



SCOUT: Munich Urban Street-level CO₂ Sensor Network

Daniel Kühbacher¹, Jia Chen¹, Patrick Aigner¹, Adrian Wenzel¹, Pascal Rubli², Dominik Brunner², Lukas Emmenegger², and Stuart K. Grange^{2,3}

¹Professorship of Environmental Sensing and Modeling, Technical University of Munich (TUM), Theresienstr. 90, 80333 Munich, Germany

²Laboratory for Air Pollution/Environmental Technology, Swiss Federal Laboratories for Materials Science and Technology (Empa), Überlandstr. 129, 8600 Dübendorf, Switzerland

³now at Climate and Environmental Physics, Physics Institute, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland

Correspondence: Daniel Kühbacher (daniel.kuehbacher@tum.de) and Jia Chen (jia.chen@tum.de)

Abstract. The city of Munich has been a pilot city in the ICOS Cities project, which aimed to develop greenhouse gas measurement methodologies and related services for urban areas to support climate action. A part of Munich's city observatory is SCOUT (Street-level Carbon Observatory for Urban Terrain), a network of 100 battery-powered, low-cost, non-dispersive infrared (NDIR) CO₂ sensors. This publication describes the sensor network, the applied data processing strategy, and assesses sensor performance. Data from May 2023 to December 2025 are covered. We propose a novel sensor calibration approach that accounts for the strong, non-stationary temperature dependence of the raw sensor signal, thereby improving the manufacturer's measurement accuracy by a factor of four to eight. In the first stage, a rolling ordinary least squares regression is used to estimate the time-varying, temperature-induced signal component. This is followed by a multilinear regression that infers the final CO₂ concentration and corrects for humidity and pressure effects. Evaluated across 21 sensors over three to twelve months outside the training time period at a reference site equipped with a calibrated Picarro analyzer, the proposed model achieves a median RMSE of 9.46 [6.1-12.7] ppm and a R² of 0.76 [0.55-0.89] for hourly averaged values. It also shows minimal residual dependence on environmental parameters. Long-term drift is corrected by relating the sensor signal to nearby reference stations during favorable atmospheric conditions with homogeneous CO₂ concentration across the city. We consider three distinct correction scenarios and estimate a city-wide, location-dependent correction uncertainty of σ between 2.3 and 3.1 ppm. Combined with the measurement uncertainty, this yields a total uncertainty of 9.95 [6.4-13.1] ppm. We deployed the sensor systems in pairs and present a strategy to combine the signals from duplicate measurements at the same site. The network can characterize local emission regimes: winter morning CO₂ enhancements show a high correlation with Munich's gridded 2024 emission inventory ($r = 0.79$), indicating that street-level low-cost networks can serve as proxies for local emissions.

Keywords. CO₂ monitoring, low-cost sensor network, urban emissions, NDIR sensor calibration, ICOS Cities

1 Introduction

Anthropogenic climate change, primarily caused by greenhouse gas (GHG) emissions from fossil fuel combustion, industrial processes, and land-use change, is one of the greatest challenges facing humanity. Human-caused global warming reached 1.36



°C in 2024 compared to 1850–1900 (Forster et al., 2025), and the remaining carbon budget to limit global warming to below 1.5 °C is expected to run out within the next four years at the current rate of global emissions (Friedlingstein et al., 2025).
 25 Continuing current policies are likely to result in global warming of 2.8 °C (2.1–3.9 °C), and even if all conditional nationally determined contributions (NDCs) and net-zero targets are met, an average warming of 1.9 °C (1.8–2.3 °C) must be expected (UNEP, 2025).

Cities play a crucial role in both exacerbating and mitigating this threat. A total share of 67 to 72 % of global emissions, including direct and indirect emissions from the production of electricity, goods, and services that are consumed in cities, can be attributed to urban areas (IPCC, 2023). City administrations worldwide are implementing climate protection strategies to reduce CO₂ emissions, often with more ambitious targets than national entities (Matsumoto et al., 2019). Measurement, reporting, and verification capabilities are vital in guiding these efforts, aiding in establishing targeted strategies, tracking progress, and assessing the outcome. Urban atmospheric CO₂ measurements provide additional constraints on direct carbon emissions and their spatial and temporal variability, thereby supporting urban climate action planning (Lian et al., 2023).
 35 Measurements provide valuable near-real-time insights, whereas accurate data required to assemble emission inventories are often only available with a multi-year time delay (Ke et al., 2023; Hu et al., 2025).

The ICOS Cities PAUL project (Pilot Application in Urban Landscapes) developed frameworks to measure and model GHG emissions in urban areas and to deliver related services (ICOS, 2025). Various GHG measurement techniques have been tested in three pilot cities: Paris, Munich, and Zurich. In Munich, the total-column network MUCCnet (Dietrich et al., 2021) has been operating since 2019 and has now been complemented by the ACROPOLIS mid-cost sensor network (Aigner et al., 2026a), eddy covariance (EC) and relaxed eddy-accumulation (REA) measurements from a central tall tower (Kunz et al., 2025; Ecosystem Thematic Centre et al., 2026), and biogenic measurements (Li et al., 2026b, 2025a, b; Lan et al., 2025). The monitoring networks MUCCnet and ACROPOLIS are designed to measure CO₂ concentrations at the rooftop level, where the signal is comparatively well mixed and representative of a broader footprint, making them suited for trend analysis and inverse
 45 modeling (Li et al., 2026a).

This study adds another component to the city observatory: SCOUT (Street-level Carbon Observatory for Urban Terrain). Street-level CO₂ measurements provide a near-surface perspective influenced by local emission sources, including traffic, residential heating, and biogenic processes. The SCOUT network comprises 100 commercial, battery-powered, low-cost CO₂ sensors (Decentlab DL-LP8P), which were first deployed in a scientific network by Müller et al. (2020) in the Swiss Carbosense project and subsequently deployed in a dense network in Zurich (ZiCOs-L) (Creman et al., 2026). The built-in sensor Senseair LP8 costs around 60 € in retail, while its successor Senseair Sunrise is available for less than 40 €. Combined with supporting electronics for power supply and data communication, the sensor system is orders of magnitude less expensive than the GHG analyzers used in traditional atmospheric measurement networks, enabling dense deployments with negligible installation effort and minimal site requirements. The sensor systems can easily be co-located with street-level air quality sensors or traffic detectors, enabling extended correlation analysis and site-to-site comparisons. Nevertheless, the manufacturer-specified accuracy is typically ± 50 ppm (Senseair AB), rendering the sensors unsuitable for urban outdoor measurements without careful calibration and data processing.



The performance potential of low-cost NDIR CO₂ sensors has been investigated in several studies, and various calibration methods have demonstrated significant improvements in accuracy. Martin et al. (2017) applied multiple linear regression to six Senseair K30 sensors trained and evaluated during a four-week outdoor side-by-side campaign, while Dubey et al. (2024) assessed the performance of three sensors and six different sensor models during a three-month outdoor deployment. Both studies demonstrated that individual sensor calibration can substantially improve the accuracy. However, these investigations were limited in spatial extent and duration and do not address the challenges associated with large-scale, multi-year deployments. Cai et al. (2025) extended the evaluation period to 30 months and found that environmental correction alone reduced the RMSE from 5.9 to 1.6 ppm for a Senseair K30 sensor, but long-term drift necessitates regular recalibration. Unlike the Senseair LP8, which uses pulsed light to drastically reduce energy consumption, the K30 uses a continuous light source, making it more stable but unsuitable for battery-powered devices. The Senseair LP8 was used by Müller et al. (2020) and Creman et al. (2026), who both proposed individually calibrated linear sensor models combined with comprehensive approaches for drift correction and outlier removal. Their pioneering large-scale deployments in Switzerland and Zurich highlighted key challenges of long-term outdoor operation, including evolving response characteristics, gradual and abrupt sensor drift, and humidity- and temperature-induced artifacts.

Building on this, we introduce a novel, time-adaptive calibration model that explicitly estimates the sensor's temperature-induced signal component. The dependence arises from multiple factors, including temperature-dependent detector sensitivity, variations in the infrared light source, thermal stress that induces mechanical changes in low-cost optical components, and degradation of optical performance over time. This produces a non-stationary temperature response that static calibrations fail to capture. To this end, our model uses rolling ordinary least squares regression (ROLS) to estimate the temperature-induced signal component in a first stage, followed by a multilinear regression to correct for humidity and pressure effects and to retrieve CO₂ mole fractions. We further present a multi-scenario drift-correction strategy to maintain stable absolute measurements over time. Paired sensor deployments at each site are leveraged for automated fault detection and outlier identification, and the effect of a solar radiation shield on measurement accuracy is evaluated under real-world conditions. Finally, we demonstrate the network's capability to assess local emission patterns.

2 Methods

2.1 Sensor system and data transmission infrastructure

The sensor node, Decentlab DL-LP8P, is built around the Senseair LP8 non-dispersive infrared (NDIR) CO₂ sensor. It is complemented by a Sensirion SHT21 for temperature and humidity measurements, as well as a Bosch BMP280 for atmospheric pressure sensing. The added pressure sensor is the only significant difference compared to the sensors in the Carbosense and ZiCOs-L networks. The LP8 uses an indium antimonide (InSb) detector with a built-in optical band-pass filter (Asahi Kasei Microdevices Corporation (AKM), 2016) and an infrared light source manufactured by OSHINO Lamps. Figure 1 shows a complete and a disassembled CO₂-Sensor which consists of three parts, a circuit board hosting the detector, lamp, and electronic components, and a two-piece optical cell. The air to be measured diffuses through a filter membrane behind the light

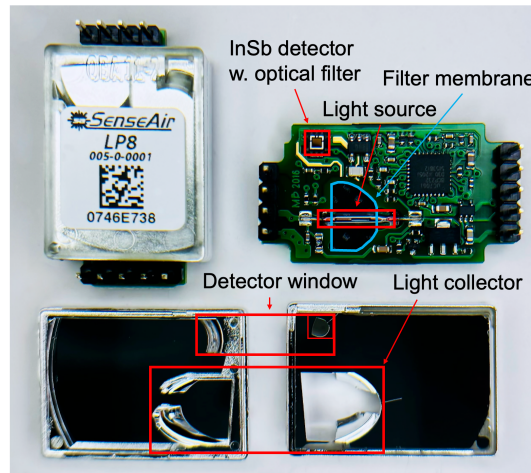


Figure 1. Disassembled Senseair LP8 NDIR CO₂ sensor. The main parts are a mid-infrared InSb detector with an optical bandpass filter, a broadband light source, and a reflective optical cell. Ambient air diffuses into the cell through a filter membrane located below the light source. The top left shows the complete sensor, comprising the circuit board and the optical assembly. The top-right shows the circuit board, and the bottom view depicts the two parts of the optical assembly with the light collector and detector window facing downward.

source into the measurement cell. The Senseair LP8 sensor is highly energy-efficient, allowing the node to run on two alkaline C batteries (LR14, Varta Longlife Power, 1.5 V/ 7800 mAh) for between 2.4 and 5.1 years, depending on the strength of the network signal (Decentlab GmbH, 2018).

The sensor node measures CO₂ and the environmental parameters at 1-minute intervals and transmits 10-minute averages
95 via a public LoRaWAN network. LoRaWAN is a particularly energy-efficient communication network protocol designed to connect battery-powered devices over a long range (up to 15 km) (LoRa Alliance®, 2025). We connected 77 of our nodes to The Things Network¹ (TTN), a free, community-based LoRa network that offers good coverage in Munich. At locations with poor TTN reception, we switched to the paid Helium network², which costs approximately 0.5 € per node per year. Data packets are forwarded to an InfluxDB hosted by the device manufacturer Decentlab, where the data are stored permanently. On
100 their platform, it can be visualized as dashboards and directly accessed via an API. The metadata, including the deployment location, configuration used, and any relevant additional comments, is stored in JSON-formatted files within a version-controlled repository on GitHub.

2.2 Sensor characterization and network deployment

Prior to deployment in Munich, all sensor nodes were placed in a climate chamber, where varying temperatures, humidity, and
105 CO₂ levels were set, along with a calibrated reference analyzer (Picarro G2401). These tests were conducted in collaboration with Empa using identical procedures as Creman et al. (2026) and yielding the same conclusions: data from controlled tests

¹<https://www.thethingsnetwork.org/>

²<https://www.helium.com/iot>



(a) Individual deployment on lamp post (b) Co-location w. low-cost air quality sensor

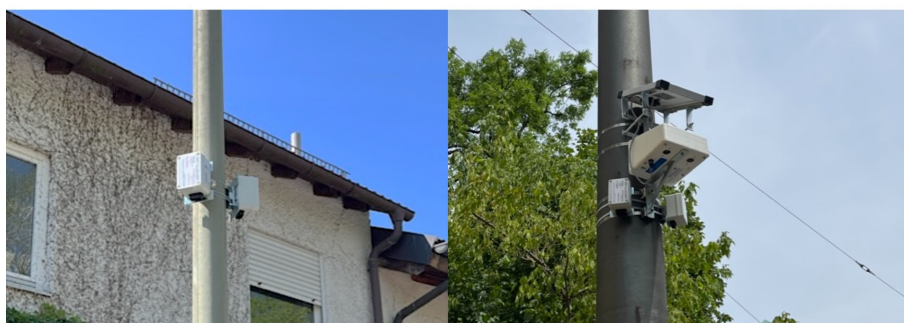


Figure 2. The sensors were mounted on lamp posts at a height of 3 m above ground level in Munich. They were deployed in pairs at independent sites (a) or co-located with other instruments, such as low-cost air-quality sensors (b).

provide valuable insights into sensor behavior. However, an outdoor side-by-side campaign with a reference analyzer is more effective for calibrating a sensor model. Following the chamber tests, side-by-side measurements were then carried out under field conditions for three months at the Empa-Dübendorf reference air-quality monitoring site, which is equipped with a calibrated Picarro G2301 analyzer. An additional month of measurements was performed using a calibrated Picarro G2301 at a rooftop site at the Technical University of Munich (TUMR). This reference data are used to optimize the sensor model parameters for each sensor. The calibrations of the reference analyzers were performed against the WMO CO₂ X2019 calibration scale (Hall et al., 2021).

Field deployment started in May 2023, and the sensors were mounted on streetlamp posts in pairs enabling fault detection and quality control (Fig. 2). For many sites, we opted for co-location with additional measurement devices, like traffic detectors (8 sites), governmental air quality stations (5 sites), low-cost air quality sensors (7 sites), rooftop mounted, calibrated mid-cost CO₂ sensors (Aigner et al., 2026a) (6 sites), a high-precision CO₂ analyzer (Picarro G2301, 1 site), and a meteo station (1 site). For the remaining 21 sites, we chose locations with different local characteristics (e.g., traffic-dominated, urban/suburban, and biogenic-dominated). Figure 3 shows the location of all selected sites in Munich.

