Response to the reviewers

Dear Editor and Reviewers,

We sincerely appreciate the valuable and constructive comments you have provided. In response, we have carefully revised the manuscript to address the comments and suggestions. We have responded to each point in detail to clarify how each concern has been addressed. Your insightful feedback has significantly improved the clarity, rigor, and overall quality of this manuscript. For ease of review, the original comments are presented in italics, our responses are shown in blue font, and the corresponding revisions in the manuscript are marked in red.

Response to Reviewer 1:

Major comments

Comment 1: *In Section 1, CO and CH4 are largely introduced independently. However, these two species are chemically and dynamically coupled through oxidation pathways (e.g., OH consumption), emission sources (e.g., combustion, biomass burning), and atmospheric transport. Since the study emphasizes co-monitoring and joint reconstruction, the introduction should explicitly discuss the scientific rationale for simultaneous reconstruction of XCO and XCH4. Clarifying this interaction would significantly strengthen the conceptual foundation of the work.*

Response 1: Thank you for this valuable suggestion. We substantially revised the Introduction to better clarify the scientific rationale for the simultaneous reconstruction of XCO and XCH_4. Specifically, we added a new paragraph emphasizing the chemical and dynamical coupling between CO and CH_4 through shared oxidation pathways, common emission sources, and atmospheric transport processes. The relevant text has been added to Lines 57–72 of the revised manuscript.

Revised text:

Although CO and CH_4 are often discussed separately in terms of their sources and impacts, they are chemically and dynamically coupled in the atmosphere. Both gases are predominantly removed by the hydroxyl radical (OH), meaning that changes in one can affect the atmospheric lifetime and concentration of the other through OH competition (Lelieveld et al., 2016). Moreover, they share major emission sources, particularly biomass burning and fossil-fuel-related activities such as incomplete combustion and flaring, where they can be emitted simultaneously (Andreae and Merlet, 2001; Akagi et al., 2011; Andreae, 2019). Their co-emission and co-transport imply that the spatial patterns of CO and CH_4 may be correlated near fire hotspots, industrial regions, and other combustion-influenced areas. Therefore, joint reconstruction of XCO and XCH_4 offers two key advantages: (1) shared spatial structure and common meteorological drivers can provide complementary constraints for deep learning models, improving generalization in data-sparse regions; (2) simultaneous monitoring can support more consistent estimation of emission ratios and source attribution, as demonstrated by studies using CH_4/CO or related trace-gas ratios to constrain biomass-burning and fossil-fuel methane emissions (Worden et al., 2013; Worden et al.,

2017; Ramsden et al., 2022). A comprehensive understanding of these coupled gases is fundamental to improving air quality predictions and formulating climate mitigation policies.

Comment 2: *The introduction does not clearly define the current limitations in satellite data fusion and gap-filling studies. Specifically: (1) What is the current state-of-the-art spatial and temporal resolution? (2) What level of continuity (temporal or spatial) is typically achieved? (3) What limitations remain in existing model–satellite fusion frameworks? Without a clear statement of the unresolved problem, the novelty and advancement of this work are difficult to assess. The authors should explicitly articulate the existing methodological gap and quantify how their approach advances beyond prior studies.*

Response 2:

Thank you for this insightful comment. We agree that the Introduction should more clearly define the current methodological limitations and the specific advancement of our study. In the revised manuscript, we have rewritten the relevant part of the Introduction to summarize existing gap-filling and fusion studies in terms of spatial resolution, temporal continuity, and remaining methodological gaps. The relevant text has been added to lines 95–133 of the revised manuscript.

