Response to Referee #1

We would like to thank reviewer #1 for taking the time to review this manuscript and for providing valuable, constructive feedback and corresponding suggestions that helped us to further improve the manuscript.

In this author's comment, all the points raised by the reviewer are copied here one by one and shown in black color, along with the corresponding reply from the authors in blue.

This manuscript presents a short-term measurement campaign in Xining, a city located on the eastern edge of the Qinghai-Tibetan Plateau (QTP), using a portable FTIR instrument (EM27/SUN) to retrieve column-averaged concentrations of CO_2, CH_4, and CO. The study touches on several topics, including satellite validation for CH_4 and CO, CAMS product evaluation, combustion efficiency derived from the CO:CO_2 ratio, and CO_2 emissions estimation. However, the manuscript lacks a cohesive narrative and frequently shifts between topics without adequately developing or concluding each one. As a result, it reads more like a collection of loosely connected sub-studies rather than a focused, hypothesis-driven investigation.

We are grateful to the referee for this insightful comment, which we acknowledge is fundamental to improving the manuscript. We have thoroughly revised the paper to address the lack of a cohesive narrative and to better integrate the sub-studies.

And the study does not offer significant methodological innovation or new scientific insights and lacks discussion part. The only potentially unique aspect is the absence of previous atmospheric column observations in the suburban area of Xining city. However, this alone does not justify publication unless the authors can thoroughly address the concerns outlined above through major revisions.

We appreciate the referee's feedback and the concerns raised regarding the originality and scientific contribution of the study. We believe that our study presents a novel approach for estimating CO_2 emissions by combining collaborative ground-based and space-based observations. While TROPOMI offers high spatial coverage for CO measurements,TROPOMI does not provide CO_2 data, and satellites like OCO-2/3 or GOSAT have limited coverage in this region. The use of the EM27/SUN instrument, which probes both columnar CO and CO_2 concentrations in the study area, is a key strength, as it allows for a more direct linkage between CO and CO_2 emissions. Furthermore, the ground-based measurements offer a valuable dataset for validating satellite observations and improving model accuracy especially for this specific region, where satellite coverage is limited and which is difficult from the viewpoint of satellite measurements, as it is orographically complex.

We have emphasized this unique advantage and the potential of the proposed methodology to improve our understanding of urban emission sources in the revised manuscript. Furthermore, we have expanded the discussion section to address the broader scientific implications and how this combined approach could contribute to more accurate carbon cycle studies in regions with limited satellite data coverage.

Major Comments:

1. The observational period was very short, only 8 days in early June 2024, but the rationale for selecting this specific timeframe is unclear. Was there a particular emission event or atmospheric condition of interest during this period? The motivation for the campaign is not well explained. Additionally, the measurements were conducted in a suburban area, but it is not clear whether this site is representative, whether the data can inform future carbon cycle studies, or what the broader scientific significance is. Summer conditions are typically associated with various interfering factors, yet these are neither acknowledged nor discussed in the manuscript.

Thank you for the valuable comment. The choice of early summer for the observational period was intentional, as it helps minimize the influence of heating emissions from the cold season, which typically lead to higher CO emissions due to residential and industrial combustion. These emissions are difficult to separate from those of transportation and other sources during colder months. By selecting the early summer period, we aim to focus on emissions that are more representative of typical urban activity, without the confounding influence of winter heating.

Additionally, the EM27/SUN measurement site was located to the west of the city to capture emissions across the entire urban area, as the predominant wind direction in this region is easterly. This placement allows for a more comprehensive representation of the city's emission profile.

2. The manuscript aims to evaluate satellite retrievals using ground-based observations. In Section 2.2, a detailed description of the general COCCON product is provided, but there is almost no information about the specific EM27/SUN instrument used in this campaign. Key details such as the instrument's stability before and after the measurement period, the configuration of retrieval parameters, and whether any calibration was performed using TCCON, AirCore, or aircraft measurements are missing. Additionally, the measurement uncertainty is not discussed. As the ground-based observations serve as the reference for satellite validation, it is essential to present their accuracy and reliability clearly, rather than focusing only on general background information.

We thank the reviewer for highlighting the need for a more detailed description of the performance and quality assurance of EM27/SUN. In response, we have substantially expanded Section 2.2 of the manuscript to provide a comprehensive account of the specific EM27/SUN spectrometer used in this campaign, directly addressing the points raised concerning instrumental stability, retrieval configuration, calibration, and measurement uncertainty.

