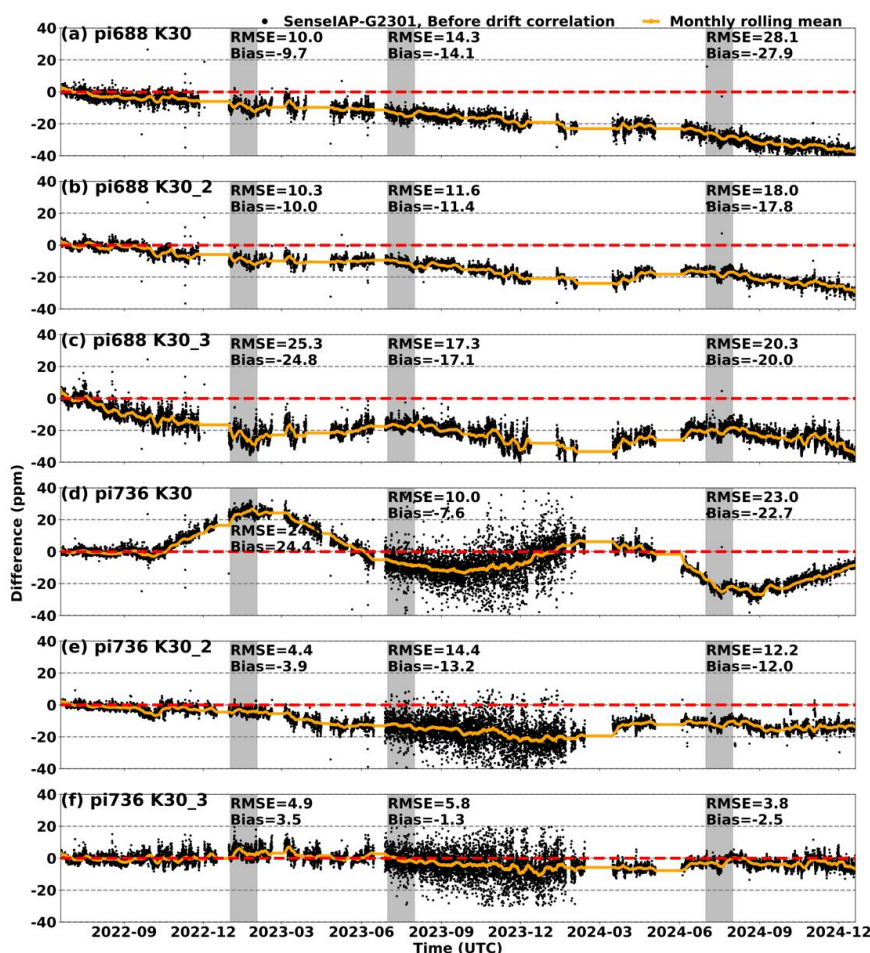
Figure 7: Time series of ΔCO_2 for the six sensors at the Beijing-IAP from June 2022 to Dec 2024 (black), with the monthly rolling mean (yellow line). The gray shadow represents the one-month time range used to evaluate the RMSE and bias for each sensor after half a year, one year, and two years of deployment.

From the perspective of drift magnitude, a significant bias of 5 ppm (approximately 1% of the ambient CO_2 concentration) typically occurred within 3-10 months after calibration, with most cases observed within 5 months. The seasonal drift cycle occurred on a six-month scale, with maximum errors typically occurring in winter and summer.

260 Therefore, we recommend that the long-term drift calibration frequency of SENSE-IAP should be no less than three months and no longer than six months. In addition, drift calibration should be performed at least once during both winter and summer seasons. If the target monitoring accuracy is within 3 ppm, the drift calibration frequency should be at least every two months, as a 3-ppm bias typically develops within 2-5 months.