2.2.1 Installation of solar radiation shields

During the first summer in the field (June to September 2023), several sensors recorded temperatures significantly higher than ambient air temperature. This discrepancy was caused by solar radiation, leading to elevated temperatures inside the enclosures relative to the surrounding environment. Furthermore, different sensor orientations led to temperature variations between the two units deployed as a pair at the same location. To mitigate these effects, sensors at 35 sites were upgraded with solar radiation shields. We used an Electro-Mation EMF 8001 weather and radiation protection shield with an internal diameter of 110 mm and a height of 210 mm (ELECTRO - MATION GmbH, o.D.). It fits two sensor nodes into a self-developed internal bracket, using only the circuit board and removing the manufacturer's casing. Figure 4 shows the radiation shield and the internal bracket. An assessment of the radiation shield's effect is presented in section 3.5.

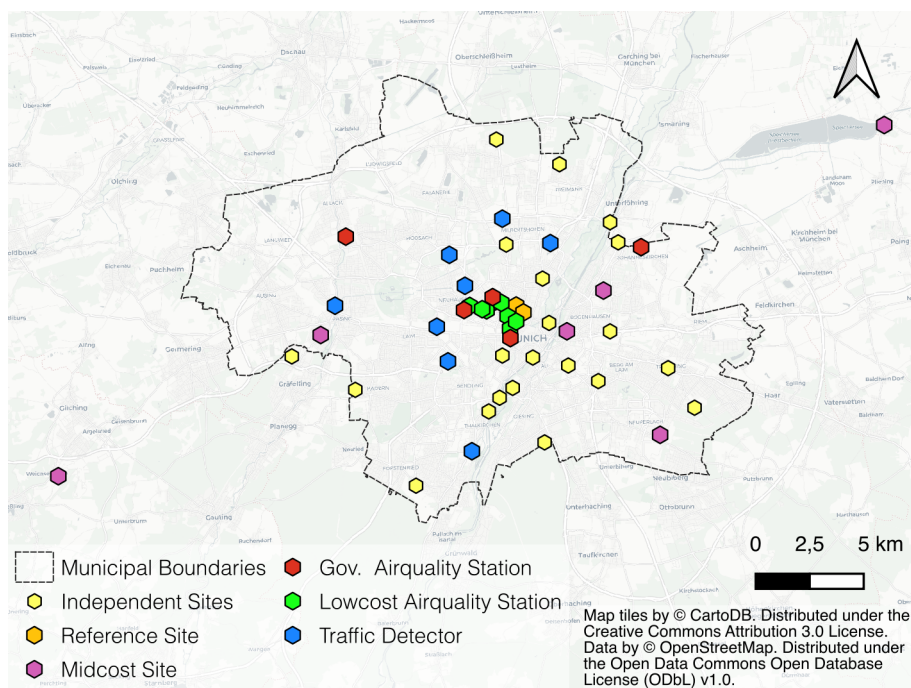


Figure 3. Map of the city of Munich showing the municipal boundaries and the locations of the SCOUT sensor sites. The nodes are densely deployed in the city center, and many are co-located with other measurement devices, such as traffic detectors, governmental and low-cost air-quality stations, as well as calibrated mid-cost CO₂ sensors.

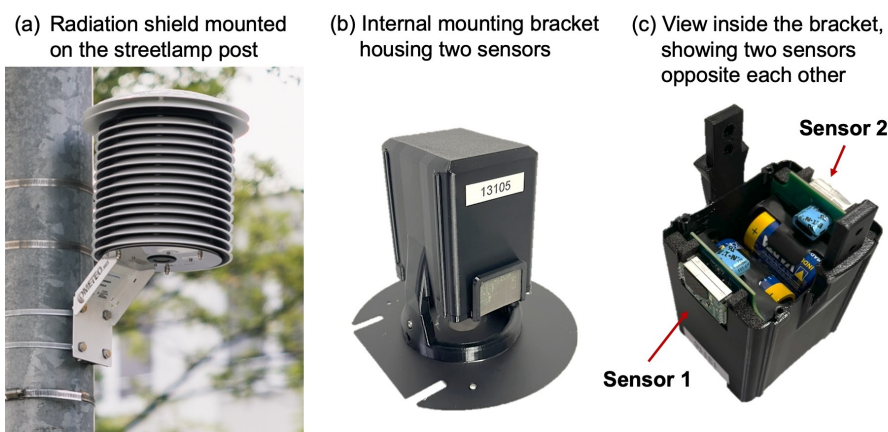


Figure 4. To avoid excess heat, sensors at 35 sites were mounted in pairs inside a solar radiation shield (a) (image source: ICOS RI). A custom-designed, 3D-printed bracket fits two nodes inside the shield. Image (b) shows the fully assembled configuration. Image (c) provides a view inside the bracket, with two sensor nodes oriented in opposite directions.

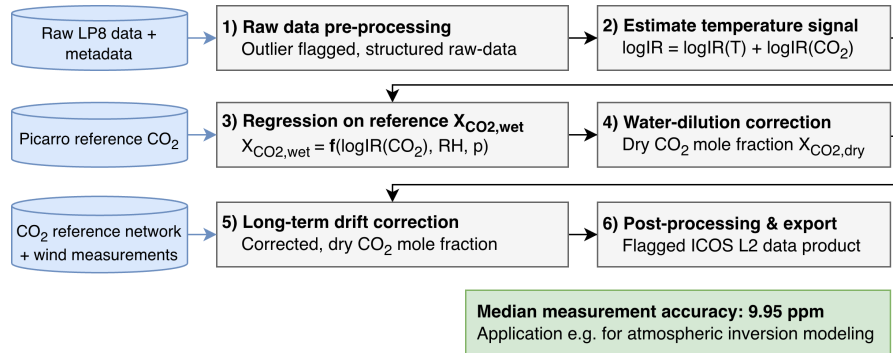


Figure 5. Overview of processing pipeline applied to the sensor data. The six steps are sequential and explained in detail in the subsequent section.

2.3 Data processing pipeline

130 The low-cost sensors are not placed in a temperature- or humidity-controlled environment, but are exposed to changing outdoor conditions. These have a major impact on the measurement performance and data quality, which must be handled through careful data processing. The process starts with downloading the raw data into a local database, followed by offline processing steps outlined below, and concludes with uploading the processed data to the ICOS Carbon Portal. Figure 5 provides an overview of the sequential processing steps applied. We developed a calibration model based on the sensor's log-transformed

135 raw IR measurements (log IR) rather than on the CO₂ concentration reported by the sensor's firmware, thereby establishing a physical relationship with gas absorption via the Beer-Lambert law. The LP8 sensor uses a pulsed light source with correlated double sampling: the raw IR signal, which is defined as the difference between the illuminated and dark detector readings, decreases with increasing CO₂ number density in the optical path (Appendix A). The raw IR signal is predominantly dependent on the ambient temperature, governed by the temperature-dependent sensitivity of the infrared detector. In addition, thermal

140 stress can induce mechanical changes in the low-cost sensor optics. In addition, the sensor characteristics can change during field use through drift in the electronics or the buildup of contamination on the low-cost optics. These effects contribute to drift in the raw IR signal. Achieving robust performance across seasons and varying temperature conditions would ideally require a side-by-side reference period that captures the full range of possible ambient variability. In practice, this is rarely justifiable given the significant time and effort required, underscoring the need for a calibration model with good generalization

145 capabilities.

2.3.1 Step 1: Raw data pre-processing

The sensor system transmits 10-minute averaged measurements to the vendor's data platform at nominal 10-minute intervals. Due to LoRaWAN's ALOHA-based MAC protocol (Augustin et al., 2016), transmissions occur at non-deterministic times, leading to irregular data arrival. If the data cannot be sent or received, it is not buffered but lost, and the sensor starts a new



150 measurement cycle. Therefore, after downloading the data via the corresponding API, the raw data were resampled to a regular 10-minute interval to enable convenient synchronization with other datasets (e.g., reference and meteorological data). Bins with no observations at this 10-minute resolution were treated as missing values and were dropped. Then, an outlier-detection algorithm proposed by Müller et al. (2020) was applied to flag outliers solely based on the sensor's raw data. We refer to this algorithm as the Müller-Filter hereafter. It considers the typical behavior of individual sensors in the field to derive multiple
 155 statistical measures for outlier identification. For this, the detector's raw infrared measurements, along with temperature and humidity measurements, are required. Data loss and outlier statistics are provided in section 3.1.

2.3.2 Step 2: Estimating the temperature-induced signal component

The time-varying, temperature-dependent component of the IR signal $\log \text{IR}(T)$ is estimated using a rolling-window ordinary least squares regression (Statsmodels RollingOLS (Seabold and Perktold, 2010)) against concurrent temperature measure-
 160 ments. The raw IR signal is dominated by temperature ($r > 0.95$; Fig. 6a), while the correlation of the IR signal with the CO_2 mole fraction is weak ($r < 0.3$). Accurately describing the temperature dependence is critical for isolating the residual dependence on CO_2 , which is ultimately the signal of interest. Figure 6b depicts the strong relationship between the raw, log-transformed IR signal and the ambient temperature. It also indicates the drift in temperature response characteristics between the training phase and the subsequent deployment phase, which must be taken into account.

165 For this calculation, the data are first filtered for outliers using the Müller-Filter. Observations with relative humidity above 94 % are additionally excluded to avoid high humidity affecting the IR signal, likely through condensation in the optical cell (Müller et al., 2020). We apply a rolling window of $m = 4000$ subsequent samples (≈ 28 days). If values are missing, at least 80 % of the window length must be present ($\text{min}_{\text{obs}} = 4000 \cdot 0.8 = 3200$). Otherwise, the window is expanded.

For each time index $t \geq m$, the model:

$$170 \quad \log \text{IR}_i = \mathbf{x}_i \beta_t + e_i = \log \text{IR}(T)_i + e_i, \quad \text{where} \quad \mathbf{x}_i = [1, T_i, T_i^2]^\top \quad (1)$$

is fitted to the most recent m observations ($i = t - m + 1, \dots, t$), with β_t obtained via the least squares optimization. Subtracting this explicitly modeled temperature-induced component from the raw IR measurement yields a residual e_i that primarily correlates with CO_2 variations ($r \approx 0.75$):

$$e_i = \log \text{IR}_i - \log \text{IR}(T)_i = \log \text{IR}(\text{CO}_2)_i \quad (2)$$

175 Although CO_2 and temperature are not statistically independent, particularly at diurnal timescales, the omitted variable bias from excluding the CO_2 term is negligible for a rolling window regression spanning multiple weeks. Therefore, a long time window is generally preferred, as it reduces the correlation between temperature and CO_2 . Furthermore, CO_2 trends at those timescales are corrupted by drift and unrecoverable from the sensor signal alone and therefore delegated to the drift correction stage (Sec. 2.3.5). Finally, a centered rolling median with a symmetric 35-day window is subtracted from each observation to
 180 remove low-frequency temporal trends. As no real-time processing is required, this non-causal filter is considered appropriate.

We parameterized the rolling window size, the median filter length, and the humidity threshold using all available deployment data where sensors were co-located with a reference analyzer. The proposed drift correction is active only during

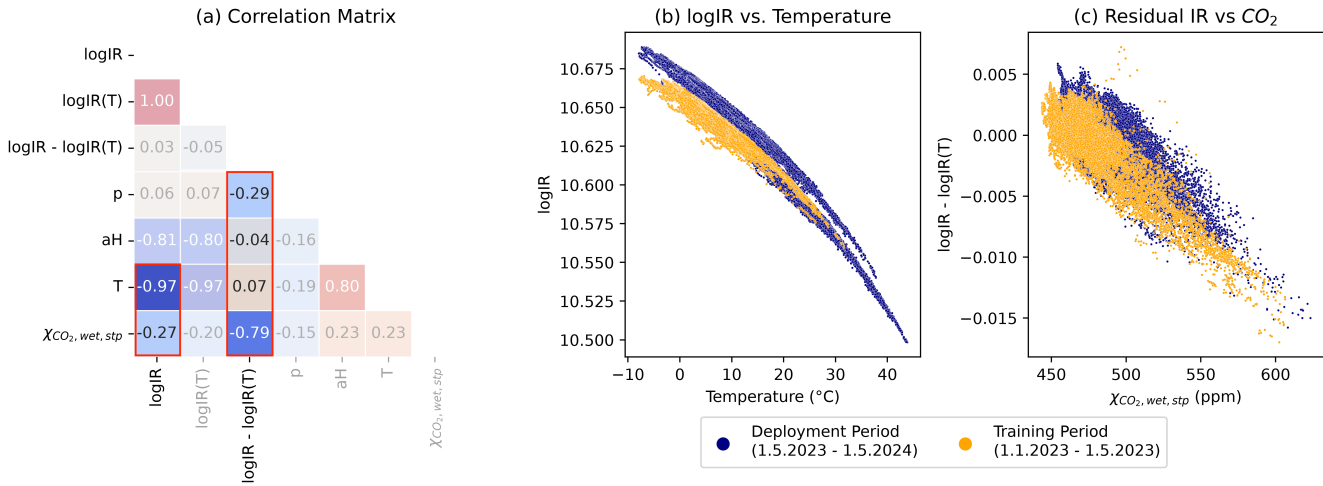


Figure 6. Correlation between a single, well-working, representative sensor’s log-transformed infrared signal (ID:13156) ambient values, and the wet CO_2 molar mixing ratio at standard conditions derived from a reference instrument. Panel (a) shows a correlation matrix between all measured and computed values. The raw log-transformed IR value shows a strong correlation with temperature, but the relationship is not stationary (panel b). After applying the rolling regression and removing the temperature component, the residual IR signal ($\log IR_i - \log IR(T)_i$) is linearly related to the wet CO_2 mixing ratio at standard conditions $\chi_{CO_2, wet, stp}$ (Fig. c).

deployment, and its interaction with the calibration model should be accounted for in the parameterization. Additionally, only the deployment period spans all seasons.

185 The generalization of the selected parameters was assessed with one-at-a-time sensitivity tests, varying each parameter while holding the others fixed. The sweep shows that all selected parameters show a broad plateau around the selected values, rather than a sharp optimum. The sensor-to-sensor variability of 6 to 8 ppm strongly exceeds the change in the median RMSE, which varies by less than 0.5 ppm (Appendix B) across the tested parameter range.

190 The residual $\log IR(CO_2)$ is physically proportional to the CO_2 number density via the Beer-Lambert law. Referencing the mole fraction to standard conditions holds the moist-air number density constant, so the residual becomes linearly related to the wet CO_2 mixing ratio at standard conditions $\chi_{CO_2, wet, stp}$. Figure 6c shows $\log IR(CO_2)$ plotted against $\chi_{CO_2, wet, stp}$ from the reference analyzer, indicating a linear relationship as well as a strong overlap between the training and deployment time period.