The EM27/SUN spectrometer used in this study operates within the COCCON network, whose instruments are well-established as robust and reliable through numerous international field campaigns (Butz et al., 2017; Frey et al., 2019; Klappenbach et al., 2015; Sha et al., 2025; Tu et al., 2020). Following the standard COCCON protocol, our specific instrument was calibrated prior to the campaign against the TCCON reference spectrometer at KIT. This calibration ensures accuracy by establishing and applying instrument-specific calibration factors for each target gas species (Frey et al., 2019).

To continuously monitor the instrument's stability and characterize its performance—key to assessing measurement uncertainty—we employed the two primary quality assurance methods proposed by the COCCON community: the Instrumental Line Shape (ILS) and the XAIR ratio. The ILS, which characterizes the spectrometer's spectral response function, is determined through laboratory open-path measurements using a nonlinear least-squares spectral fitting algorithm (Frey

et al., 2015; Hase et al., 1999). The XAIR ratio, an indicator of overall instrumental stability, is derived from the measured vertical columns of O_2 and H_2O in conjunction with surface pressure data (Alberti et al., 2022). Systematic monitoring of these parameters is critical for detecting any deviations from the expected instrumental performance.

Specifically, an ILS characterization was performed in December 2024. The results indicated a minor change: one key parameter, the modulation efficiency (ME), had decreased by approximately 1.5% compared to its baseline value measured at KIT. This deviation falls well within the accepted uncertainty range for such instruments. We attribute this small change primarily to mechanical stresses incurred during the long-distance shipment from Germany to China and subsequent domestic transports. A potential minor contribution from systematic errors associated with the specific light source and lens used in the ILS setup in China cannot be entirely ruled out. A follow-up ILS measurement in September 2025 showed no change in the ME value, confirming the instrument's stability.

Additionally, the XAIR ratio served as an independent indicator of instrumental performance throughout the measurement period. The mean XAIR value during the entire field campaign was 1.0012 ± 0.0024. Values consistently close to 1.0 signify highly stable instrument operation. This stability is further underscored by the nearly identical XAIR values measured before (1.0005) and after (1.0005) the campaign, indicating no significant instrumental drift.

In summary, we believe that the continuous monitoring of both ILS and XAIR provides a robust framework for quality assurance, effectively verifying the stability and reliability of the EM27/SUN spectrometer during the campaign. This approach is particularly valuable for ensuring data quality in remote field deployments where frequent re-calibration against a TCCON station is not feasible. All detailed information regarding the instrument's calibration, stability, retrieval parameters, and the associated quality control procedures has been incorporated into Section 2.2 as suggested.

3. The manuscript also attempts to evaluate CAMS simulation results using ground-based observations. However, the approach raises several questions. A 20 km radius was used for CAMS product validation—but why? Since CAMS provides data at specific grid points, the rationale for selecting a 20 km averaging radius is unclear. Large-area averaging is typically applied in satellite validation to reduce observational noise and improve sampling statistics, but it is not obvious why a 20 km radius was appropriate or necessary in this case.

We thank the referee for this critical point. The initial 20 km radius was indeed arbitrary. To address this, we have revised the collocation criteria to use the CAMS data from the nearest grid cells in which the EM27/SUN instrument was located. As the instrument was situated near the border of two grid cells, we used the average value from both. The relevant figures and text have been updated throughout the manuscript.

Moreover, the analysis in this section is not sufficiently developed. The authors conclude that CAMS performs well in simulating CH_4 in this region, while its performance for CO and CO_2 is poor. But what are the broader implications of this result? Does this indicate that CAMS is better suited for CH_4 studies over the Qinghai-Tibetan Plateau? Could this support the case for establishing long-term observation sites in the area?

We thank the referee for this valuable suggestion. We have expanded the discussion in this section in the revised manuscript to address these findings:

"When comprising the absolute column amounts, CAMS shows an approximately 0.4% underestimation for XCO_2 and 0.2% for XCH_4 relative to the COCCON data. However, the underestimate of CO is more pronounced, with a bias of 35%.

CAMS XCH_4 compares much better with COCCON observations than XCO_2 and CO, as the higher value of R^2 and the better agreement of the slope of the regression line indicates. We conclude that the variability of XCO and XCO_2 dominating the variability as detected by the ground-based observation is generated on a smaller spatial scale which is not properly resolved in the simulation. In contrast, the variability of the CH_4 model field seems to be dominated by extended sources distributed in a wider area, which therefore can be properly depicted by the model.