Revised text:

The generation of continuous TROPOMI XCO and XCH_4 products has been addressed through a range of gap-filling approaches. First, machine learning-based interpolation methods (X. Chen et al., 2022; Hu et al., 2022; Valerio et al., 2025; Wei et al., 2025) reconstruct missing observations by learning nonlinear spatiotemporal relationships. However, their performance strongly depends on the quality and representativeness of the training data, and biased or insufficient samples may lead to unreliable predictions. Second, enhanced spectral fitting algorithms (Borsdorff et al., 2019; Guanter et al., 2015; Schneising et al., 2023; Wang et al., 2020) improve retrieval continuity by refining the spectral inversion process, for example through higher-order polynomial fitting to reduce reflectance-induced biases, albeit at increased computational cost. Third, physics–data–driven coupling approaches, including model fusion, data assimilation, and hybrid machine learning frameworks, integrate satellite retrievals with chemical transport models or reanalysis datasets (Fritz et al., 2022; Schneising et al., 2023; Sicard et al., 2021). By combining physical constraints with data-driven learning or observational corrections, these methods improve spatial consistency and temporal continuity while reducing retrieval bias. However, their performance remains dependent on uncertainties in prior model simulations and the representativeness of assimilated observations (Inness et al., 2022; Wang et al., 2023).

Despite these advances, current methods face challenges in balancing computational efficiency with physical consistency, particularly in regions with persistent cloud cover. To overcome these limitations, this study proposes a novel signal-domain guided spatio-temporal fusion framework. Our core objective is to generate daily global and regional continuous XCO and XCH_4 products (2019–2023) at high resolution (0.25° globally and 0.05° over China) by effectively leveraging complementary information from chemical transport modeling and frequency-domain representations.

The proposed method specifically addresses the shortcomings of existing approaches in three key aspects: (1) Addressing Data Sparsity and Non-linearity: Unlike pure machine learning interpolation that may struggle with unbalanced datasets, we employ a two-stage strategy. We first use 3D Discrete Cosine Transform (DCT) and Singular Value Decomposition (SVD) to perform low-rank signal-domain reconstruction (Rao and Yip, 2014; Wall et al., 2003). This step exploits the shared spatiotemporal structure between GEOS-Chem simulations and TROPOMI observations to approximate missing values in the frequency domain, providing a robust prior; (2) Enhancing Physical Consistency: To correct residual biases without overfitting, we implement a lightweight residual U-Net for pixel-level refinement (Ronneberger et al., 2015; Tang et al., 2022). Crucially, instead of using the TROPOMI data mask as a direct input feature, we incorporate it as a weighting mask in the loss function (Wei et al., 2022). This masked learning strategy ensures that the model is supervised exclusively by valid observations, thereby preserving the physical consistency of the DCT/SVD-derived signal-domain reconstructed fields; (3) Improving Efficiency and Accuracy: By fusing model-driven priors and observational constraints in both frequency and spatial domains, our method significantly reduces computational load compared to enhanced spectral fitting while avoiding the "black-box" limitation of pure data-driven methods.

Comment 3: *The manuscript does not describe any use of averaging kernels (AKs) when comparing or integrating TROPOMI retrievals with GEOS-Chem simulations. Given the differences in vertical sensitivity between satellite retrievals and model profiles, neglecting AK application may introduce representativeness errors.*

Response 3:

Thank you for emphasizing that averaging kernels (AKs) have an influence when comparing TROPOMI retrieval results with GEOS-Chem simulation results. Because satellite retrievals and model profiles differ in their vertical sensitivity, representativeness errors may be introduced if averaging kernels (AKs) are not considered when comparing satellite data with model results. To address this issue, we added a sensitivity analysis of the TROPOMI CO AKs in the revised manuscript. We compare both the original GEOS-Chem XCO and the AK-smoothed GEOS-Chem XCO with TROPOMI XCO (unit: ppb).

The results show that applying the TROPOMI CO column averaging kernels slightly increases the R^2 from 0.543 to 0.565, while the RMSE increases from 28.17 ppb to 28.66 ppb and the mean bias changes from −17.09 ppb to −18.34 ppb. These results indicate that AK treatment has a limited influence on the column-concentration-based fusion results.