2.3.3 Step 3: Inference of the CO_2 mole fraction

195 To infer the gas concentration from the temperature-corrected sensor signal $\log IR(CO_2)$, we optimized individual multilinear regression models for each sensor using Scikit-Learn (Pedregosa et al., 2011). The model uses relative humidity and atmospheric pressure as additional input features. The relative humidity is expanded to a quadratic polynomial. Atmospheric pressure is normalized to p_0 using $(p - p_0)/p_0$. The regression target is the wet mole fraction at standard conditions, $\chi_{CO_2, wet, stp}$.



The formulation of the regression model is shown in Eq. 3. For optimization, we use a weighted least squares approach to increase the weight of infrequent, high CO₂ concentrations. Weights were determined from the histogram of the regression target and clipped at the upper and lower range to avoid very small and very large weights. Finally, the result is converted to ambient atmospheric conditions using the ideal gas law (Eq. A4).

$$\chi_{CO_2, wet, i} = \frac{p_0 \cdot T}{p \cdot T_0} \cdot (\mathbf{x}_i \boldsymbol{\beta}_t + e_i) \quad \text{where} \quad \mathbf{x}_i = [1, \log \text{IR}(CO_2), RH, RH^2, \frac{p - p_0}{p_0}]^T \quad (3)$$

2.3.4 Step 4: Water dilution correction

To account for the water content in the sampled gas, we determine dry-air mole fractions of CO₂, referred to as $\chi_{CO_2, dry}$. First, the saturation water vapor pressure p_σ given the sensor temperature T , the critical temperature T_c and pressure p_c of water is calculated using the Wagner equation (Eq. 4, Wagner and Pruß (2002)).

$$p_\sigma = p_c \cdot \exp \left[\frac{T_c}{T} (a_1 \tau + a_2 \tau^{1.5} + a_3 \tau^3 + a_4 \tau^{3.5} + a_5 \tau^4 + a_6 \tau^{7.5}) \right] \quad (4)$$

using

$$\tau = \left(1 - \frac{T}{T_c} \right) \quad T_c = 647.096 \text{ K} \quad p_c = 22.064 \times 10^6 \text{ Pa}$$

$$a_1 = -7.85951783 \quad a_2 = 1.84408259 \quad a_3 = -11.7866497 \quad a_4 = 22.6807411 \quad a_5 = -15.9618719 \quad a_6 = 1.80122502$$

The measured relative humidity RH is then used to calculate the partial pressure of water vapor. Applying Dalton's law of partial pressures, the measured ambient pressure p is used to determine the mole fraction of water vapor, χ_{H_2O} (Eq. 5). Finally, the dry mole fraction $\chi_{CO_2, dry}$, is calculated accordingly (Eq. 6).

$$\chi_{H_2O} = \frac{p_\sigma \cdot \frac{RH}{100}}{p} \quad (5)$$

$$\chi_{CO_2, dry} = \frac{\chi_{CO_2, wet}}{1 - \chi_{H_2O}} \quad (6)$$

2.3.5 Step 5: Long-term drift correction

Low-cost CO₂ sensors are limited in their measurement accuracy, particularly in long-term stability. As a remedy, more advanced systems use controlled injections of reference gas for regular calibration (Aigner et al., 2026a; Grange et al.; Lian et al., 2023). A more cost-efficient and lightweight approach relies on specific atmospheric conditions (e.g. high wind speeds), in which horizontal gradients are expected to be low, and the sensor's measurement signal can be related to a nearby reference station (Creman et al., 2026; Müller et al., 2020; Delaria et al., 2021). We follow the latter option and have examined three



scenarios for Munich, each requiring a different configuration of reference stations: (1) No reference station in the city but
 225 one in the region, (2) a single reference station in the city, (3) a network of reference stations in the city. We apply all three
 scenarios because the city reference station and the network of reference stations were not available from the beginning of the
 deployment period.

In addition to the CO₂ measurements, we use wind data from a meteorological station in the city center, available at 1-
 minute intervals. An exploratory analysis indicated that wind speed thresholds provide the most robust heuristics, whereas
 230 wind direction and other meteorological parameters offered no substantial, testable advantage in any of the presented scenarios.
 For scenarios 1 and 2, we define wind speed thresholds to select periods when horizontal CO₂ gradients are small, and the
 reference signal can be related to all low-cost sensors in the city. In scenario 3, the network of reference stations enables
 a direct assessment of when horizontal gradients are small, as indicated by a low maximum-to-minimum spread across the
 network of reference stations. We additionally restrict the data to afternoon hours (12:00–15:00 UTC) for scenarios 2 and 3,
 235 when the convective boundary layer is deep and vertical gradients between the rooftop reference stations and the street-level
 SCOUT sensors can be assumed negligible. In scenario 1, this restriction would result in a low calibration frequency. Instead,
 we rely on the high wind speed threshold, as turbulent mixing under these conditions suppresses vertical gradients.

A quantitative analysis of the scenarios is presented in section 3.3.2. For all presented thresholds, a sensitivity analysis was
 conducted to balance uncertainty and the criterion’s frequency.

240 **Scenario 1: A single reference station in the region.**

The closest ICOS atmospheric station to Munich is at Hohenpeissenberg (HPB), approximately 55 km southwest of the city,
 and it also corresponds to the prevailing wind direction. HPB is an elevated site situated on a small hill and hosts a mobile
 network tower from which air samples are collected at three heights: 50, 93, and 131 m above ground level. We use the mea-
 surements at 50 m above ground level. The elevated position decouples the site from local surface fluxes, so the sampled air
 245 is representative of the regional background. Quality-controlled and near-real-time data can be accessed through the ICOS
 Carbon Portal (Kubistin et al., 2026). At high wind speeds, the concentration difference between the mean urban CO₂ con-
 centration and HPB decreases. We define a threshold of $\overline{ws}_{3h} \geq 5 \text{ m s}^{-1}$ on the 3-hour rolling-mean wind speed $\overline{ws}_{3h}$. An
 average bias of $b_{\text{HPB}} = 4.0 \text{ ppm}$ between the city and HPB (Fig. 12a) is also accounted for to ensure a bias-free result at the
 given conditions. We note that this bias can only be quantified using reference measurements within the city, without which a
 250 persistent network-wide mean bias must be expected.

Scenario 2: A single reference station in the city.

The second scenario requires a single reference station within the city, which, in our case, is located in the city center on the
 university building (Site ID: TUMR). In this setup, a more relaxed wind-speed criterion can be applied while maintaining a
 comparable standard deviation across all urban stations. We define a threshold of $\overline{ws}_{3h} \geq 4 \text{ m s}^{-1}$ for the 3-hour rolling mean
 255 wind speed and restrict the data to afternoon hours ($t \in [12:00, 15:00]$). An average bias of $b_{\text{TUMR}_P} = -0.6 \text{ ppm}$ (Fig. 12b)
 across the city relative to the single reference station is again accounted for, as in scenario 1.



Scenario 3: A network of reference stations in the city.

The ACROPOLIS mid-cost CO₂ sensor network in Munich was gradually established alongside the deployment of the SCOUT network. It allows direct assessment of the homogeneity of the concentration field by calculating the maximum-to-minimum spread across spatially distributed stations. We define a maximum spread of 9 ppm across at least eight active stations in the city. In addition, we filter for afternoon hours, when the vertical gradients are small. Equation 7 shows the full formulation of the condition. $S(t)$ denotes the number of active stations i at time t . All threshold values were determined empirically to balance the frequency of the criterion and the associated uncertainty. When the criteria are satisfied, the mean concentration across all stations is used as a reference value. As it represents the mean city signal $\chi_{\text{CO}_2,i}(t)$, no further bias correction must be applied (Fig. 12c).

$$\max_{i \in S(t)} \chi_{\text{CO}_2,i}(t) - \min_{i \in S(t)} \chi_{\text{CO}_2,i}(t) < 9 \text{ ppm}, \quad |S(t)| \geq 8, \quad t \in [12:00, 15:00] \text{ UTC} \quad (7)$$

Combining Scenario 1–3: City baseline

A city baseline (Fig. 7) is constructed from the data and heuristics described above to determine each sensor's drift at the respective time period. Overlapping scenarios are resolved using a fixed-priority order ($3 > 2 > 1$), reflecting their increasing location-dependent uncertainty as quantified in section 3.3.2. The city baseline is formalized in Eq. 8. At the individual sensor level, the drift correction is applied to outlier-filtered, hourly averaged dry CO₂ mole fractions. Between calibration intervals, drift estimates are linearly interpolated, and the resulting time series is smoothed with a centered 3-hour rolling mean filter.

$$\chi_{\text{CO}_2,\text{base}}(t) = \begin{cases} \overline{\chi_{\text{CO}_2,i}(t)} & \text{if Eq. (7)} \\ \chi_{\text{CO}_2,\text{TUMR_P}}(t) + b_{\text{TUMR_P}} & \text{else if } \overline{\text{ws}_{3\text{h}}} \geq 4 \text{ m s}^{-1}, t \in [12:00, 15:00] \text{ UTC} \\ \chi_{\text{CO}_2,\text{HPB}}(t) + b_{\text{HPB}} & \text{else if } \overline{\text{ws}_{3\text{h}}} \geq 5 \text{ m s}^{-1} \\ \text{undefined} & \text{otherwise} \end{cases} \quad (8)$$

2.3.6 Step 6: Data post-processing

For the final site-level data products, we average 10-minute data to hourly values, flag untrusted measurements, and combine measurements from paired sensor nodes into a single signal. In some cases, sensors were also replaced at specific locations, requiring the merging of data sets from multiple sensors in chronological order. Final data files for each site are published on the ICOS Carbon Portal. The post-processing requires three steps, as described below.

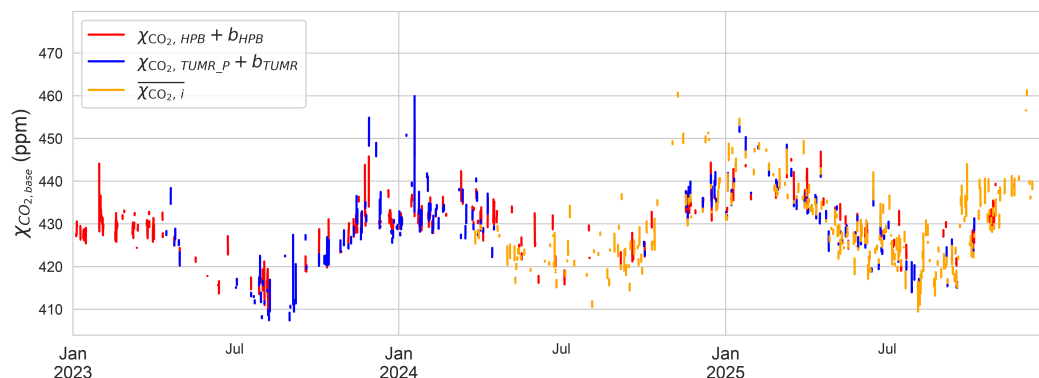


Figure 7. City baseline assembled from three scenarios in which higher-grade reference stations are related to the SCOUT sensor signal. $\chi_{CO_2,HPB}$ denote the reference mole fraction at HPB, corrected by $b_{HPB} = 4.0$ ppm to eliminate the residual mean bias. $\chi_{CO_2,TUMR_P}$ represents the Picarro measurement at our reference site TUMR, corrected by $b_{TUMR_P} = -0.6$ ppm. $\bar{\chi}_{CO_2,i}$ represents the mean concentration across eight independent sites i in the city. This baseline is used to correct the SCOUT sensor signal for its long-term drift.

Temporal aggregation to hourly measurements

280 We report hourly sensor measurements to the ICOS Carbon Portal, in line with the timescale commonly used for atmospheric modeling. First, we drop all data points identified as outliers by previous processing steps. We average the measured values for $\chi_{CO_2,dry}$, temperature, relative humidity, and air pressure to hourly means. We also count the number of data points and calculate the standard deviation of $\chi_{CO_2,dry}$, which are added as separate columns to the data set.

Combining the sensor signals

285 If redundant sensors are available from the same location and both sensors work properly, we average their measurements and report the combined standard deviation and number of data points in the final dataset. As a fault indicator, we compute the absolute difference between the two sensors' hourly measurements and apply a 6-day centered rolling mean to detect gradual divergence. An “acceptable” threshold is defined by the performance assessment (Sec. 3.3) and set to 17 ppm ($\approx 3 \times \text{MAE}$). For each hourly time step in which the rolling absolute mean difference exceeds this threshold, a counter is incremented. If it falls
 290 below the threshold, the counter is decremented. When the counter reaches the defined limit of 48 (equivalent to 2 days), a changepoint is marked, while the counter is capped at 96 (4 days) to prevent runaway accumulation of the value.

For each time period between successive changepoints, we calculate the mean absolute difference (MAD) between both sensors and the outlier ratio of each individual sensor, defined as the number of outliers (see Sect. 2.3.1) divided by the total number of measurements for that period. If the MAD is below the threshold of 17 ppm, we report the averaged signal. If the
 295 MAD is above the threshold, we select the sensor with the lower outlier ratio. The rationale is that when the MAD diverges, one sensor typically reports anomalous signals with large amplitudes, likely caused by condensation in the optical path. These



anomalous signals are detected as outliers, so the outlier ratio serves as a useful proxy for identifying the affected sensor. If both sensors are affected by this issue, we can observe it in the final visual inspection and flag it as untrusted.

Flagging

300 The final dataset is flagged as **O** for trusted sensor data after automatic processing and as **K** for untrusted data. The **K**-flag implies that there are either (1) more than 50 % of the expected data is missing (< 3 data points for a single sensor and < 6 data points for redundant sensors), (2) the CO_2 value is outside a reasonable range (< 400 ppm; verified using data from the ACROPOLIS network), or (3) the outlier ratio of the sensor is above 0.5 (more than 50 % outliers). The **K**-flag was also applied to locations where two nodes are typically deployed, but only data from one sensor is available. In this case, 305 the aforementioned averaging and related quality control cannot be performed. Finally, a visual inspection of the data was conducted, and any additional obvious outliers were manually flagged before the data were released.

3 Results and Discussion

3.1 Raw data coverage

Over the period from January 2023 to December 2025 (36 months), we collected 14 million measurements. This dataset includes 310 four months of outdoor side-by-side measurements for all sensors (Jan. 2023–April 2023), an additional twelve months of outdoor side-by-side measurements for a subset of 11 nodes (May 2023–May 2024), and up to 32 months of field measurements (May 2023–Dec. 2025). The dataset that was used for sensor model calibration (outdoor side-by-side period for all sensors) spans CO_2 concentrations from 419 to 625 ppm, temperatures from -8.6 to 33 °C, relative humidity from 23 to 100 %, and atmospheric pressure from 93486 to 99152 Pa. After applying the Müller filter, 80.2 % of all measurements are classified 315 as valid. A tabular overview of the collected data and its use for model training and sensor performance assessment can be found in appendix C.