In terms of broader implications, the relatively strong performance of CAMS in simulating CH_4, especially when compared to CO and CO_2, indicates that CAMS may be more reliable for CH_4 studies in such regions. This finding could potentially support the case for establishing long-term observation sites in this area to help with satellite validation and improve model accuracy, particularly for species like CO, which appear to require more refined simulations."

In addition, satellite retrievals usually offer a higher number of soundings than ground-based instruments. How does the CAMS product compare with satellite observations in terms of coverage and consistency? Do the conclusions drawn from CAMS agree with those from the COCCON dataset? These questions are not clearly addressed in the manuscript.

We appreciate the referee's valuable comment regarding the comparison between CAMS and satellite observations. To address this, we regridded TROPOMI data to the same spatial resolution as the CAMS GHG forecast over the study area. The figures below shows the correlation between these two datasets.

In general, CAMS agrees well with TROPOMI CO data, with CAMS showing approximately 3% higher CO amounts compared to TROPOMI. However, larger biases are evident over the northern and eastern parts of the study area, which correspond to the main valley region (see Figure 1-right below and Figure A5 in the manuscript). These discrepancies may be due to overestimation in the CAMS simulations or potential underestimation in the TROPOMI observations in the complex terrain of the region. This also highlights the importance of establishing long-term ground-based observations in such regions to better constrain model outputs and support satellite validation.

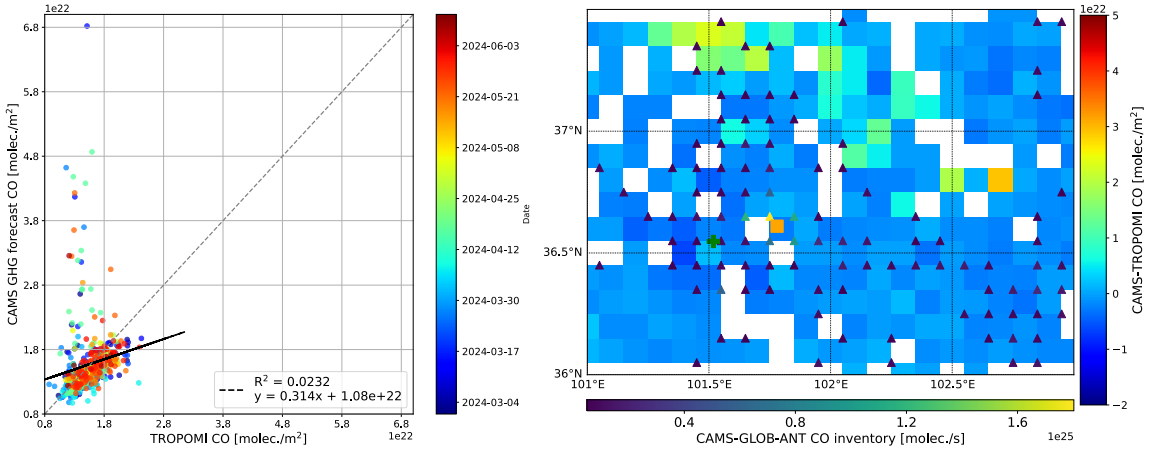

Figure 1 left: correlation between CAMS and TROPOMI for CO, right: spatial distribution of difference in columnar CO amounts between CAMS and TROPOMI.

For collocated XCH_4, CAMS shows good agreement with TROPOMI, with a bias of -0.18%. However, the availability of TROPOMI XCH_4 data in this region is limited (see Figure 2 right), with only about 3000 observations collected over five years. This relatively small dataset also makes it challenging to estimate CH_4 emissions accurately from the TROPOMI dataset in this region.