265 Table 1. The Long-term drift trend of six sensors (Unit: ppm/mo)

SENSE-IAP	pi688			pi736		
Slope (ppm/mo)	s1	s2	s3	s1	s2	s3
drift in the first year	-1.2	-1.0	-1.4	-0.6	-1.1	-0.1
drift in the second year	-1.2	-0.5	-0.2	-1.3	0.1	-0.1

According to functions 6-7, we illustrate our long-term drift calibration method by focusing on the first year of observation. At the start of the observation period (Jun 2022), we adjusted the baseline of the six sensors. We designated the initial calibration time point as S_{cor} for the first observation period (Jun 2022 to Jan 2023). The inflection point of the drift trend in Feb 2023 was identified as E_{cor} for the first period and as S_{cor} for the second period of the observation (Jan 2023 to 270 Sep 2023). The E_{cor} for the second period was set to Sep 2023 in this study. We applied the linear calibration method between these two time points for both periods of the drift trend. After calibration, the CO_2 concentrations from the six sensors showed strong consistency with the Picarro, with an RMSE ranging from 2.4-3.0 ppm (Fig. 8).
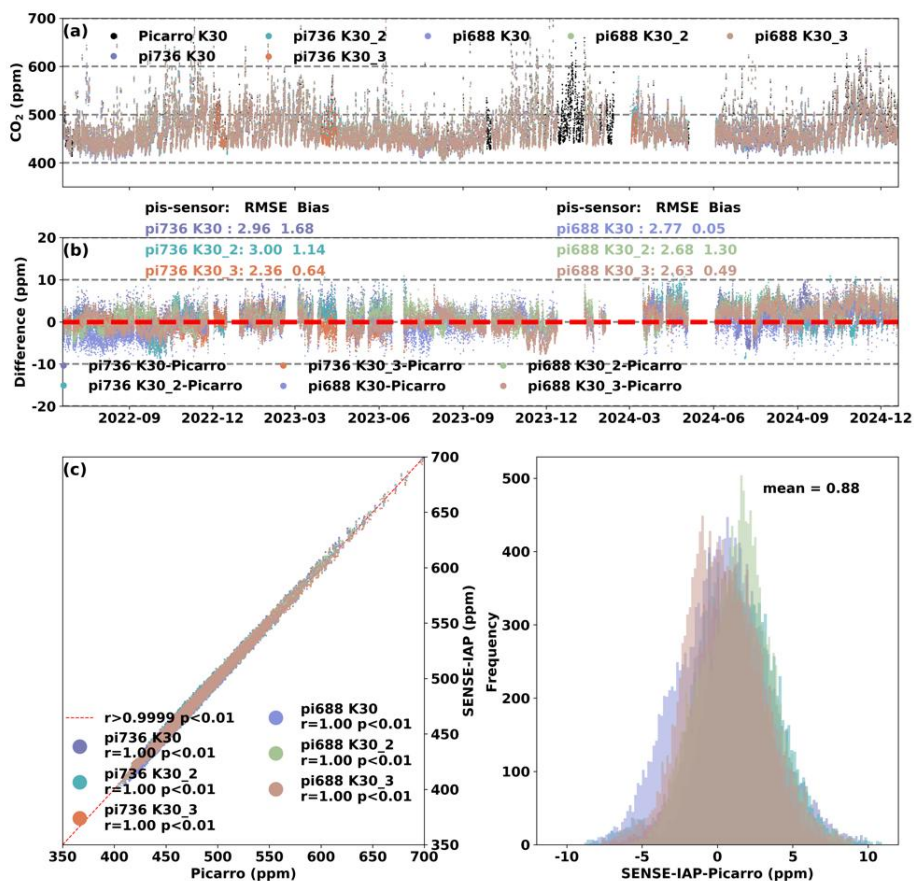
Figure 8: (a) Comparison of hourly CO_2 concentrations measured by six sensors and Picarro at Beijing- IAP from June 2022 to Dec 2024; (b) the time series of ΔCO_2; (c) scatter plot of SENSE-IAP and Picarro; (d) histogram plot of the ΔCO_2.

It should be noted that due to the loose connection of ventilation pipeline between Jul 2023 to Mar 2024, the CO_2 concentration measured by pi736 and Picarro were not strictly synchronized. This issue led to a relatively lower short-term monitoring accuracy during this period, primarily due to the lag effect caused by air diffusion. To assess the actual hourly monitoring accuracy of SENSE-IAP, we excluded data from this period. Without this exclusion, the RMSE for pi736 would have been 3.9-4.5 ppm (Fig. S3). However, to ensure the completeness and robustness of the long-term drift analysis, we retained samples from this period, allowing for a more comprehensive evaluation of drift trends over time.