Figure 8 shows the number of data points received and indicates valid and outlier-flagged data. A total number of 144 measurements is expected for each active sensor per day, with an average of 6.4 % of the measurements not received due to LoRaWAN's send-and-forget protocol. We observe a high fraction of outliers during humid periods in fall and winter: 21.65 320 % between September 2023 and February 2024, and 37.48 % between September 2024 and February 2025. Comparing the meteorological conditions of both periods reveals a significantly higher average humidity and slightly lower temperature during Sept.-Feb. 2024/25 (82.0 %, 7.07 °C) compared to Sept.-Feb. 2023/24 (73.8 %, 9.3 °C). Additionally, the second winter has double the number of hours with relative humidity above 90 % RH, the humidity range where the number of outliers increases sharply. A further impact is exerted by the solar radiation shield, which will be discussed in more detail in Sect. 3.5. The share 325 of outliers among sensors within radiation shields during the second winter is slightly higher (38.6 %) than among devices without a shield (36.9 %). Some sensors began running out of battery in August 2025, after almost 3 years of continuous

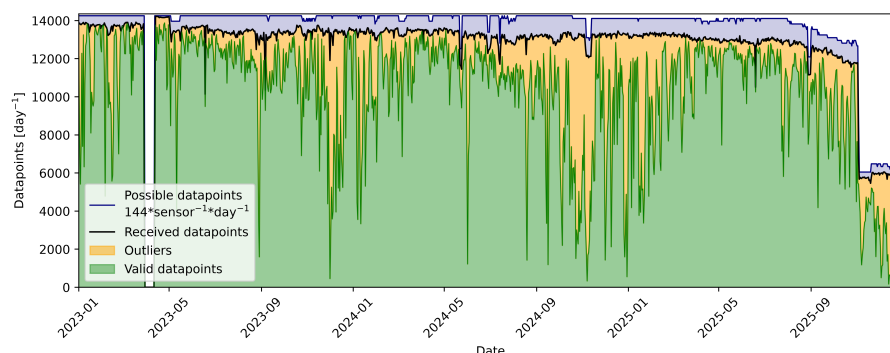


Figure 8. Raw data availability for the time period from January 2023 until October 2025. Each active node sends measurements every 10 minutes, resulting in a total of 144 possible data points per day. On average, 6.4 % of the expected data points were not received due to poor LoRaWAN connectivity at certain locations. A high number of outliers occurred during fall and winter, as humid conditions negatively affect the sensor measurements. Several weeks of bad weather, therefore, result in a low yield of valid data. Towards the end of the project in December 2025, several sites were also decommissioned because they had been exhibiting poor data quality since September 2025.

operation. The sensor network was also partially dismantled in November 2025, particularly sensors that exhibited anomalous signals after visual inspection of the data.

3.2 Selection of well-performing sensors

330 As a first-level quality metric, we assess the correlation between the temperature-corrected infrared signal $\log \text{IR}(\text{CO}_2)$ and the
 reference CO_2 mole fraction at standard conditions $\chi_{\text{CO}_2, \text{wet}, \text{stp}}$ for the training time period (compare Fig. 6). The evaluation
 of all 100 sensor nodes revealed a substantial performance spread, with some nodes consistently performing worse than others.
 This performance spread is commonly observed in similar studies involving multiple sensors (Mueller et al., 2021; Creman
 et al., 2026) and can be attributed to manufacturing variations in low-cost optical components, as well as to the spread in
 335 the specificity and sensitivity of the detector and optical filter unit (Dinh et al., 2016; Ueno et al., 2013). Figure 9 shows the
 distribution of the correlation coefficients across all sensors for the training time period. The temperature-corrected raw signal
 negatively correlated with the CO_2 mole fraction, as higher absorption results in a lower sensor signal. Applying an empirically
 chosen threshold of $r < -0.7$ identifies seven poorly performing nodes that deviate significantly from the rest of the sensors.
 All subsequent analyses exclude these seven bad-performing sensors.

340 3.3 Sensor performance

3.3.1 Measurement accuracy

The sensor performance was evaluated by comparison with measurements from a calibrated reference analyzer (Picarro G2301)
 installed on the rooftop of our university building (site: TUMR). The sensor installation started on 2 May 2023 and continued

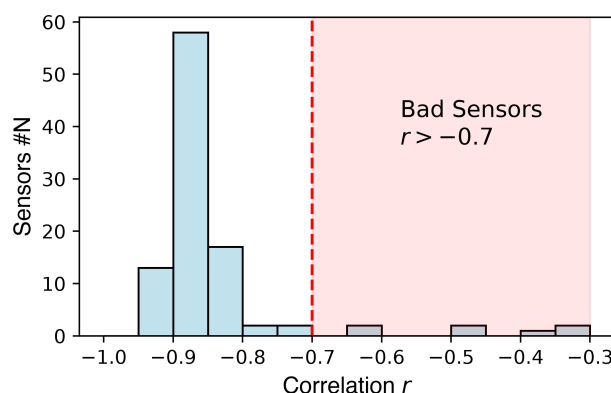


Figure 9. Histogram of the correlation coefficient between $\log \text{IR}(\text{CO}_2)$ and $\chi_{\text{CO}_2, \text{wet}, \text{stp}}$ to identify poor-performing sensors based on the raw sensor signal. High CO_2 mole fractions result in a lower sensor signal (anticorrelation). Applying a threshold of $r < -0.7$ identified seven out of 100 sensors that perform substantially worse than the rest. They likely suffer from manufacturing variations with the low-cost optics or detector.

for several months. During the installation phase, some sensors remained side by side with the reference unit for multiple
 345 months before their deployment. For performance evaluation, we selected 21 nodes that were co-located for more than three months between May 2023 and May 2024. A subset of 11 nodes was co-located for the full year to capture seasonal variability and support robust performance evaluation. For comparison, we applied two additional calibration models substituting step 2 and 3 in our processing pipeline: a slope and bias correction of the sensor's reported CO_2 concentration and the calibration model proposed by Müller et al. (2020). The Müller model is physically based on the Beer–Lambert equation and uses the
 350 sensor's raw IR measurements, temperature atmospheric pressure as model input. The raw data pre-processing (step 1), water dilution and drift correction (step 4-5) remain identical across all models. The performance assessment was carried out on outlier- and drift-corrected hourly averaged values.

Figure 10 shows scatter plots against environmental influences of all 21 nodes, corrected with three different models. The slope and bias correction achieves a substantial improvement compared to the manufacturer's stated accuracy of ± 50 ppm, but
 355 systematic biases can be observed at both temperature extremes (Fig. 10a), which also correspond to outliers in humidity and across the pressure range. The Müller model achieves a better performance than the slope and bias correction, but also shows a systematic deviation, particularly at high temperatures (Fig. 10e). However, it is important to note that Müller's training dataset included climate-chamber measurements up to 52°C and 1000 ppm CO_2 , covering a broader range of possible ambient conditions and potentially providing a better fit at high temperatures and thereby avoiding these errors. However, the maximum
 360 temperature in our climate chamber experiments is 32°C since the significant heating caused by solar radiation was not taken into account when planning the experiments.

Finally, our model shows minimal systematic dependence of the residuals on atmospheric parameters but the reference CO_2 mole fraction, indicating an underestimation of higher values (Fig. 10l). The deviation can be explained by the non-linearity



of the LP8's NDIR response at higher mixing ratios. We deliberately fit a first-order linear model, as the available data at high concentrations are too sparse to robustly constrain higher-order terms and avoid overfitting. Polynomial or spline expansions can capture these deviations in principle, but tend to overfit in the well-sampled low-concentration region without improving network-wide performance. We therefore retain the linear model as the most generalizable and robust choice across all sensors. A second sensor-intrinsic effect is the finite temporal response, which dampens high-frequency signals that are well captured by Picarro's fast sampling (1 sec). The response time required for the sensor signal to reach 90 % of its final value after a step change in concentration equals ten minutes under laboratory conditions (Senseair AB). Emulating this response by low-pass filtering the reference measurements reduces the residual slope, confirming the effect alongside the non-linearity. A third, calibration-related contribution arises during multi-day CO₂ accumulation periods under persistent stable conditions, which produce a rising urban background signal. Under these conditions, the atmospheric situations required to anchor the drift correction are absent. This leaves accumulated drift uncorrected and produces a systematic deviation. Users of the data can draw on the city baseline to evaluate the time elapsed since the last calibration and thereby assess the associated increase in uncertainty. This systematic deviation at higher concentrations is even stronger for the other models (Fig. 10d,h).

Finally, Fig. 11 compares the field performance of the three calibration models across all sensor's RMSE, MAE, and R². ROLS achieves a median RMSE of 9.46 [6.1-12.7] ppm, the Müller model 10.31 [5.3-25.0] ppm, and the slope and bias correction 11.7 [8.2-25.5] ppm. Similarly, the MAE decreases, and R² increases across the sensor models. ROLS achieves a MAE of 5.59 [4.3-9.5] ppm and R² of 0.76 [0.55-0.89]. The consistently narrower distribution of ROLS across multiple sensors and all metrics demonstrates its suitability for the entire network.

Further performance assessment was conducted at sites with co-located ACROPOLIS mid-cost sensor systems. But at three of six sites, problems were observed due to difficult measurement conditions on exposed rooftops and erroneous ACROPOLIS measurements. At some other locations, we installed nodes that were initially classified as bad sensors (see 3.2) and replaced their sensor element. We conducted additional 2-month side-by-side measurements with a Picarro reference analyzer and the replaced sensor element and deployed the devices accordingly. However, a leak in the air hose of the reference analyzer was detected, leading the reference instrument to measure air from inside the measurement cabinet. In summary, it was not possible to fit a meaningful sensor model for these devices. For the ACROPOLIS co-location at Bogenhausen and Klinikum Rechts der Isar, we calculate RMSEs of 12.2 and 14.4 ppm, mean biases of 2.2 and 9 ppm, and R² values of 0.48 and 0.64. This also includes the measurement uncertainty of the ACROPOLIS system, which was evaluated at 1.16 ppm for hourly values.

3.3.2 Location dependent bias

We tested the proposed drift correction scenarios with quality-controlled data from the ACROPOLIS sensor network (Aigner et al., 2026b), which has a stated accuracy of 1.16 ppm, alongside Picarro data from the city-center reference station. The resulting mean and variance provide estimates of the location-dependent bias and uncertainty. Figure 12a shows residuals between all reference stations within the city relative to the HPB ICOS station. Filtering for wind speeds $\bar{w}s > 5 \text{ m s}^{-1}$ minimizes the differences but leaves a positive bias of 4.01 ppm. When creating the city baseline, we have subtracted this value to obtain an unbiased baseline for this scenario. The standard deviation across all city locations relative to HPB is 2.87 ppm

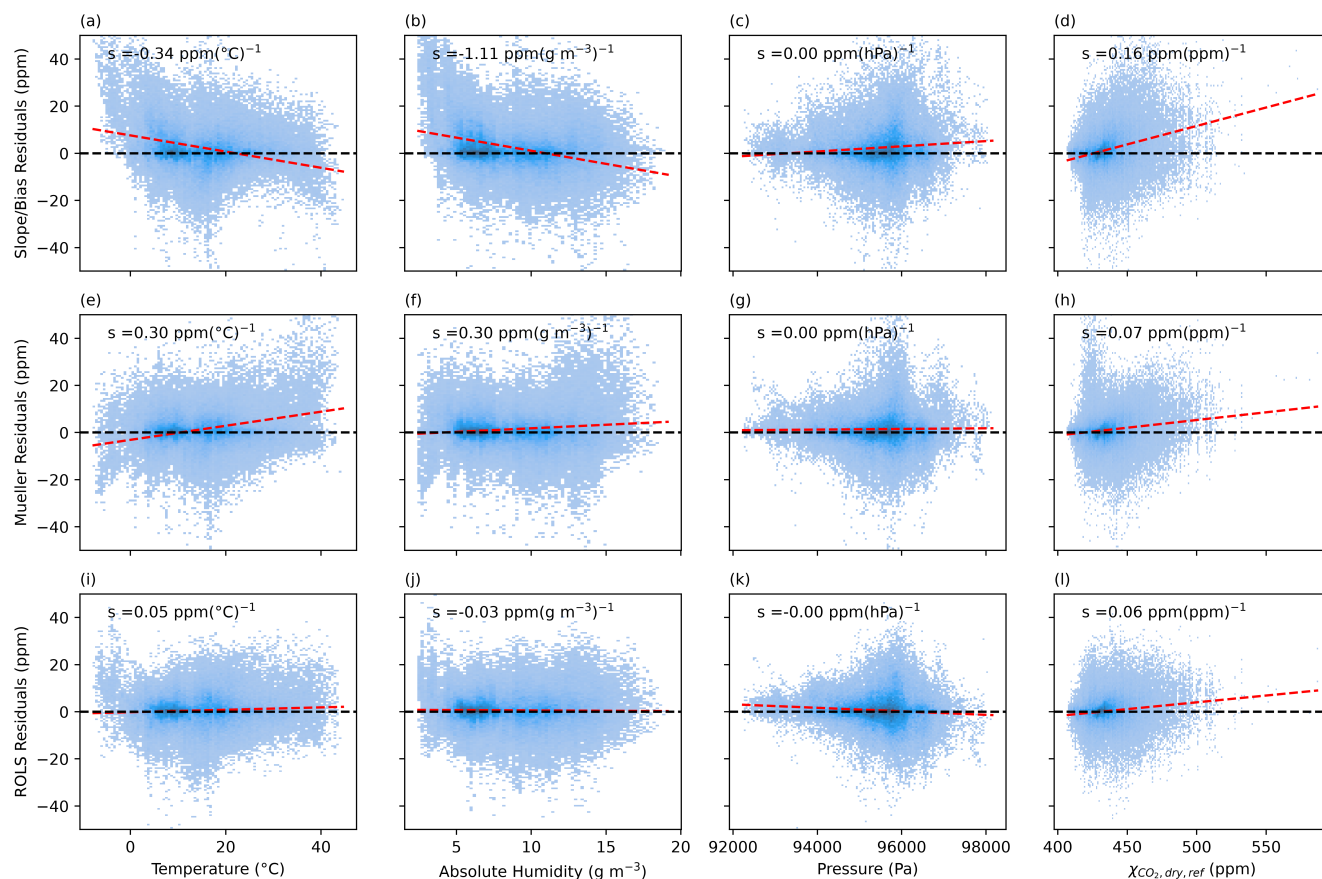


Figure 10. Residual scatter plots from 21 sensors available for side-by-side comparison with a reference analyzer. Residuals for three calibration models are plotted against temperature, absolute humidity, atmospheric pressure, and the CO₂ mole fraction. Residuals are defined as the difference between the Picarro reference and the corrected sensor value.