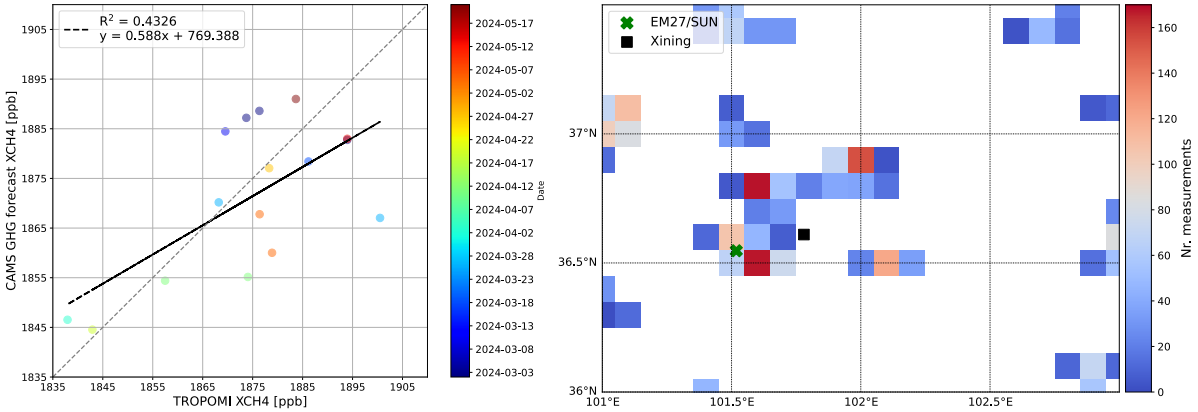

Figure 2 left: similar to Figure 1-left, but for XCH_4, right: total number of measurements in each grid during May 2018 – May 2024.

4. The manuscript estimates CO_2 emissions using CO fluxes derived from TROPOMI and EM27/SUN observations. However, several points need clarification. First, the emission estimation appears to be based on multiplying values observed by wind speed. If so, does this method account for transport processes and particle dilution along the plume? A brief explanation of this approach in the Methods section would be helpful. In addition, how does the CO:CO_2 ratio used in this study compare with values reported in emission inventories? This could be further discussed, especially in relation to combustion efficiency and source attribution.

Moreover, there appears to be an inconsistency in logic. In the CAMS evaluation, CO and CO_2 simulations were shown to perform poorly, while CH_4 agreed well with observations. However, in this section, the calculated CO emissions match CAMS values, and CO_2 emissions align closely with inventory estimates. Given the earlier performance issues with CO and CO_2 in CAMS, this raises questions about the reliability of the derived fluxes. What about CH_4 emissions in this context? Without addressing this discrepancy, the conclusions are difficult to reconcile.

Thank the referee for these comments.

(1) The method considers the transport along the plume. We have added the explanation of this approach in Section 2.6:

"2.6 Dispersion model and wind-assigned anomaly method

For a single point source, the total emission is calculated by multiplying the measured total column enhancement (ΔCO) by the area of the affected plume (Babenhauserheide et al., 2020). This plume area is modeled as an evenly distributed cone, representing the long-term averaged dispersion (Tu et al., 2022a). The relationship is given by the following equation:

$$\varepsilon = \Delta CO \times d \times v \times \partial \qquad Eq.\ 1$$

where ΔCO represents the enhanced CO column observed at the downwind site, d is the distance from the source to the measuring site and v is the wind speed.

To estimated averaged emissions from satellite observations over a region, the wind-assigned method was applied (Tu et al., 2022a, 2022b, 2023, 2024b). This technique fits the anomalies between the satellite observations and the dispersion model by analyzing enhancements under opposing wind sectors. Specifically, the wind-assigned anomaly is defined as the difference in observed enhancements between two opposite wind fields (e.g., E: 0°–180° and W: 180°–360°). A key advantage of this approach is that it inherently eliminates the uncertainty associated with background concentration calculations for long-lived gases like CO, thereby significantly improving the reliability of the resulting emission estimates."

(2) A linear fit to the EM27/SUN observations yields a $\Delta XCO/\Delta XCO_2$ slope of 0.035 ppb/ppb. This observed enhancement ratio is compared to the CAMS emission inventory ratio of 0.021 for CO/CO_2 over the same region. In this study, we find that the CAMS model underestimates the forecast CO column by an average factor of 1.6127 compared to the ground-based observations. Compensating for this bias by scaling the CAMS CO emission rate increases the emission ratio to 0.034, which aligns closely with the observed enhancement ratio (0.035 ppb/ppb). We have added this ratio comparison to the manuscript.

Additionally, we have incorporated the following explanation regarding combustion into the manuscript in section 3.2:

"CO emissions from vehicle exhaust, a major contributor to air pollution, is closely related to fossil fuel combustion (Naus et al., 2018). Gao et al. (2025) reported that CO emissions increased significantly with altitude, observing nearly twice the emission levels in Xining (2320 m) compared to those at an altitude of 20 m. This trend can be attributed to the decline in atmospheric pressure and air density at higher elevations (Fattah et al., 2019). Under such conditions, engines draw in less air per cycle, which alters the air-fuel ratio and leads to suboptimal combustion. As altitude rises, the excess air ratio decreases because more diesel fuel is injected into the cylinders, but less air is captured per cycle. Therefore, the combustion between fresh air and diesel fuel becomes incomplete, resulting in substantial CO emissions."