We have added the description of the AK sensitivity test in Section 2.2.2 and included the corresponding results in the Discussion of uncertainties and limitations section. The detailed statistics are provided in Table S11 of the Supplementary Materials.The revisions in the revised manuscript are as follows:

Revised text:

2.2.2. Averaging-kernel sensitivity test

We conducted an additional sensitivity test to examine the influence of TROPOMI averaging

kernels (AKs) on the comparison between satellite retrievals and model simulations. For 2019, quality-screened TROPOMI CO Level-2 retrievals (qa_value > 0.5) were collocated with GEOS-Chem CO vertical profiles. Two GEOS-Chem XCO datasets were then generated for comparison. For the first dataset, GEOS-Chem XCO was diagnosed from the simulated CO profiles by vertically averaging the dry-air mole fractions using pressure-layer thickness as weighting, consistent with the standard post-processing procedure described above. For the second dataset, the GEOS-Chem CO profiles were first interpolated to the TROPOMI pressure layers and then smoothed using the TROPOMI CO column averaging kernels. Finally, both the original GEOS-Chem XCO and the AK-smoothed GEOS-Chem XCO were compared with TROPOMI XCO in ppb. (Lines 271-280)

AKs are an important source of uncertainty in comparisons between satellite retrievals and model simulations because they describe the vertical sensitivity of TROPOMI retrievals. To evaluate this effect, we conducted an AK sensitivity test using the 2019 CO data by comparing GEOS-Chem XCO with TROPOMI XCO before and after applying the TROPOMI CO column averaging kernels. As shown in Table S11, applying the TROPOMI CO column averaging kernels slightly increased the R^2 from 0.543 to 0.565, while the RMSE increased from 28.17 ppb to 28.66 ppb and the mean bias changed from −17.09 ppb to −18.34 ppb. These results indicate that AK treatment has a limited influence on the column-concentration-based fusion results in this experiment. (lines 800-808)

Table S11. Overall statistics of the TROPOMI CO averaging-kernel sensitivity test in 2019, with RMSE and μ in parts per billion (ppb).

Comparison	R^2	RMSE	μ
Original GEOS-Chem vs. TROPOMI	0.543	28.17	-17.09
AK-smoothed GEOS-Chem vs. TROPOMI	0.565	28.66	-18.34

Comment 4: *In Section 3.1, the validation examples appear to focus on sites with relatively dense valid TROPOMI retrievals. While this may demonstrate successful reconstruction under favorable conditions, the core objective of this study is gap filling. Therefore, performance in regions with sparse coverage (e.g., persistent cloud contamination) is more critical. The authors should: (1) Include representative low-coverage regions; (2) Quantify performance as a function of data availability; (3) Demonstrate robustness under high missing-rate conditions.*

Response 4:

We thank you for this valuable comment. We agree that, because gap filling is a core objective of this study, the validation should clearly demonstrate the performance of the fused datasets under sparse TROPOMI coverage conditions. In the original manuscript, the validation examples were already based on high-missing-rate site-day collocation datasets; however, this was not clearly described and may have caused ambiguity. In the revised manuscript, we have clarified the definition of the missing rate and the site-selection criteria. We have also included representative low-coverage regions, quantified data availability using tables, and demonstrated the robustness of the model under high-missing-rate conditions.

First, we explicitly defined the TROPOMI missing rate (MR) for each site-day collocation dataset. MR was defined as the fraction of grid cells within a 2° radius around each site that had no valid TROPOMI retrieval after quality filtering. Site-day collocation datasets with MR > 0.5, corresponding to less than 50% valid TROPOMI coverage within the collocation window, were used to evaluate the fused XCO and XCH_4 datasets under sparse TROPOMI coverage conditions. Figures 5 and 6 in the original manuscript, corresponding to Figs. 4 and 5 in the revised manuscript, were selected based on this criterion, but this was not clearly stated in the original manuscript. Therefore, these figures actually present the fusion performance under low TROPOMI coverage conditions.

Second, we calculated R^2, RMSE, and μ between the fused data and TCCON, as well as between TROPOMI and TCCON, under three missing-rate intervals: MR < 0.3, 0.3 ≤ MR ≤ 0.5, and MR > 0.5. This analysis was used to quantify the performance of the fused datasets under different levels of TROPOMI coverage. Based on the analysis of these metrics, the fused XCO and XCH_4 datasets show performance comparable to that of the TROPOMI datasets under low or moderate missing-rate conditions (MR < 0.3 and 0.3 ≤ MR ≤ 0.5). When TROPOMI retrievals are sparse (MR > 0.5), the fused datasets exhibit relatively high reliability. The corresponding results are presented in Tables S6 and S7.