(Fig. 12d). Combined in quadrature with the 1.16 ppm measurement uncertainty of the ACROPOLIS stations, this yields a total uncertainty of 3.09 ppm. Figure 12b shows residuals between all ACROPOLIS stations and the TUM Picarro. For wind speeds above 4 m s⁻¹ during the afternoon hours, the differences are minimal, and a negative bias of 0.63 ppm persists, which again was subtracted from the city baseline. The standard deviation of 2.48 ppm yields a total uncertainty of 2.88 ppm when added in quadrature with the measurement uncertainty. Figure 12c shows the network spread defined as the maximum-minimum of at least eight active stations during afternoon hours. Restricting to a maximum spread of 9 ppm and afternoon hours yields a standard deviation of 1.9 ppm with zero bias, and a total uncertainty of 2.3 ppm. We conclude that the location-dependent bias remains below 3.1 ppm across all three scenarios, well below the measurement accuracy of the low-cost sensors and therefore sufficient to constrain the sensor-dependent drift.

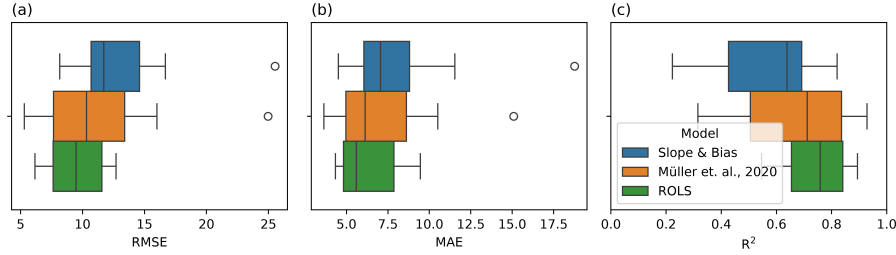


Figure 11. Field performance of 21 individual sensors co-located with a reference analyzer at a single site for at least three months between May 2023 and May 2024. Boxplots show the distribution of RMSE, MAE, and R^2 across individual sensors ($n = 21$) for three calibration models: slope and bias, the Müller model (Müller et al. (2020)), and the proposed ROLS model. ROLS achieves the lowest median RMSE and MAE and the tightest performance distribution, indicating robust and consistent behavior across devices.

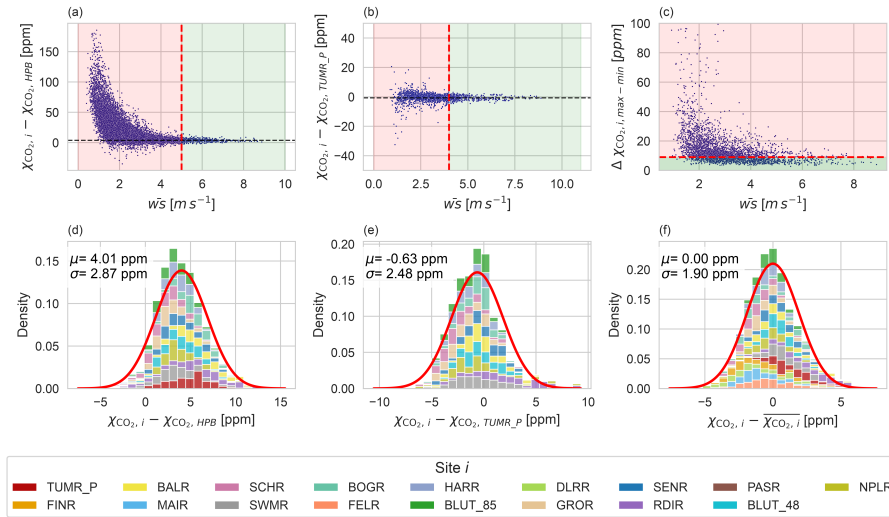


Figure 12. Scatter plots (a-c) and corresponding residual histograms (d-f) for three drift-correction scenarios. Data from individual ACROP-OLIS sites distributed across the city, and from a Picarro reference site (TUMR_P) in the city center, were used for the evaluation (TUMR_P only for scenario 1). The scenarios were evaluated over 22 months, from 1 February 2024 to 31 December 2025, during which the Picarro reference site TUMR_P was unavailable for 8 months. The 3-hour mean wind speed, $\bar{w}_s$, measured at the central meteorological station is shown on the x-axis of panels a-c. Scenario 1 (a, d) uses data from the nearby Hohenpeissenberg ICOS station during wind speeds $> 5 \text{ m s}^{-1}$; a remaining positive bias of 4 ppm is observed. Scenario 2 (b, e) relies on a single Picarro reference station at TUMR, allowing for a lower wind-speed threshold of 4 m s^{-1} . Scenario 3 (c, f) applies the mean concentration of all sites within the city boundary $\overline{X_{CO_2,i}}$ when the maximum difference among at least eight stations is below 9 ppm during the afternoon (12:00-15:00 UTC).



Table 1. Final sensor performance, combining the measurement uncertainty with the location-dependent bias introduced by the drift correction, for three drift correction scenarios. Measurement accuracy and field performance are given as median [minimum, maximum].

	Scen. 1	Scen. 2	Scen. 3
Location dependent bias	3.09 ppm	2.88 ppm	2.3 ppm
Measurement accuracy	9.46 [6.1-12.7] ppm		
Field performance	9.95 [6.41-13.1] ppm		

3.3.3 Overall sensor performance

The measurement accuracy of the sensors was evaluated at a single, well characterized site (TUMR). To determine the field performance, we add the uncertainty from the drift correction in quadrature to the measurement uncertainty resulting in an expected median sensor performance of 9.95 ppm. In the best case (well-functioning sensor combined with scenario 3 for drift correction) this reduces to 6.41 ppm. In the worst case it reaches 13.1 ppm. This upper limit is optimistic, as it presumes independent measurement and drift-correction errors and drift corrections frequent enough to prevent accumulation. A shortcoming of our experiment is that no street-level reference data are available in Munich. We therefore cannot assess sensor performance across multiple independent sites with distinct characteristics and local source contributions. For the drift correction, omitting the vertical gradient between roof and street level is nonetheless justified by the filtering criteria applied. High winds (scenario 1) and afternoon hours (scenarios 2 and 3) promote strong mechanical and convective mixing, which suppresses vertical mole fraction differences. Under these conditions the roof-level reference is representative of street level, so the omission introduces negligible bias. We acknowledge the limitations of assessing network-wide sensor performance at a single site and recommend that similar studies incorporate reference measurements at multiple street- and roof-level sites to obtain sufficient validation data. An overview of the findings of the sensor performance is provided in table 1.

3.4 Quality-controlled dataset

Following the logic presented in Sect. 2.3.6, we report the averaged signal from sites where two sensor nodes are located. The post-processed data conform to the ICOS ATC data format and are classified as Level 2B quality, meaning that measurements from the primary CO₂ sensor are corrected using co-located on-board sensors (temperature, humidity, pressure) and external reference data (Schneider et al., 2019). Figure 13 depicts the process for an exemplary site (ERHV). The 6-day rolling mean absolute difference (Fig. 13a) is below the defined threshold until one sensor starts showing abnormal, outlier-prone behavior. Then the MAD repeatedly exceeds the defined threshold, triggering changepoints that partition the timeseries into distinct periods. In each period, the sensor with the lower average outlier ratio (Fig. 13b) is selected, filtering out episodes where one sensor degrades relative to the other. Panel (d) shows the final signal, which underwent further manual visual inspection to identify remaining artifacts and outliers not captured by the automated process. The finalized datasets for 50 sites have been uploaded to the ICOS Carbon Portal (Kühbacher and Chen, 2026a) and are publicly available. In the final, quality-controlled files 90.51 % of all observations are O-flagged, after automatic and manual quality control.

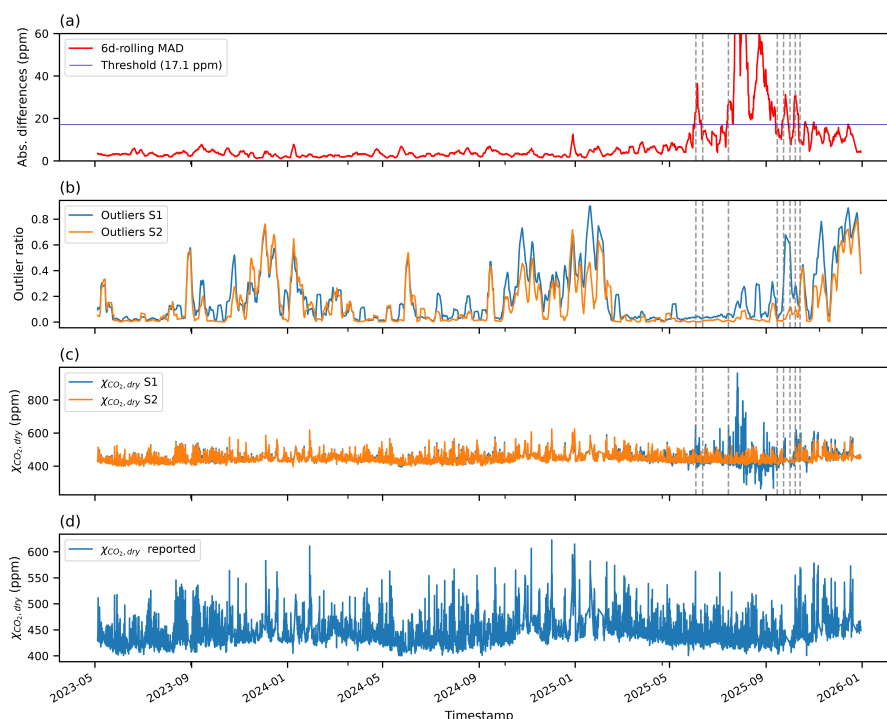


Figure 13. Illustration of the post-processing logic to extract a single sensor signal from two nodes at the same location. Panel (a) shows the 6-day rolling absolute difference between both nodes, which is used to determine changepoints (gray vertical lines). For time periods between two changepoints, the sensor signals are average if the mean absolute difference remains below a threshold of 17 ppm. If the MAD is above the threshold, we select the sensor with the lower outlier ratio (panel b) is reported (panel d). This example shows two nodes (IDs 13095 and 13141) at the ERHV site.

3.5 Impact of the solar radiation shield

At an early project stage, we applied the calibration model proposed by Müller et al. (2020) to improve the measurement accuracy. However, it consistently reported incorrect values during summer, when temperatures exceeded the temperature range observed during our side-by-side campaigns at Empa-Dübendorf and the site TUMR. To keep the sensor temperature closer to ambient conditions, sensor nodes at 35 sites were deployed in solar radiation shields (see Sec. 2.2.1). The paired nodes within the shield show excellent agreement with each other ($\Delta T = 0.01 \pm 0.12^\circ\text{C}$; $\Delta RH = 0.25 \pm 2.4\% \text{ RH}$), evaluated across all sensors deployed in radiation shields from June 2024 to December 2025.

To assess the effect of the naturally ventilated solar radiation shield on temperature and humidity measurements, three sensor nodes were co-located with a reference-grade meteorological station: two housed within the radiation shield and one in the manufacturer's original casing. The shielded sensors cover the period from May 2024 to Nov. 2025, while the unshielded sensor was only co-located until Nov. 2024. For comparison, the measurements of both sensors inside the shield were averaged.



Relative to the reference station, the shielded sensors exhibit a mean temperature bias of -0.29°C and an absolute humidity
 445 offset of 0.31 g m^{-3} . Under high solar radiation ($> 50\text{ W m}^{-2}$), shielded measurements (Fig. 14a-c) are typically $\sim 1^{\circ}\text{C}$ lower
 than the reference, while at low radiation ($< 50\text{ W m}^{-2}$) both agree well. This negative bias during sunshine can be attributed
 to the faster response of the reference station due to its active ventilation and to evaporative cooling from moisture retained
 within the radiation shield. In contrast, the unshielded sensor (Fig. 14d-f) exhibits extreme radiative heating with temperature
 differences up to 10°C .

450 Both configurations measure a higher average absolute humidity than the reference. However, the relative humidity inside
 the radiation shield is notably higher ($+2.05\% \text{ RH}$) than for the unshielded setup ($-0.7\% \text{ RH}$), particularly under solar radi-
 ation and consistent with the lower temperature inside the shield reducing the saturation vapor pressure. Both setups strongly
 underestimate humidity above $80\% \text{ RH}$ and exhibit pronounced nonlinear deviations in their relative-humidity measurements.
 Wind speed had no significant observable effect on the residuals of temperature and humidity.

455 In summary, the shield significantly improves temperature accuracy, whereas sensors without it exhibit substantial excess
 heating when solar radiation is present. The humidity is, on average, higher inside the shield. Furthermore, as shown exem-
 plarily in appendix D1, a lot of dirt accumulated in some of the radiation shields, and the slightly rusty screws suggest high
 humidity or stagnant moisture, which creates a humidity-buffering effect and leads to additional outliers. Although the sensor
 inside the inner bracket was not contaminated, the dirt on the outside has likely affected the measurement.

460 In conclusion, we recommend protecting the sensors from strong solar radiation using suitable radiation shields, whilst en-
 suring adequate ventilation (active ventilation) and regular maintenance to ensure a clean and stable measurement environment.
 In our case, the decision was made in the summer of 2023, when we applied slope and bias correction or the Mueller model
 during data processing, both of which produced very large outliers at high summer temperatures (see Fig. 10a,e), which were
 counteracted by the radiation shield. This issue could also be solved with the proposed ROLS model that provides a more
 465 robust temperature correction across the range.