(3) We thank the referee for raising this important point regarding the discrepancy between CO and CO_2 emissions in the CAMS inventory and the column-based observations.

When directly comparing columnar gas ratio between COCCON and CAMS, we find that the ratios for XCO_2, XCH_4, and CO are 1.0042, 1.0022, and 1.6127, respectively. As discussed in (2), adjusting the CAMS CO emission rate by a factor of 1.6127 compensates for the discrepancy in CO, resulting in an emission ratio between CO and CO_2 that aligns more closely with the column ratio observed by the EM27/SUN measurements.

It is also important to note that the calculated CO emission rate is based on TROPOMI XCO data, which is approximately 33% underestimated compared to the EM27/SUN measurements. The average ratio between EM27/SUN and TROPOMI is 1.54, similar to that between EM27/SUN and CAMS. Adjusting for this underestimation also brings the CO/CO_2 emission ratio more in line with the column ratio observed by EM27/SUN.

As discussed in Comment #3, the limited TROPOMI XCH_4 data in this region makes it challenging to estimate emission rates accurately. Furthermore, the correlations between XCH_4 and XCO_2 or XCO from the ground-based observations are relatively weak (see Figure 3), complicating the estimation of CH_4 emissions based on these correlations. Longer observation periods may help improve these correlations and provide more reliable CH_4 emission estimates.

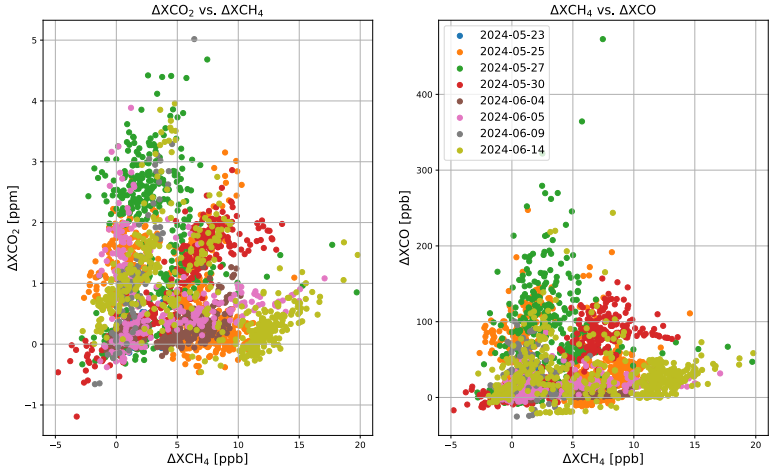

Figure 3 correlation between ΔXCO and ΔXCH_4 (left), and between ΔXCH_4 and ΔXCO (right).

Minor Comments:

Line 53–55: I still do not fully understand the claim that surface observations are influenced by surface exchange but limit the ability to estimate sources and sinks. Given that surface measurements are sensitive to near-surface fluxes, wouldn't they actually be more effective for detecting local sources and sinks? This statement needs clarification.

The referee is correct in noting that surface observations are more sensitive to near-surface fluxes and are therefore effective in representing local sources and sinks, particularly at smaller spatial scale. In contrast, column-based measurements, such as those from FTIR observations, are better suited for capturing source and sink information over broader, regional areas. We acknowledge that the original wording may have been misleading, and have revised the sentences accordingly:

"These stations conduct in-situ measurements and provide highly accurate surface observations and valuable insights into local fluxes; however, they are influenced by surface exchanges and vertical transport, which can limit their ability to estimate sources and sinks over larger spatial scales. When combined with other observation types, such as FTIR, which capture emissions and transport on a broader scale, they become complementary, together offering a more comprehensive understanding of sources and sinks at both local and regional levels (Callewaert et al., 2022)."

Line 161: This point raises concerns. The measurement period spans only 8 days, while coal mining activity is often highly episodic. As shown in the study by [Author] (https://doi.org/10.1016/j.ecolind.2023.110454), regions similar to the study area in Qinghai are known to have significant coal mining emissions. Therefore, concluding that there is no such influence based solely on this short observational window appears premature.

We thank the referee for raising this important point. As suggested, we have removed the premature conclusion from line 161 and have added a discussion on this limitation and the cited the abovementioned literature in the revised manuscript.