Third, from MR < 0.3 to 0.3 ≤ MR ≤ 0.5 and then to MR > 0.5, the R^2 between the fused data and TCCON decreases only slightly, from 0.87 to 0.85 and then to 0.83. The overall change is small, demonstrating the robustness of our fused data under high-missing-rate conditions.

The relevant text has been added to Lines 581–590 and Lines 644–662 of the revised manuscript. The relevant figures and tables have been added as Figs. 4 and 5 in the main text and Tables S6 and S7 in the Supplement.

Revised text:

we further evaluated site-level statistical agreement using scatterplots at selected representative stations under sparse TROPOMI coverage conditions in Figs. 4 and 5. This selection was based on the number of valid collocated samples after applying the missing-rate criterion, rather than on validation performance. Specifically, for each TCCON station, we first applied the qa_value > 0.5 quality filter, the 13:30 ± 1 h temporal window, the 2° spatial collocation radius, and the site-day missing-rate criterion (MR > 0.5). Stations with relatively sufficient high-missing-rate collocations were then selected as representative examples for visualization, whereas the overall validation statistics were calculated using all qualified high-missing-rate site-day collocations rather than only these representative stations.

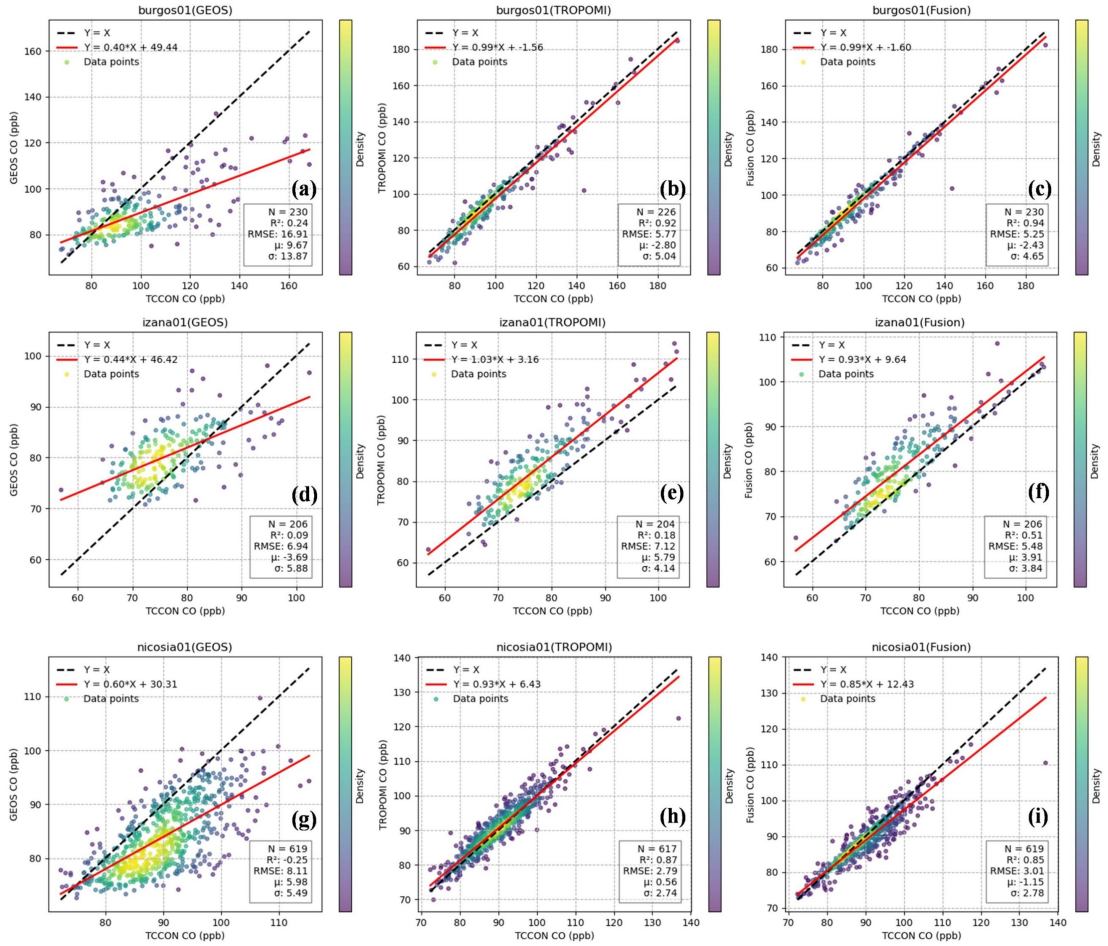


Figure 4. Density scatterplots of the independent validation results for GEOS-Chem, TROPOMI, and fused XCO data at burgos01, izana01, and nicosia01 under sparse TROPOMI coverage conditions (MR > 0.5). Panels (a, d, g), (b, e, h), and (c, f, i) correspond to GEOS-Chem, TROPOMI, and fused XCO data, respectively. Black dotted and red solid lines indicate the 1:1 line and fitted regression line, respectively. X: TCCON data; Y: GEOS-Chem, TROPOMI, and fused data. Units: ppb for RMSE, μ, and σ.