6. Comparison of data quality across multiple levels over a long-term scale

As previously mentioned, the correction system effectively corrected the sensors' short-term environmental dependence. By applying a linear calibration method at least every three months, we can calibrate the time-dependent drift of LCSs and

resolve residual seasonal cycles that the environmental correction system cannot fully resolve. This seasonal variation is prominently reflected in the original electrical signal (raw signal). As shown in Figure S4(b), we eliminated the influence of short-term environment changes through a 24-hour running mean, revealing that the seasonal drift cycle in the raw signal can reach up to 100ppm, with an RMSE of 38.2 and a bias of -21.1 ppm (Table 2).

290   The defining characteristic of medium-precision low-cost sensors is the presence of long-term drift. This drift, which exhibits a downward trend, is observed in all sensors, with variations only in the drift rate (Fig.S5). Long-term drifts resulted in an RMSE of 15.8 and a bias of -12.0 for the LCS, respectively (Table 2). Although seasonal cycle does not significantly alter the overall trend of long-term drift, seasonal environmental variations can introduce errors of up to 25 ppm if baseline calibration are performed only annually (Fig. 7). Therefore, we recommend that baseline calibration be conducted at least

295   every six months, ideally during both winter and summer.

The data provided by the SenseAir manufacture were corrected for temperature sensitivity using the default temperature parameters. The long-term drift was calibrated using a so-called ABS algorithm, which employs periodic one-point calibration in SenseAir, assuming that the minimum value of CO_2 concentration is 400 ppm in fresh air (SenseAir-Corrected data). However, due to the constant assumption of fresh air concentration, coupled with the carbon absorption of vegetation

300   in summer and the higher emissions in winter, the SenseAir-Corrected data exhibit a fluctuating trend of approximately 20 ppm, with higher values in summer and lower values in winter. Consequently, the bias of SenseAir-Corrected data can be -3.3±1.4 ppm, with an RMSE of 12.1±2.0 ppm (Fig. S6, Table 2). In contrast, the drift-calibrated SENSE-IAP data demonstrate a much smaller bias (0.8±0.4 ppm) and an 80% improvement in accuracy, with an RMSE of 2.4±0.2 ppm (Fig.S7, Table 2).

305

Table 2: Evaluation parameters of the CO_2 concentration measured by K30 sensors compared to those from Picarro, including SenseAir-Corrected values, Raw signal, and SENSE-IAP at the Beijing-IAP from June 2022 to December 2024 (unit: ppm).

Data Type	SenseAir-Corrected		Raw signal		SENSE-IAP-Env-Corrected		SENSE-IAP-Env+Drift-Corrected	
Sensors	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias
Pi688 K30	15.9	-5.7	34.4	-20.2	20.1	-16.9	2.2	0.1
Pi688 K30_2	10.1	-3.1	35.1	-19.0	15.7	-13.6	2.4	1.3
Pi688 K30_3	10.2	-3.3	45.2	-24.2	21.5	-20.0	2.3	0.5
Pi736 K30	10.8	-1.1	22.0	-11.8	13.5	-3.5	2.6	1.1
Pi736 K30_2	13.0	-2.6	38.4	-23.0	14.3	-10.8	2.8	1.0
Pi736 K30_3	12.5	-4.2	53.9	-28.1	9.8	-7.0	2.2	0.7
Mean	12.1	-3.3	38.2	-21.1	15.8	-12.0	2.4	0.8
SD	2.0	1.4	9.9	-5.1	4.0	5.6	0.2	0.4

*According to the statistical results of 24-hour running means.