3.6 Use cases

3.6.1 Diurnal cycles and location characteristics

Averaged seasonal diurnal cycles for six selected sites across distinct urban microenvironments (Fig. 15) demonstrate that the
 SCOUT network resolves CO_2 variability across sites with distinct emission and uptake characteristics. The regional back-
 470 ground site (HPB) is shown for reference and remains largely invariant throughout the day across all seasons, with mole
 fractions consistently below those observed at the urban network sites. In winter, all sites show a pronounced bimodal cycle
 with a sharp morning peak around 06:00–08:00 UTC and an evening peak around 17:00–19:00 UTC. The morning peak ampli-
 tude increases with traffic volume, with higher amplitudes at a traffi-heavy site (LANC), intermediate amplitudes at light-traffic
 a site (LOTG), and lower amplitudes at a suburban site (SMIA). This is consistent with the combined effect of local source
 475 strength and a shallow nocturnal boundary layer. A shallow boundary layer with similar diurnal behavior can be observed
 in autumn. In contrast, the suburban/biogenic site (SMIA) display the most pronounced afternoon drawdown in spring and

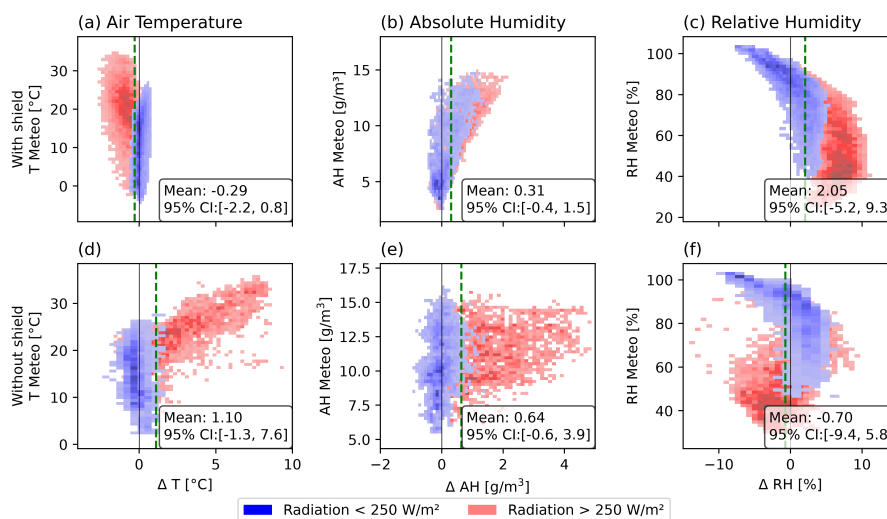


Figure 14. Effect of solar radiation on temperature and humidity measurements inside the manufacturer enclosure and solar radiation shield under varying radiation conditions. Panels a–c show residuals between a shielded sensor and a reference meteorological station for air temperature (ΔT), absolute humidity (ΔAH), and relative humidity (ΔRH). Panels d–f show the corresponding residuals for an unshielded sensor. Mean biases and 95 % confidence intervals are indicated in each panel. Under high radiation ($> 50 \text{ W m}^{-2}$), the unshielded sensor exhibits pronounced positive temperature biases due to radiative heating, whereas the shielded sensor remains close to reference conditions with a slight negative bias. The shield also affects humidity: relative humidity inside the shield is, on average, 2 % higher than without.

summer, attributable to photosynthetic uptake, together with elevated nighttime concentrations driven by ecosystem respiration under stable boundary-layer conditions. Seasonally, winter cycles show the highest mean concentrations and the largest morning peaks, while summer cycles are shaped by biogenic activity and enhanced vertical mixing. Overall, the consistency
 480 between sensors across urban sites with similar emission characteristics confirms that the SCOUT network captures temporal CO_2 variations associated with specific urban emission and uptake regimes.

3.6.2 Winter morning enhancement as local emission proxy

The morning enhancement before the boundary rises (03:00–07:00 LT) during wintertime (Dec.-Feb.) provides a suitable regime for a qualitative inference of local emission strength from low-cost sensor data (Shusterman et al., 2018). A shall-
 485 low nocturnal boundary layer and suppressed vertical mixing trap emissions near the surface, while biospheric fluxes are strongly reduced by low temperatures and vegetation dormancy. Local combustion sources such as residential heating and road traffic, therefore, dominate the signal, producing enhancements of several tens of ppm that lift it well above the expected low-cost sensor uncertainty. Averaging over the three winter months further suppresses day-to-day variability in meteorology and source activity, yielding a robust mean enhancement representative of typical local emission strength. We compare the winter
 490 2023/2024 morning enhancement with Munich’s high-resolution 2024 emission inventory (Kühbacher et al., 2025), gridded at

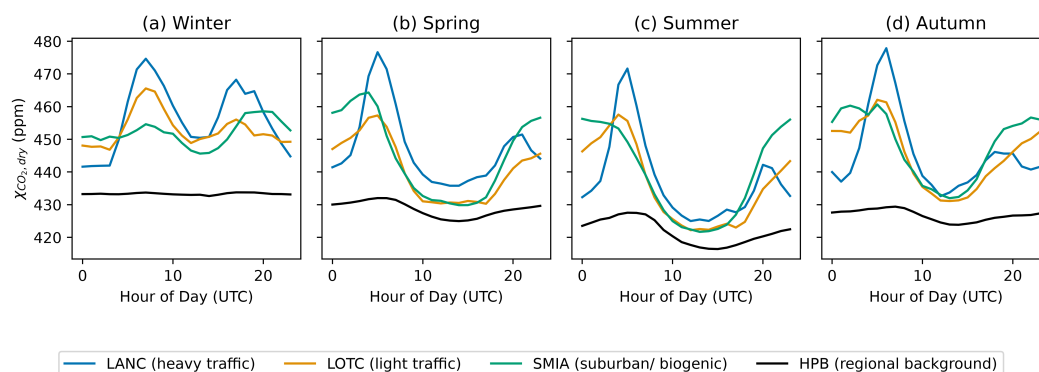


Figure 15. Seasonal diurnal cycles for six selected sites with different characteristics and the regional background (HPB, black). Hourly mean CO_2 mole fractions are averaged over each season and shown as a function of hour of day (UTC) for (a) winter, (b) spring, (c) summer, and (d) autumn. Sites are grouped by dominant emission and uptake characteristics: heavy traffic (LANC, blue), light traffic (LOTC, orange), and suburban/biogenic (SMIA, green). The regional background site HPB (black) is shown as a reference.

100 × 100 m. Inventory emissions are aggregated within a 300 m buffer around each sensor location. Sensitivity tests indicated that the correlation with the observed enhancement remains similarly high for buffer radii up to 500 m and decreases beyond, while disaggregating by sector (e.g., road transport) does not improve the correlation. This suggests that the signal is dominated by local sources and integrates across multiple emission sectors. Figure 16a shows the correlation r between the inventory CO_2 emissions and the observed morning enhancement, while panel (b) presents the measured morning enhancement superimposed on a spatial map of the emission inventory. The strongest enhancement of 34 ppm was observed at LANC, close to a busy road (Mittlerer Ring) with around 100,000 vehicles on an average weekday. The co-located governmental air quality monitoring station is also known for its high measured annual average nitrogen oxide levels (Bayerisches Landesamt für Umwelt, 2025), which are attributable to both heavy traffic and local accumulation effects in the surrounding urban canyon. The DOMV and PETV sites are located near roads with similar traffic levels but are not surrounded by a narrow urban canyon. Therefore, they can be expected to have better ventilation. Overall, all sites with a strong morning enhancement are heavily traffic-dominated. The lowest enhancements (< 4 ppm) were observed at LITA, ASAA, and MARA. These sites are located in the city's southern suburbs, away from busy roads, in residential areas with lower population density, consisting mainly of single-family homes. The red dashed fit line indicates a minimum enhancement of 3.3 ppm (y-intercept) across all sites beyond the 300 m buffer, which is affected by urban emissions. The slope of the line is 1.37 ppm of morning enhancement per kt of emissions in the inventory. We conclude that the sensors provide useful qualitative information on local emission strength and that the derived emission proxy is consistent with the emission inventory. A more detailed analysis that accounts for meteorological conditions, including boundary-layer height, temperature and wind, could be part of a future study.

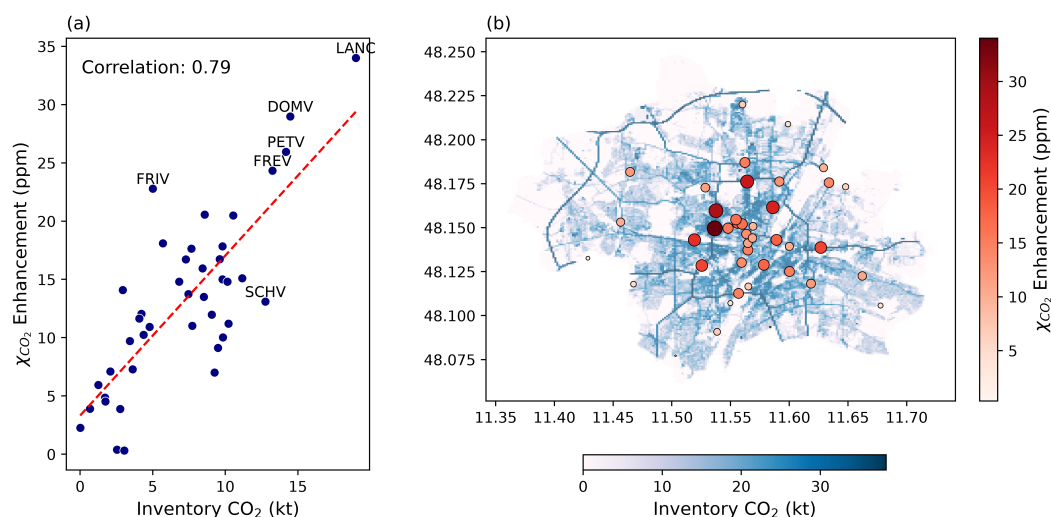


Figure 16. Correlation of the Munich city emission inventory and the winter morning enhancement (03:00–07:00 LT) measured by the SCOUT network in Nov. 2023 until Feb. 2024. Panel (a) shows the correlation between the measured enhancement and emissions of the inventory within a buffer distance of 300 m around the sensor location. A maximum enhancement of 34 ppm was observed at LANC, a heavily traffic-dominated site. Similarly, DOMV and PETV are close to busy roads and show strong morning enhancements. The red-dashed fit line ($y = 3.3 + 1.37x$) suggests a minimal enhancement of 3.3 ppm for all stations and a 1.37 ppm enhancement per kt CO_2 emission. In panel (b), the sensor locations and corresponding morning enhancements are superimposed on Munich’s high-resolution emission inventory (Kühbacher et al., 2025), with sizes and colors varying with the measured enhancement strength.

4 Conclusion

510 This study presents the deployment, data-processing strategy, and measurement accuracy of SCOUT, a low-cost CO_2 sensor network with 100 nodes deployed in Munich. The nodes are mounted in pairs on street-lamp posts for redundancy and fault detection, and sensors at 35 sites were mounted in solar radiation shields to prevent excessive heating. In many locations, the sites are co-located with governmental air quality stations, low-cost air quality stations, traffic detectors, and the mid-cost CO_2 sensor network ACROPOLIS. We propose a novel data-processing strategy based on a rolling regression model to correct for

515 the strong, non-stationary temperature dependence of the raw sensor signal. The processed raw signal is fed into a multiple linear regression model, optimized using 4 months of reference data to calculate the CO_2 mole fraction. Model accuracy was assessed at a single reference station equipped with a Picarro analyzer, and a median RMSE of 9.46 [6.1–12.7] ppm was calculated for a subset of 21 sensors over the time period of 3 to 12 months. Our proposed model shows negligible residual dependence on temperature, humidity, and atmospheric pressure. A long-term drift-correction strategy based on three reference

520 scenarios is presented and evaluated using data from a Picarro reference station and the ACROPOLIS network in Munich. After drift correction, a location-dependent residual bias of $\sigma < 3.1$ ppm remains. The total measurement uncertainty, combining contributions from the drift-correction method and intrinsic sensor noise, is estimated at 9.95 [6.4–13.1] ppm for all sensors.



We collected a total of 14 million measurements from January 2023 until December 2025. After applying an outlier detection algorithm on the raw 10-minute data, 19.8 % of the data points were flagged with a higher share during winter, caused by adverse atmospheric conditions for the measurements, primarily by high humidity. The sensors perform better under dry environmental conditions. The winter 2023/2024 shows a lower fraction of outliers than winter 2024/2025, which is related to significantly higher humidity in the second winter and to a humidity-buffering effect of the solar radiation shield. The unventilated radiation shielding trapped moisture and lead to the build-up of debris. Nevertheless, the radiation shield provides a more stable measurement environment for the sensors by reducing diurnal temperature amplitudes and demonstrates significant improvements in temperature measurements. This makes the data also useful for other urban climate related studies, such as those on the urban heat island effect.

Applying the data in two use cases to explore local emission characteristics and spatial variability shows that they can provide valuable CO₂ observations at street level, complementing the roof-level measurements of ACROPOLIS and MUCCnet in Munich. While the associated measurement uncertainty is higher, applying individual sensor models and careful calibration can achieve measurement accuracy sufficient to resolve signals from strong local sources (e.g., traffic or biogenic processes). It captures the traffic-related and near-surface variability that rooftop monitoring misses, making the network suitable for high-resolution urban emission modeling.

Drawing on the findings of SCOUT, we propose the following recommendations for similar networks:

- Carefully control for the temperature-induced signal component e.g. through our novel algorithm to explicitly separate the temperature-dependent signal component.
- Co-deploy sensors in pairs to enable cross-checking and fault detection. Given the low unit cost, this redundancy can be achieved economically.
- House sensors in a radiation protective shield with active ventilation to prevent radiative heating and buffering of moisture.
- Apply annual to bi-annual maintenance and, where possible, periodic recalibration with higher-grade reference instruments to ensure good measurement performance over a long time period.
- Take specific care of sensors co-located with reference instruments, and (if possible) incorporate co-located reference measurements at multiple street- and roof-level sites to collect data for model and method validation.

In future studies, we will make use of the vast number of co-located sensors measuring air pollutants and traffic volume to explore further use cases for the sensor data. It provides unique, direct observations of the urban emissions environment with high spatial density and offers new opportunities to resolve spatial variability of emissions. Furthermore, experiments with the sensor's successor, the Senseair Sunrise, show promising results for future low-cost CO₂ networks. Building on this, we propose two complementary paths for next-generation low-cost networks. First, integration with low-cost air-quality sensors toward a more general, multi-species monitoring system. Second, operation within a temperature-controlled, reference-gas-corrected platform such as ZiCOS-M or ACROPOLIS, which could yield a substantial performance gain.



Appendix A: Sensor working principle

The physical working principle of optical gas measurement as used in this NDIR sensor is the Beer-Lambert Law:

$$\log\left(\frac{I_0}{I_1}\right) = \epsilon_\lambda \cdot c \cdot d \quad (\text{A1})$$

I_1 and I_0 represent the detected and emitted light. ϵ_λ is the wavelength-dependent molar attenuation coefficient [$\text{m}^2 \text{mol}^{-1}$], c the number density of the gas in the light path [mol m^{-3}], and d the path length of the light beam through the optical cell. Carbon dioxide has a strong absorption in the mid-infrared range at 4.2-4.3 μm (Gordon et al., 2022).