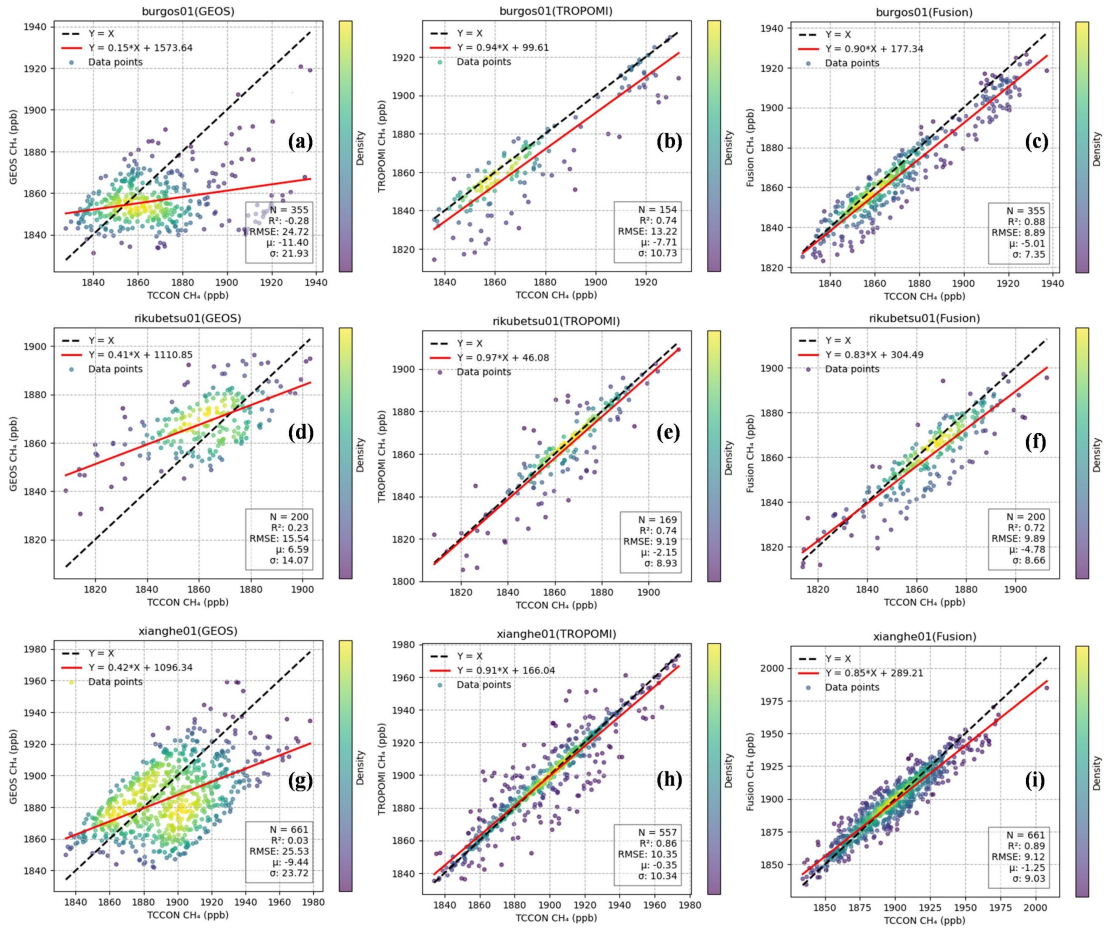


Figure 5. Density scatterplots of the independent validation results for GEOS-Chem, TROPOMI, and fused XCH_4 data at burgos01, rikubetsu01, and xianghe01 under sparse TROPOMI coverage conditions (MR > 0.5). Panels (a, d, g), (b, e, h), and (c, f, i) correspond to GEOS-Chem, TROPOMI, and fused XCH_4 data, respectively. Black dotted and red solid lines indicate the 1:1 line and fitted regression line, respectively. X: TCCON data; Y: GEOS-Chem, TROPOMI, and fused data. Units: ppb for RMSE, μ, and σ.

To demonstrate the performance of our fused datasets under different TROPOMI coverage, we conducted experiments using different MR thresholds. Tables S6 and S7 further quantify how the performance of the fused datasets varies with TROPOMI coverage. For XCO, when the missing rate is low or moderate (MR < 0.3 and 0.3 ≤ MR ≤ 0.5), TROPOMI shows slightly higher R^2 values and lower RMSE values than the fused dataset, indicating the high reliability of TROPOMI retrievals under relatively good coverage conditions. However, under sparse TROPOMI coverage conditions (MR > 0.5), the performance of TROPOMI decreases, with its R^2 decreasing to 0.89 and RMSE increasing to 6.16 ppb. In contrast, the fused XCO data maintain a higher R^2 value of 0.91 and a lower RMSE value of 5.65 ppb, indicating greater robustness under high-missing-rate conditions. For XCH_4, a similar pattern is observed. TROPOMI performs better than the fused product under low and moderate missing-rate conditions, with higher R^2 and lower RMSE. However, when MR > 0.5, TROPOMI performance decreases markedly, with R^2 declining to 0.81 and RMSE increasing to 15.84 ppb. The fused XCH_4 data show a slightly higher R^2 of 0.83 and a lower RMSE of 15.40 ppb under the same high-missing-rate condition. Although the fused XCH_4

bias is larger than that of TROPOMI, the higher R^2 and lower RMSE suggest that the fused dataset reduces overall errors under sparse-observation conditions. Overall, under conditions of good TROPOMI coverage, the fused XCO and XCH_4 datasets perform comparably to the TROPOMI datasets. When TROPOMI retrievals are sparse, the fused datasets show relatively high reliability. These results indicate that our fusion framework is effective for gap-filling applications.