7. Seasonal drift cycle effects on SENSE-IAP

As shown in Fig. S8, the seasonal variations observed before instrument linear calibration were correlated with T, RH and P. The ΔCO_2 between pi736-K30 and Picarro significantly correlated with all three environmental factors, with r values of -0.58, -0.46, and 0.33 against T, RH and P, respectively. In contrast, the relationship between pi688-K30_3 and environmental factors was opposite to that of pi736-K30, with r values of 0.33, 0.5, and -0.6 against T, RH and P, respectively. Considering seasonal phase differences between CO_2 concentration changes and environmental factors, this seasonal deviation was likely attributable not only to insufficient environmental compensation but also to the influence of seasonal effects on the instrument's physical properties. For instance, changes in the sensor's optical cavity size caused by thermal expansion and cold contraction can change the optical path-lengths, subsequently affecting the pressure within the optical cavity and the strength of infrared CO_2 absorption (Yao et al., 2023). However, this hypothesis cannot fully explain why the two sensors, pi688 K30_3 and pi736 K30, exhibited opposite drift directions during the same season.

8. Conclusions

We evaluated low-cost NDIR CO_2 sensors using Picarro as a reference instrument. Our environmental correction system effectively corrected the impact of short-term daily environmental changes by assigning unique environmental

sensitivity parameters to each sensor. This approach reduces the short-term RMSE from 5.9±1.2 ppm for SenseAir to

325   1.6±0.5 ppm for SENSE-IAP. The correction system demonstrates robustness against seasonal environmental variations,

maintaining a daily RMSE of 1-3 ppm.

Based on a 30-month observation, we recommend that the calibration frequency for long-term drifts not exceed six

months. For optimal performance and to ensure the target monitoring accuracy remains within 1% of the ambient CO_2

concentration, a three-month calibration interval is recommended. If standard instruments, standard gases (which are

330   generally easier to obtain), or other reliable concentration references such as model simulations are available, long-term drift

can be linearly corrected at the seasonal scale.

Consequently, after deployment, even with significant environmental changes around the instrument, there is no need to

frequently bring the instruments back to the laboratory for re-correction of environmental impacts. After the long-term drift

calibration, the RMSE of SENSE-IAP remains 2.4 ± 0.2 ppm even after 30 months of operation. This performance enables

335   long-term deployment of the instruments, significantly reducing the maintenance costs associated with LCS.

Supporting information

The supplementary information is available online.

Data availability statement

340   The data used to generate the figures in this manuscript are available upon reasonable request to the corresponding authors.

Competing interests

The authors declare that they have no conflicts of interest.

Funding

This research was supported by the National Key R&D Program of China (No. 2023YFC3705500 and 2017YFB0504000);

345   Jinan Carbon Monitoring and Evaluation Pilot Project (grant no. SDGP370100000202202001740); the Qiluzhongke Institute

of Carbon Neutrality Program of Jinan Dual Carbon Simulator; and the CAS Proof of Concept Program: Carbon

Neutrality-oriented Urban Carbon Monitoring System and Its Industrialization (grant no. CAS-GNYZ-2022).

Author contributions

NZ, PFH and QXC conceived and designed the study. QXC and PFH collected and analyzed the datasets. XYY, CX, and

350   ZJW discussed the sensor results. QXC led the writing of the paper with contributions from all the coauthors. All coauthors

contributed to the descriptions and discussions of the manuscript.

Acknowledgments

We thank Mr. Cory Martin, Mr. Di Liu, Mr. Yinan Wang, Mr. Ming Cui, Mr. Jin Guo, Mr. Zhimin Zhang, and Mr. Yang Zi

for their help in the SENSE-IAP instrument development, calibration and deployments.

355

References

SenseAir: LP8 Product Sheet: https://rmtplusstoragesenseair.blob.core.windows.net/docs/Dev/publicerat/PSH1233.pdf, last access: 6 March 2019.

SenseAir: K30 products sheets: https://rmtplusstoragesenseair.blob.core.windows.net/docs/publicerat/PSP110.pdf, last access: 6 March 2022.

Picarro G2301 Analyzer Datasheet: https://www.picarro.com/environmental/support/library/documents/g2301_analyzer_datasheet, last access: 30 January 2023.

BAO, Z., HAN, P., ZENG, N., LIU, D., CAI, Q., WANG, Y., TANG, G., ZHENG, K., and YAO, B.: Observation and modeling of vertical carbon dioxide distribution in a heavily polluted suburban environment, Atmospheric Ocean. Sci. Lett., 13, 371–379, https://doi.org/10.1080/16742834.2020.1746627, 2020.