The molar mixing ratio χ_{CO_2} of CO_2 in moist air is related to the number density c_{CO_2} via the number density of air c_{air} and subsequently the number of moles $n = c/V$. The number of moles in air depends on the atmospheric pressure p and temperature T , as defined by the ideal gas law.

$$\chi_{\text{CO}_2} = \frac{c_{\text{CO}_2}}{c_{\text{air}}} = \frac{n_{\text{CO}_2}}{n_{\text{air}}} \quad (\text{A2})$$

$$n_{\text{air}} = \frac{R \cdot T}{p \cdot V} = n_{p_0, T_0} \cdot \frac{p \cdot T_0}{T \cdot p_0} \quad (\text{A3})$$

$$\chi_{\text{CO}_2} = c_{\text{CO}_2} \cdot \frac{V}{n_{p_0, T_0}} \cdot \frac{T \cdot p_0}{p \cdot T_0} \propto c_{\text{CO}_2} \cdot \frac{T \cdot p_0}{p \cdot T_0} \quad (\text{A4})$$

with p , T , and V denoting the pressure, temperature, and volume of the gas, $p_0 = 1013.25 \text{ hPa}$, and $T_0 = 273.15 \text{ K}$, and n_{p_0, T_0} the number of moles in the air at standard conditions. The volume and number of moles at standard conditions are constant and therefore, the molar mixing ratio is proportional to the number density through the ratio $\frac{T \cdot p_0}{p \cdot T_0}$.

A1 Electrical sensor characteristics

The detector and light source remain off except during measurement. The detector is first powered on, and after a settling time, a dark-current measurement is taken with a specified integration time. After 80 ms, the light source emits a 50 ms light pulse through the optical cell, and a second measurement is taken. In the presence of CO_2 , absorption reduces the signal increase. Subtracting the dark-current measurement from the illuminated measurement yields a detector signal that is proportional to the CO_2 number density (Eq. A1). This technique, known as correlated double sampling (CDS), is illustrated in Fig. A1 using oscilloscope measurements of the sensor. Information on the exact integration time, readout logic, and path length of the optical cell is not publicly available.

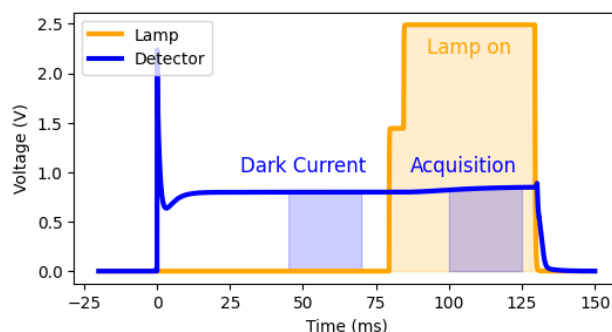


Figure A1. Measured detector and lamp voltage waveform of an LP8 NDIR CO₂ sensor. After powering the detector and a settling period, a dark current measurement is taken. Then, the lamp is powered, and an acquisition measurement is taken. The difference between the two measurements is proportional to the CO₂ number concentration c_{CO_2} along the light path between the lamp and the detector. The blue-shaded areas illustrate exemplary 20 ms intervals during which the measurements are integrated.

Appendix B: Hyperparameter sensitivity test

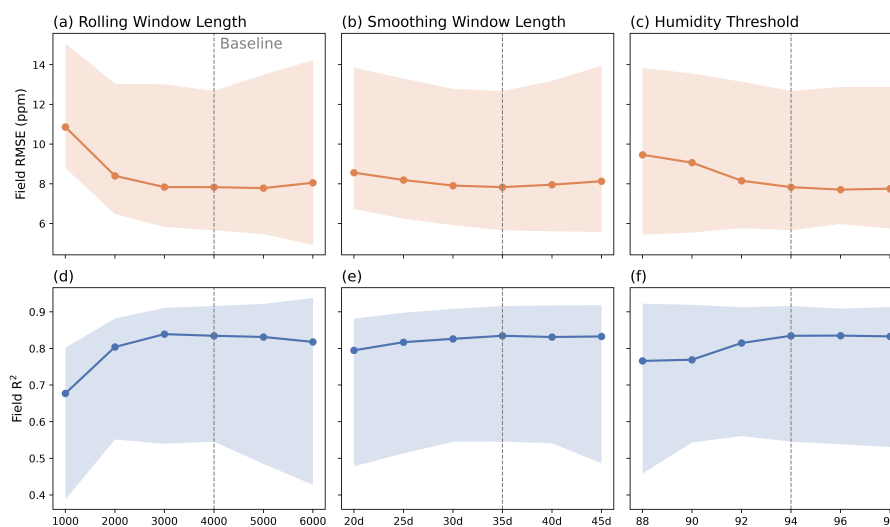


Figure B1. One-at-a-time sensitivity of the field performance to the three hyperparameters in the calibration model, evaluated across 21 sensors co-located with the reference analyzer. Each parameter is varied while the remaining two are held at their selected values (dashed lines: rolling window 4000 samples, smoothing window 35 d, humidity threshold 94 % RH). Lines show the median RMSE (top) and R^2 (bottom) across sensors. Shaded bands indicate the minimum and maximum values across sensors. All three parameters exhibit a broad plateau around the selected values, with the median RMSE varying by less than 0.5 ppm within the plateau region. Across all settings, the sensor-to-sensor variability of 6 to 8 ppm exceeds the median's variation by an order of magnitude, indicating that sensor-to-sensor variability, rather than hyperparameter selection, dominates the performance spread.



Appendix C: Overview of the training and deployment datasets

Table C1. Overview of the measurement data time periods used for sensor model training and evaluation of the sensor performance.

Time period	Site	Usage
2023-01-01 to 2023-29-03	Outdoor side-by-side for all sensors at Empa-Dübendorf (DUEB)	Sensor model training
2023-04-12 to 2023-05-01	Outdoor side-by-side for all sensors at the Technical University of Munich (TUMR)	Sensor model training
2023-05-02 to 2024-04-16	Outdoor side-by-side for 10 sensors at TUMR	Evaluation of the sensor performance; Parameterization of the rolling window and median filter length
2023-05-02 to 2025-12-31	Field deployment	Deployment to the selected locations; Sensors that stayed at the reference site TUMR for more than 3 months are included in the sensor evaluation.

Appendix D: Inside the radiation shield

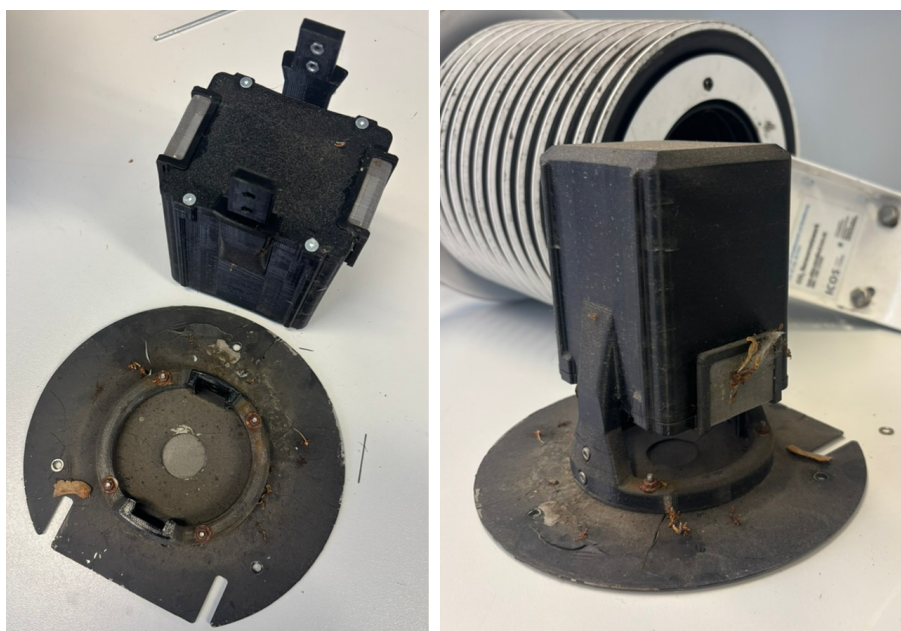


Figure D1. Example of a contaminated sensor inside its radiation shield after approximately two years of field deployment. Dirt and particulate matter accumulated over this period may have substantially affected the sensor's data quality.



Author contributions. DK developed the methods, supported by SG, PR, LE, DB and JC. PA, AW, DB and JC provided extensive feedback in discussions on the methods. DK carried out the network setup, with support from AW and PA. DK carried out the network operations and implementation of all methods presented. PA, AW, and JC have carried out the preliminary steps of the project. DK wrote the manuscript, and all authors reviewed the manuscript. JC conceived the concept and supervised the project as the project PI. JC and LE acquired the funding.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Atmospheric Measurement Techniques. The authors declare no other competing interests.

Code and data availability. Code and computational notebooks are available on Zenodo (Kühbacher and Chen, 2026b). Processed, quality controlled datasets are published on the ICOS Cities Carbon Portal (Kühbacher and Chen, 2026a).

Disclaimer. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Acknowledgements. We kindly acknowledge the support of our technician Klaus Kürzinger, for setting up and operating the network. Furthermore, we thank all students who contributed to the work, namely Ali Ahmad Khan, Julian Bärtschi, and Eroll Selmani. We would also like to thank our colleagues Moritz Makowski and Felix Böhm for their help with software development, and Dr. Andreas Luther for his input in discussions. Furthermore, we would like to thank the LMU Meteorological Institute, and in particular Markus Garhammer, for providing access to meteorological data and to the meteorological laboratory. We also thank the Deutscher Wetterdienst Meteorological Observatory Hohenpeissenberg and the Laboratory for Climate and Environmental Sciences (LSCE) for providing a Picarro reference instrument for the duration of the project. Finally, our thanks go to the City of Munich for its ongoing and generous support of the project. We acknowledge the assistive use of the AI tools Grammarly, DeepL and Claude Code to improve the writing and assist coding tasks.

Financial support. This work has been funded by the ICOS PAUL project: PAUL, Pilot Applications in Urban Landscapes - Towards integrated city observatories for greenhouse gases (ICOS Cities), funded by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 101037319. Furthermore, the work is supported by the ERC consolidator grant CoSense4Climate (Grant no.: 101089203, PI: Jia Chen) and by the Bundesministerium für Forschung, Bildung und Raumfahrt (MCube Phase 2, project DatSim 2.0, FKZ 03ZU2105HA)



References

- Aigner, P., Chen, J., Böhm, F., Chariot, M., Emmenegger, L., Frölich, L., Grange, S., Kühbacher, D., Kürzinger, K., Laurent, O., Makowski, M., Rubli, P., Schmitt, A., and Wenzel, A.: ACROPOLIS: Munich urban CO₂ sensor network, *Atmospheric Measurement Techniques*, 19, 745–773, <https://doi.org/10.5194/amt-19-745-2026>, 2026a.
- Aigner, P., Chen, J., and Kürzinger, K.: ICOS Cities 2025-12 release mid-cost CO₂ data from the Munich network ACROPOLIS, <https://doi.org/10.18160/HWNB-7CD8>, 2026b.
- Asahi Kasei Microdevices Corporation (AKM): AK9710AEF01 – IR Sensor for NDIR CO₂ Sensing: Datasheet, <https://www.akm.com/content/dam/documents/products/gas-sensing/infrared-sensor-for-ndir-gas-sensor/ak9710aef01/ak9710aef01-en-datasheet.pdf>, last access: 12 August 2026, 2016.
- Augustin, A., Yi, J., Clausen, T., and Townsley, W.: A Study of LoRa: Long Range & Low Power Networks for the Internet of Things, *Sensors*, 16, 1466, <https://doi.org/10.3390/s16091466>, 2016.
- Bayerisches Landesamt für Umwelt: Lufthygienischer Jahreskurzbericht 2024, Lufthygienischer jahreskurzbericht, Bayerisches Landesamt für Umwelt (LfU), Augsburg, Germany, https://www.lfu.bayern.de/luft/immissionsmessungen/lufthygienische_berichte/doc/jahreskurzberichte/jk24.pdf, last access: 28 May 2026, 2025.
- Cai, Q., Zeng, N., Yang, X., Xu, C., Wang, Z., and Han, P.: A 30-Month Field Evaluation of Low-Cost CO₂ Sensors Using a Reference Instrument, 18, 4871–4884, <https://doi.org/10.5194/amt-18-4871-2025>, 2025.
- Creman, L., Grange, S. K., Rubli, P., Fischer, A., Brunner, D., Hueglin, C., Emmenegger, L., and Bernet, L.: The Zurich Low-cost CO₂ sensor network (ZiCOS-L): data processing, performance assessment and analysis of spatial and temporal CO₂ dynamics, *Atmospheric Measurement Techniques*, 19, 1441–1463, <https://doi.org/10.5194/amt-19-1441-2026>, 2026.
- Decentlab GmbH: DL-LP8P Datasheet: CO₂, Temperature, Humidity and Barometric Pressure Sensor for LoRaWAN®, Decentlab GmbH, Kriesbachstrasse 30, 8600 Dübendorf, Switzerland, <https://cdn.decentlab.com/download/datasheets/Decentlab-DL-LP8P-datasheet.pdf>, last access: 15 October 2026, 2018.
- 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 CO₂ Network: Field Calibrations of Sensor Temperature Dependence and Assessment of Network Scale CO₂ Accuracy, 14, 5487–5500, <https://doi.org/10.5194/amt-14-5487-2021>, 2021.
- Dietrich, F., Chen, J., Voggenreiter, B., Aigner, P., Nachtigall, N., and Reger, B.: MUCCnet: Munich Urban Carbon Column Network, 14, 1111–1126, <https://doi.org/10.5194/amt-14-1111-2021>, 2021.
- Dinh, T.-V., Choi, I.-Y., Son, Y.-S., and Kim, J.-C.: A review on non-dispersive infrared gas sensors: Improvement of sensor detection limit and interference correction, *Sensors and Actuators B: Chemical*, 231, 529–538, <https://doi.org/10.1016/j.snb.2016.03.040>, 2016.
- Dubey, R., Telles, A., Nikkel, J., Cao, C., Gewirtzman, J., Raymond, P. A., and Lee, X.: Low-Cost CO₂ NDIR Sensors: Performance Evaluation and Calibration Using Machine Learning Techniques, 24, 5675, <https://doi.org/10.3390/s24175675>, 2024.
- Ecosystem Thematic Centre, Holst, C., Lan, C., Jahn, C., Christen, A., Hilland, R., Stagakis, S., Nicolini, G., and Mauder, M.: ETC L2 Fluxes from Oberpostdirektion, 2022-12-31–2026-03-31, https://hdl.handle.net/11676/HyC8aZLr8BoB5JH7j_AnbnNWt, 2026.
- ELECTRO - MATION GmbH: Wetterschutzgehäuse Serie EMF 8000 – Technisches Datenblatt, <https://magento2.electromotion.powered-by-rackspeed.de/media/productattachments/files/datenblatt-wetterschutzgehaeuse-emf-8000.pdf>, last access: 26 November 2025, o.D.