Table S6. Validation statistics of XCO under different TROPOMI missing-rate intervals. GEOS-Chem results were bias-corrected before validation, and RMSE and μ are reported in parts per billion (ppb).

MR	R^2			RMSE			μ		
	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION
MR < 0.3	0.65	0.97	0.94	16.60	2.04	5.69	-0.22	-0.43	-1.10
0.3 ≤ MR ≤ 0.5	0.63	0.96	0.93	13.44	2.96	5.16	0.02	-0.88	-1.21
MR > 0.5	0.48	0.89	0.91	11.99	6.16	5.65	0.79	-0.87	-1.40

Table S7. Validation statistics of XCH_4 under different TROPOMI missing-rate intervals. GEOS-Chem results were bias-corrected before validation, and RMSE and μ are reported in parts per billion (ppb).

MR	R^2			RMSE			μ		
	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION
MR < 0.3	-0.15	0.95	0.87	29.10	3.05	9.66	-2.87	-0.57	-4.37
0.3 ≤ MR ≤ 0.5	-0.17	0.94	0.85	28.39	3.84	10.78	-3.19	-0.65	-2.17
MR > 0.5	0.34	0.81	0.83	29.84	15.84	15.40	-0.80	-5.48	-7.30

Comment 5: *Section 3.2 provides qualitative descriptions of spatial patterns but lacks quantitative analysis of underlying drivers. For example: can reconstructed signals be linked to emission inventories, meteorology, or policy changes? The discussion should move beyond descriptive mapping and include quantitative diagnostics or attribution analysis.*

Response 5:

We thank you for your constructive comment. We agree that the discussion should not be limited to qualitative analysis, but should also include quantitative analysis. In the revised manuscript, we added a quantitative analysis of short-term temporal variations in XCO and XCH_4 over the major rice-growing regions of Northeast and Southeast China. This analysis aims to evaluate whether the high-resolution fused datasets can capture regional temporal signals in agricultural emission-related regions, rather than merely reproduce spatial distribution patterns.