Bigi, A., Mueller, M., Grange, S. K., Ghermandi, G., and Hueglin, C.: Performance of NO, NO_2 low cost sensors and three calibration approaches within a real world application, Atmospheric Meas. Tech., 11, 3717–3735, https://doi.org/10.5194/amt-11-3717-2018, 2018.

Cai, Q., Han, P., Pan, G., Xu, C., Yang, X., Xu, H., Ruan, D., and Zeng, N.: Evaluation of Low-Cost CO_2 Sensors Using Reference Instruments and Standard Gases for Indoor Use, Sensors, 24, https://doi.org/10.3390/s24092680, 2024.

Cai, Q., Wang, Z., Han, P., Zeng, N., Nie, X., Yang, X., Wang, Z., and Gao, S.: The Enhancement and Variation in the Carbon Dioxide Concentration in a Typical Industrial Park in the Northern China, Aerosol Sci. Eng., https://doi.org/10.1007/s41810-025-00286-4, 2025.

Delaria, E. R., Kim, J., Fitzmaurice, H. L., Newman, C., Wooldridge, P. J., Worthington, K., and Cohen, R. C.: The Berkeley Environmental Air-quality and CO_2 Network: field calibrations of sensor temperature dependence and assessment of network scale CO_2 accuracy, Atmospheric Meas. Tech., 14, 5487–5500, https://doi.org/10.5194/amt-14-5487-2021, 2021.

Grimmond, C. S. B., King, T. S., Cropley, F. D., Nowak, D. J., and Souch, C.: Local-scale fluxes of carbon dioxide in urban environments: methodological challenges and results from Chicago, Environ. Pollut., 116, S243–S254, https://doi.org/10.1016/S0269-7491(01)00256-1, 2002.

Gurney, K. R., Razlivanov, I., Song, Y., Zhou, Y., Benes, B., and Abdul-Massih, M.: Quantification of Fossil Fuel CO_2 Emissions on the Building/Street Scale for a Large U.S. City, Environ. Sci. Technol., 46, 12194–12202, https://doi.org/10.1021/es3011282, 2012.

Hagan, D. H., Isaacman-VanWertz, G., Franklin, J. P., Wallace, L. M. M., Kocar, B. D., Heald, C. L., and Kroll, J. H.: Calibration and assessment of electrochemical air quality sensors by co-location with regulatory-grade instruments, Atmospheric Meas. Tech., 11, 315–328, https://doi.org/10.5194/amt-11-315-2018, 2018.

Han, P., Yao, B., Cai, Q., Chen, H., Sun, W., Liang, M., Zhang, X., Zhao, M., Martin, C., Liu, Z., Ye, H., Wang, P., Li, Y., and Zeng, N.: Support Carbon Neutral Goal with a High-Resolution Carbon Monitoring System in Beijing, Bull. Am. Meteorol. Soc., 105, E2461–E2481, https://doi.org/10.1175/BAMS-D-23-0025.1, 2024.

Kellett, R., Christen, A., Coops, N. C., van der Laan, M., Crawford, B., Tooke, T. R., and Olchovski, I.: A systems approach to carbon cycling and emissions modeling at an urban neighborhood scale, Landsc. Urban Plan., 110, 48–58, https://doi.org/10.1016/j.landurbplan.2012.10.002, 2013.

Lateb, M., Meroney, R. N., Yataghene, M., Fellouah, H., Saleh, F., and Boufadel, M. C.: On the use of numerical modelling for near-field pollutant dispersion in urban environments–A review., Environ. Pollut., 208 Pt A, 271–283, 2016.

Liu, D., Sun, W., Zeng, N., Han, P., Yao, B., Liu, Z., Wang, P., Zheng, K., Mei, H., and Cai, Q.: Observed decreases in on-road CO_2 concentrations in Beijing during COVID-19 restrictions, Atmospheric Chem. Phys., 21, 4599–4614, https://doi.org/10.5194/acp-21-4599-2021, 2021.