- Forster, P. M., Smith, C., Walsh, T., Lamb, W. F., Lamboll, R., Cassou, C., Hauser, M., Hausfather, Z., Lee, J.-Y., Palmer, M. D., von
645 Schuckmann, K., Slangen, A. B. A., Szopa, S., Trewin, B., Yun, J., Gillett, N. P., Jenkins, S., Matthews, H. D., Raghavan, K., Ribes, A.,
Rogelj, J., Rosen, D., Zhang, X., Allen, M., Aleluia Reis, L., Andrew, R. M., Betts, R. A., Borger, A., Broersma, J. A., Burgess, S. N.,
Cheng, L., Friedlingstein, P., Domingues, C. M., Gambarini, M., Gasser, T., Gütschow, J., Ishii, M., Kadow, C., Kennedy, J., Killick, R. E.,
Krummel, P. B., Liné, A., Monselesan, D. P., Morice, C., Mühle, J., Naik, V., Peters, G. P., Pirani, A., Pongratz, J., Minx, J. C., Rigby, M.,
Rohde, R., Savita, A., Seneviratne, S. I., Thorne, P., Wells, C., Western, L. M., van der Werf, G. R., Wijffels, S. E., Masson-Delmotte, V.,
650 and Zhai, P.: Indicators of Global Climate Change 2024: Annual Update of Key Indicators of the State of the Climate System and Human
Influence, *Earth System Science Data*, 17, 2641–2680, <https://doi.org/10.5194/essd-17-2641-2025>, 2025.
- Friedlingstein, P., O’Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Li, H., Luijckx,
I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Aas, K., Alin, S. R., Anthoni, P.,
Barbero, L., Bates, N. R., Bellouin, N., Benoit-Cattin, A., Berghoff, C. F., Bernardello, R., Bopp, L., Brasika, I. B. M., Chamberlain,
655 M. A., Chandra, N., Chevallier, F., Chini, L. P., Collier, N. O., Colligan, T. H., Cronin, M., Djeutchouang, L., Dou, X., Enright, M. P.,
Enyo, K., Erb, M., Evans, W., Feely, R. A., Feng, L., Ford, D. J., Foster, A., Fransner, F., Gasser, T., Gehlen, M., Gkritzalis, T., Goncalves
De Souza, J., Grassi, G., Gregor, L., Gruber, N., Guenet, B., Gürses, Ö., Harrington, K., Harris, I., Heinke, J., Hurtt, G. C., Iida, Y., Ilyina,
T., Ito, A., Jacobson, A. R., Jain, A. K., Jarníková, T., Jersild, A., Jiang, F., Jones, S. D., Kato, E., Keeling, R. F., Klein Goldewijk, K.,
Knauer, J., Kong, Y., Korsbakken, J. I., Koven, C., Kunimitsu, T., Lan, X., Liu, J., Liu, Z., Liu, Z., Lo Monaco, C., Ma, L., Marland,
660 G., McGuire, P. C., McKinley, G. A., Melton, J., Monacci, N., Monier, E., Morgan, E. J., Munro, D. R., Müller, J. D., Nakaoka, S.-I.,
Nayagam, L. R., Niwa, Y., Nutz, T., Olsen, A., Omar, A. M., Pan, N., Pandey, S., Pierrot, D., Qin, Z., Regnier, P. A. G., Rehder, G.,
Resplandy, L., Roobaert, A., Rosan, T. M., Rödenbeck, C., Schwinger, J., Skjelvan, I., Smallman, T. L., Spada, V., Sreeush, M. G., Sun,
Q., Sutton, A. J., Sweeney, C., Swingedouw, D., Séférian, R., Takao, S., Tatebe, H., Tian, H., Tian, X., Tilbrook, B., Tsujino, H., Tubiello,
F., Van Ooijen, E., Van Der Werf, G., Van De Velde, S. J., Walker, A., Wanninkhof, R., Yang, X., Yuan, W., Yue, X., and Zeng, J.: Global
665 Carbon Budget 2025, <https://doi.org/10.5194/essd-2025-659>, 2025.
- Gordon, I., Rothman, L., Hargreaves, R., Hashemi, R., Karlovets, E., Skinner, F., Conway, E., Hill, C., Kochanov, R., Tan, Y., Weislo, P.,
Finenko, A., Nelson, K., Bernath, P., Birk, M., Boudon, V., Campargue, A., Chance, K., Coustenis, A., Drouin, B., Flaud, J., Gamache,
R., Hodges, J., Jacquemart, D., Mlawer, E., Nikitin, A., Perevalov, V., Rotger, M., Tennyson, J., Toon, G., Tran, H., Tyuterev, V., Adkins,
E., Baker, A., Barbe, A., Canè, E., Császár, A., Dudaryonok, A., Egorov, O., Fleisher, A., Fleurbaey, H., Foltynowicz, A., Furtenbacher,
670 T., Harrison, J., Hartmann, J., Horneman, V., Huang, X., Karman, T., Karns, J., Kass, S., Kleiner, I., Kofman, V., Kwabia-Tchana, F.,
Lavrentieva, N., Lee, T., Long, D., Lukashevskaya, A., Lyulin, O., family=Makhnev, given=V.Yu., g.-i., Matt, W., Massie, S., Melosso,
M., Mikhailenko, S., Mondelain, D., Müller, H., Naumenko, O., Perrin, A., Polyansky, O., Raddaoui, E., Raston, P., Reed, Z., Rey, M.,
Richard, C., Tóbiás, R., Sadiek, I., Schwenke, D., Starikova, E., Sung, K., Tamassia, F., Tashkun, S., Vander Auwera, J., Vasilenko, I.,
Vigasin, A., Villanueva, G., Vispoel, B., Wagner, G., Yachmenev, A., and Yurchenko, S.: The HITRAN2020 Molecular Spectroscopic
675 Database, 277, 107 949, <https://doi.org/10.1016/j.jqsrt.2021.107949>, 2022.
- Grange, S. K., Rubli, P., Fischer, A., Brunner, D., Hueglin, C., and Emmenegger, L.: The ZICOS-M CO₂ Sensor Network: Measurement
Performance and CO₂ Variability across Zurich, 25, 2781–2806, <https://doi.org/10.5194/acp-25-2781-2025>.
- Hall, B. D., Crotwell, A. M., Kitzis, D. R., Mefford, T., Miller, B. R., Schibig, M. F., and Tans, P. P.: Revision of the World Meteorological
Organization Global Atmosphere Watch (WMO/GAW) CO₂ calibration scale, *Atmospheric Measurement Techniques*, 14, 3015–3032,
680 <https://doi.org/10.5194/amt-14-3015-2021>, 2021.



- Hu, H., Geng, G., Xu, R., Liu, Y., Shi, Q., Xiao, Q., Liu, X., Zheng, B., Zhang, Q., and He, K.: Notable uncertainties in near real-time CO₂ emission estimates in China, *npj Climate and Atmospheric Science*, 8, 108, <https://doi.org/10.1038/s41612-025-00991-4>, 2025.
- ICOS: ICOS Cities, i.e. Pilot Applications in Urban Landscapes (PAUL) – Towards integrated city observatories for greenhouse gases, <https://www.icos-cp.eu/media/265>, last access: 13 January 2025, 2025.
- 685 IPCC: Sections, in: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Core Writing Team, Lee, H., and Romero, J., pp. 35–115, IPCC, Geneva, Switzerland, <https://doi.org/10.59327/IPCC/AR6-9789291691647>, 2023.
- Ke, P., Deng, Z., Zhu, B., Zheng, B., Wang, Y., Boucher, O., Arous, S. B., Zhou, C., Andrew, R. M., Dou, X., Sun, T., Song, X., Li, Z., Yan, F., Cui, D., Hu, Y., Huo, D., Chang, J.-P., Engelen, R., Davis, S. J., Ciais, P., and Liu, Z.: Carbon Monitor Europe near-real-time daily CO₂ emissions for 27 EU countries and the United Kingdom, *Scientific Data*, 10, 374, <https://doi.org/10.1038/s41597-023-02284-y>, 2023.
- 690 Kubistin, D., Plaß-Dülmer, C., Arnold, S., Kneuer, T., Lindauer, M., Müller-Williams, J., Schumacher, M., Zwerschke, E., and ICOS RI: ICOS Atmosphere Level 2 data, Hohenpeissenberg, release 2026-1, <https://doi.org/10.18160/Z7YY-7C8V>, 2026.
- Kühbacher, D. and Chen, J.: ICOS Cities 2025-12 release of quality-controlled CO₂ data from the Munich network SCOUT, <https://doi.org/10.18160/FEKH-8RPB>, 2026a.
- 695 Kühbacher, D. and Chen, J.: SCOUT Data Processing and Analytics [Computer software], <https://doi.org/https://doi.org/10.5281/zenodo.21930851>, 2026b.
- Kühbacher, D., Aigner, P., Chen, J., Hinderer, J., Super, I., Denier van der Gon, H., Dröge, R., Schoenmakers, E., and Hohenberger, T.: Gridded anthropogenic emissions of CO₂, CH₄ and co-emitted pollutants for the city of Munich for the period 2019-2024, <https://doi.org/10.18160/2K5S-967C>, 2025.
- 700 Kunz, A.-K., Aigner, P., Becker, C., Borchardt, L., Chen, J., Christen, A., Della Coletta, J., Erritt, M., Gutiérrez, X., Hashemi, J., Hilland, R., Holst, C., Jahn, C., Jordan, A., Kneißl, R., Kürzinger, K., Lan, C., Legendre, V., Levin, I., Mauder, M., Preunkert, S., and Hammer, S.: REA Flux Time Series Archive (L2) from Oberpostdirektion, <https://doi.org/10.18160/98S6-A1N8>, 2025.
- Lan, C., Mauder, M., Ruiz Moncada, A., and Mittempergher, B. G.: Biogenic In-Situ Observations (L2), Munich PAR, 2023-03-23–2024-12-31, <https://doi.org/10.18160/N7D3-EW9F>, 2025.
- 705 Li, J., Chen, J., Ruiz Moncada, A., and Mittempergher, B. G.: Biogenic In-Situ Observations (L2), Munich LAI, 2024-04-08–2024-08-14, <https://doi.org/10.18160/6ARV-14R6>, 2025a.
- Li, J., Chen, J., Ruiz Moncada, A., and Mittempergher, B. G.: Biogenic In-Situ Observations (L2), Munich Soil Respiration, 2024-03-26–2024-11-27, <https://doi.org/10.18160/XJAR-DNRX>, 2025b.
- Li, J., Chen, J., Brunner, D., Öttl, D., Holst, C. C., and Tang, H.: Urban CO₂ Emission Assessment Based on High-Resolution Dispersion Simulations and MCMC Based Inversion, <https://doi.org/10.5194/egusphere-egu26-17218>, 2026a.
- 710 Li, J., Chen, J., Glauch, T., Brunner, D., Marshall, J., Ponomarev, N., Tang, H., and Stagakis, S.: Fine-Scale Estimation of Urban Biogenic CO₂ Fluxes: A Novel Framework Integrating Multiple Versions of Vegetation Photosynthesis and Respiration Models and In Situ Measurements, *Earth's Future*, 14, e2025EF007458, <https://doi.org/https://doi.org/10.1029/2025EF007458>, e2025EF007458 2025EF007458, 2026b.
- 715 Lian, J., Lauvaux, T., Utard, H., Bréon, F.-M., Broquet, G., Ramonet, M., Laurent, O., Albarus, I., Chariot, M., Kotthaus, S., Haefelin, M., Sanchez, O., Perrussel, O., Denier van der Gon, H. A., Dellaert, S. N. C., and Ciais, P.: Can we use atmospheric CO₂ measurements to verify emission trends reported by cities? Lessons from a 6-year atmospheric inversion over Paris, 23, 8823–8835, <https://doi.org/10.5194/acp-23-8823-2023>, 2023.



- LoRa Alliance ®: About LoRaWAN ®, <https://loro-alliance.org/about-lorawan/>, last access: 24 November 2025, 2025.
- 720 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 CO₂ measurements from a low-cost NDIR sensor, 10, 2383–2395, <https://doi.org/10.5194/amt-10-2383-2017>, 2017.
- Matsumoto, T., Siripurapu, S., Lee, D., and McCann, P.: An integrated approach to the Paris Climate Agreement: The role of regions and cities, Tech. Rep. OECD Regional Development Working Papers 2019/06, OECD, Paris, <https://doi.org/10.1787/96b5676d-en>, 2019.
- Mueller, K. L., Lauvaux, T., Gurney, K. R., Roest, G., Ghosh, S., Gourdji, S. M., Karion, A., DeCola, P., and Whetstone, J.: An Emerging
 725 GHG Estimation Approach Can Help Cities Achieve Their Climate and Sustainability Goals, 16, 084 003, <https://doi.org/10.1088/1748-9326/ac0f25>, 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 Switzerland, 13, 3815–3834, <https://doi.org/10.5194/amt-13-3815-2020>, 2020.
- 730 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, <http://jmlr.org/papers/v12/pedregosa11a.html>, 2011.
- Schneider, P., Bartonova, A., Castell, N., Dauge, F. R., Gerboles, M., Hagler, G. S., Hüglin, C., Jones, R. L., Khan, S., Lewis, A. C., Mijling, B., Müller, M., Penza, M., Spinelle, L., Stacey, B., Vogt, M., Wesseling, J., and Williams, R. W.: Toward a Unified Terminology of
 735 Processing Levels for Low-Cost Air-Quality Sensors, 53, 8485–8487, <https://doi.org/10.1021/acs.est.9b03950>, 2019.
- Seabold, S. and Perktold, J.: Statsmodels: Econometric and Statistical Modeling with Python, in: Proceedings of the 9th Python in Science Conference, edited by Stéfan van der Walt and Jarrod Millman, pp. 92 – 96, <https://doi.org/10.25080/Majora-92bf1922-011>, 2010.
- Senseair AB: LP8 Datasheet, Senseair AB, <https://rmtplusstoragesenseair.blob.core.windows.net/docs/Dev/publicerat/PSH1233.pdf>, last access: 12 March 2026.
- 740 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, 18, 13 773–13 785, <https://doi.org/10.5194/acp-18-13773-2018>, 2018.
- Ueno, K., Camargo, E. G., Katsumata, T., Goto, H., Kuze, N., Kangawa, Y., and Kakimoto, K.: InSb Mid-Infrared Photon Detector for Room-Temperature Operation, 52, 092 202, <https://doi.org/10.7567/JJAP.52.092202>, 2013.
- UNEP: Emissions Gap Report 2025: Off Target - Continued Collective Inaction Puts Global Temperature Goal at Risk, United Nations
 745 Environment Programme, ISBN 978-92-807-4239-8, <https://doi.org/10.59117/20.500.11822/48854>, 2025.
- Wagner, W. and Pruß, A.: The IAPWS Formulation 1995 for the Thermodynamic Properties of Ordinary Water Substance for General and Scientific Use, 31, 387–535, <https://doi.org/10.1063/1.1461829>, 2002.