"*The average observed XCH_4 concentration during the measurement period was 1898.45 ± 6.66 ppb. Notably, 300 km to the northwest of the study site lies the largest Muli coalfield in Qinghai Province, which has an estimated coal reserve of around 4 billion tons (Xiao et al., 2023). However, the maximum XCH_4 enhancement observed was approximately 10 ppb on May 25, when the wind was from northwest before noon. This enhancement is relatively lower than those reported in other coal fields, such as in Changzhi, Shanxi Province (Tu et al., 2024a), suggesting that the ground-based observations may not have fully captured the methane emissions. The modest enhancement is likely due to the relatively greater distance from the source, combined with potentially low coal mining activity during the observing period, as coal production is often highly episodic."*

Butz, A., Dinger, A. S., Bobrowski, N., Kostinek, J., Fieber, L., Fischerkeller, C., Giuffrida, G. B., Hase, F., Klappenbach, F., Kuhn, J., Lübcke, P., Tirpitz, L., and Tu, Q.: Remote sensing of volcanic CO2, HF, HCl, SO2, and BrO in the downwind plume of Mt. Etna, Atmos. Meas. Tech., 10, 1–14, https://doi.org/10.5194/amt-10-1-2017, 2017.

Frey, M., Sha, M. K., Hase, F., Kiel, M., Blumenstock, T., Harig, R., Surawicz, G., Deutscher, N. M., Shiomi, K., Franklin, J. E., Bösch, H., Chen, J., Grutter, M., Ohyama, H., Sun, Y., Butz, A., Mengistu Tsidu, G., Ene, D., Wunch, D., Cao, Z., Garcia, O., Ramonet, M., Vogel, F., and Orphal, J.: Building the COllaborative Carbon Column Observing Network (COCCON): long-term stability and ensemble performance of the EM27/SUN Fourier transform spectrometer, Atmos. Meas. Tech., 12, 1513–1530, https://doi.org/10.5194/amt-12-1513-2019, 2019.

Klappenbach, F., Bertleff, M., Kostinek, J., Hase, F., Blumenstock, T., Agusti-Panareda, A., Razinger, M., and Butz, A.: Accurate mobile remote sensing of XCO2 and XCH4 latitudinal transects from aboard a research vessel, Atmos. Meas. Tech., 8, 5023–5038, https://doi.org/10.5194/amt-8-5023-2015, 2015.

Sha, M. K., Das, S., Frey, M. M., Dubravica, D., Alberti, C., Baier, B. C., Balis, D., Bezanilla, A., Blumenstock, T., Boesch, H., Cai, Z., Chen, J., Dandocsi, A., Mazière, M. D., Foka, S., García, O., Gillespie, L. D., Gribanov, K., Gross, J., Grutter, M., Handley, P., Hase, F., Heikkinen, P., Humpage, N., Jacobs, N., Jeong, S., Karppinen, T., Kiel, M., Kivi, R., Langerock, B., Laughner, J., Lopez, M., Makarova, M., Mermigkas, M., Morino, I., Mostafavipak, N., Nemuc, A., Newberger, T., Ohyama, H., Okello, W., Osterman, G., Park, H., Pirloaga, R., Pollard, D. F., Raffalski, U., Ramonet, M., Sepúlveda, E., Simpson, W. R., Stremme, W., Sweeney, C., Taquet, N., Topaloglou, C., Tu, Q., Warneke, T., Wunch, D., Zakharov, V., and Zhou, M.: Fiducial Reference Measurements for Greenhouse Gases (FRM4GHG): Validation of Satellite (Sentinel-5 Precursor, OCO-2, and GOSAT) Missions Using the COllaborative Carbon Column Observing Network (COCCON), Remote Sensing, 17, 734, https://doi.org/10.3390/rs17050734, 2025.

Tu, Q., Hase, F., Blumenstock, T., Kivi, R., Heikkinen, P., Sha, M. K., Raffalski, U., Landgraf, J., Lorente, A., Borsdorff, T., Chen, H., Dietrich, F., and Chen, J.: Intercomparison of atmospheric CO_2 and CH_4 abundances on regional scales in boreal areas using Copernicus Atmosphere Monitoring Service (CAMS) analysis, COllaborative Carbon Column Observing Network (COCCON) spectrometers, and Sentinel-5 Precursor satellite observations, Atmospheric Measurement Techniques, 13, 4751–4771, https://doi.org/10.5194/amt-13-4751-2020, 2020.