Lopez-Coto, I., Ghosh, S., Prasad, K. R., and Whetstone, J.: Tower-based greenhouse gas measurement network design—The National Institute of Standards and Technology North East Corridor Testbed, Adv. Atmospheric Sci., 34, 1095–1105, 2017.

Martin, C. R., Zeng, N., Karion, A., Dickerson, R. R., Ren, X., Turpie, B. N., and Weber, K. J.: Evaluation and environmental correction of ambient CO_2 measurements from a low-cost NDIR sensor, Atmospheric Meas. Tech., 10, 2383–2395, https://doi.org/10.5194/amt-10-2383-2017, 2017.

405 Miles, N. L., Davis, K. J., Richardson, S. J., Lauvaux, T., Martins, D. K., Deng, A. J., Balashov, N., Gurney, K. R., Liang, J., Roest, G., Wang, J. A., and Turnbull, J. C.: The influence of near-field fluxes on seasonal carbon dioxide enhancements: results from the Indianapolis Flux Experiment (INFLUX), Carbon Balance Manag., 16, 4, https://doi.org/10.1186/s13021-020-00166-z, 2021.

Müller, M., Graf, P., Meyer, J., Pentina, A., Brunner, D., Perez-Cruz, F., Hüglin, C., and Emmenegger, L.: Integration and calibration of non-dispersive infrared (NDIR) CO2 low-cost sensors and their operation in a sensor network covering

410 Switzerland, Atmospheric Meas. Tech., 13, 3815–3834, https://doi.org/10.5194/amt-13-3815-2020, 2020.

Shusterman, A. A., Kim, J., Lieschke, K. J., Newman, C., Wooldridge, P. J., and Cohen, R. C.: Observing local CO2 sources using low-cost, near-surface urban monitors, Atmospheric Chem. Phys., 18, 13773–13785, https://doi.org/10.5194/acp-18-13773-2018, 2018.

Turner, A. J., Shusterman, A. A., McDonald, B. C., Teige, V., Harley, R. A., and Cohen, R. C.: Network design for

415 quantifying urban CO_2 emissions: assessing trade-offs between precision and network density, Atmospheric Chem. Phys., 16, 13465–13475, https://doi.org/10.5194/acp-16-13465-2016, 2016.

Vaisala: GMP343 Carbon Dioxide Probe Datasheet, https://docs.vaisala.com/v/u/B210688EN-H/en-US, last access: 6 March 2020.

Wada, R., Pearce, J. K., Nakayama, T., Matsumi, Y., Hiyama, T., Inoue, G., and Shibata, T.: Observation of carbon and

420 oxygen isotopic compositions of CO2 at an urban site in Nagoya using Mid-IR laser absorption spectroscopy, Atmos. Environ., 45, 1168–1174, https://doi.org/10.1016/j.atmosenv.2010.10.015, 2011.

Wu, L., Broquet, G., Ciais, P., Bellassen, V., Vogel, F., Chevallier, F., Xueref-Remy, I., and Wang, Y.: What would dense atmospheric observation networks bring to the quantification of city CO_2 emissions?, Atmospheric Chem. Phys., 16, 7743–7771, https://doi.org/10.5194/acp-16-7743-2016, 2016.

425 Yang, Y., Zhou, M., Wang, T., Yao, B., Han, P., Ji, D., Zhou, W., Sun, Y., Wang, G., and Wang, P.: Spatial and temporal variations of CO2 mole fractions observed at Beijing, Xianghe, and Xinglong in North China, Atmospheric Chem. Phys., 21, 11741–11757, https://doi.org/10.5194/acp-21-11741-2021, 2021.

Yao, G., Li, Y., Shang, Q., and Fan, H.: Research on Temperature Compensation of Optical Fiber MEMS Pressure Sensor Based on Conversion Method, Photonics, 10, https://doi.org/10.3390/photonics10010022, 2023.

430 Zeng, N., Han, P., Liu, Z., Liu, D., Oda, T., Martin, C. R., Liu, Z., Yao, B., Sun, W., Wang, P. C., Cai, Q., Dickerson, R., and Maksyutov, S.: Global to local impacts on atmospheric CO2 from COVID-19 lockdown, biosphere and weather variabilities, Environ. Res. Lett., 17, 2021.