Specifically, we used harmonic regression with annual and semi-annual terms to reduce the influence of the seasonal cycle and derived short-term rates of change for both TROPOMI and fused datasets. The corresponding 95% confidence intervals are provided in Tables S9 and S10. The results show that XCO exhibits negative short-term rates of change in both datasets across the two rice-growing regions, with TROPOMI and fused rates of −2.36 and −2.28 ppb/year in Northeast China and −1.29 and −1.76 ppb/year in Southeast China, respectively. In contrast, XCH_4 exhibits positive short-term rates of change, with TROPOMI and fused rates of 12.08 and

12.16 ppb/year in Northeast China and 12.16 and 11.78 ppb/year in Southeast China, respectively. These consistent signs and comparable magnitudes indicate that the fused datasets preserve the main short-term XCO and XCH_4 signals observed by TROPOMI while providing spatially and temporally continuous fields.The specific figure and table revisions are as follows.

Revised tables and figures:

Table S9. Seasonally adjusted short-term rates of change and 95% confidence intervals for XCO over rice-growing regions during 2019–2023. Unit: ppb/year.

Region	Rate of change		95% confidence interval	
	TROPOMI	Fusion	TROPOMI	Fusion
Northeast China	−2.36	−2.28	−3.30 to −1.43	−3.30 to −1.27
Southeast China	−1.29	−1.76	−2.15 to −0.43	−2.79 to −0.72

Table S10. Seasonally adjusted short-term rates of change and 95% confidence intervals for XCH_4 over rice-growing regions during 2019–2023. Unit: ppb/year.

Region	Rate of change		95% confidence interval	
	TROPOMI	Fusion	TROPOMI	Fusion
Northeast China	12.08	12.16	11.24 to 12.92	11.27 to 13.04
Southeast China	12.16	11.78	10.83 to 13.49	9.72 to 13.83

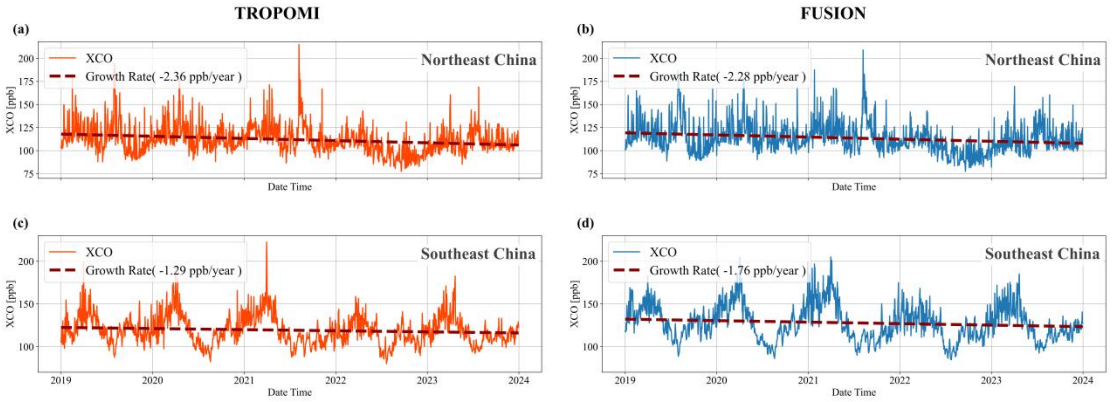


Figure 8. Daily time series of TROPOMI and fused XCO data over rice-growing regions in (a, b) Northeast China and (c, d) Southeast China. Red dashed lines indicate short-term rates of change derived from harmonic regression after accounting for the seasonal cycle, with values shown in parentheses. Unit: parts per billion per year.

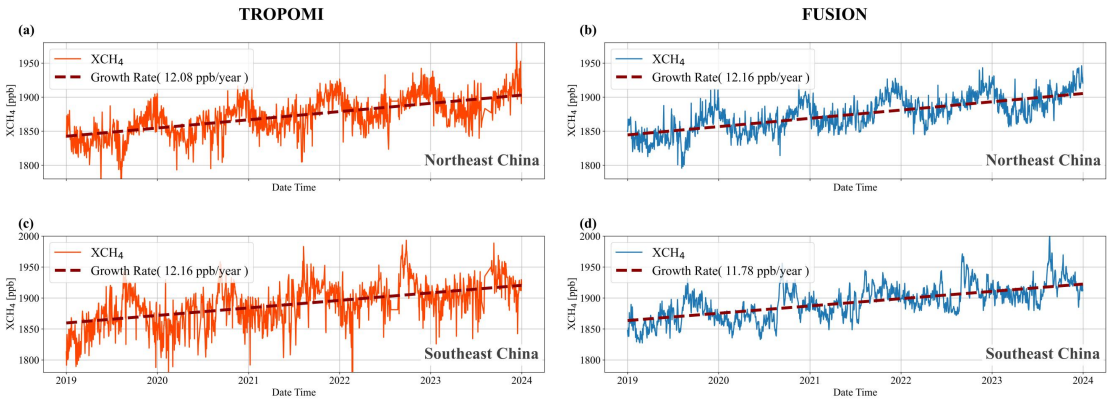


Figure 9. Daily time series of TROPOMI and fused XCH_4 data over rice-growing regions in (a, b) Northeast China and (c, d) Southeast China. Red dashed lines indicate short-term rates of change derived from harmonic regression after accounting for the seasonal cycle, with values shown in parentheses. Unit: parts per billion per year.

Comment 6: *There is a notable spatial-resolution mismatch between: GEOS-Chem simulation (2.5° globally; 0.25° for China) and final reconstructed maps (0.25° globally; 0.05° for China). The manuscript does not evaluate how this resolution mismatch affects reconstruction fidelity. Downscaling beyond model native resolution may introduce artificial fine-scale variability. I strongly suggest that the authors should perform at least one year of controlled experiments aligning model and reconstruction resolutions.*

Response 6:

Thank you for raising this important concern. We agree that the spatial-resolution mismatch between the native GEOS-Chem simulations and the final fused products may affect the interpretation of reconstruction fidelity, especially for fine-scale spatial variability. To address this issue, we added a one-year resolution-sensitivity experiment using the 2019 CO data.

In this experiment, we used GEOS-Chem simulation data at the native 2° × 2.5° resolution and aggregated TROPOMI XCO to the 2° × 2.5° resolution using the area-weighted averaging method. While keeping the fusion workflow and method unchanged, we reproduced our experiment at a resolution of 2° × 2.5°, thereby obtaining the fused XCO data at the 2° × 2.5° resolution. The obtained fused XCO data at the 2° × 2.5° resolution were then compared with the fused product obtained by aggregating the fused XCO data at the 0.25° resolution to the same 2° × 2.5° grid using the area-weighted averaging method.

Figure S12 shows that the fused XCO data generated directly at the native 2° × 2.5° resolution and the 0.25° fused XCO data aggregated to the same 2° × 2.5° grid present relatively similar global spatial distributions, with relatively high XCO over East Asia, central Africa, and parts of South America. The quantitative comparison in Figure S13 also shows high consistency, with $R^2 = 0.926$, RMSE = 6.93 ppb, and mean bias = 0.15 ppb. These results indicate that the 0.25° fused XCO data preserve the coarse-scale spatial structure and that the resolution mismatch does not have a substantial influence on the coarse-scale reconstruction fidelity of XCO. The corresponding description has been added to Section 2.2.1 and the "Discussion of uncertainties and limitations" section.

Revised text:

To evaluate the influence of the resolution mismatch between GEOS-Chem simulation results and the final fused product, we used the 2019 CO data to conduct a one-year resolution sensitivity experiment. In this experiment, we used GEOS-Chem simulation data at the native 2° × 2.5° resolution and aggregated TROPOMI XCO to the 2° × 2.5° resolution using the area-weighted averaging method. The related fusion calculation framework was directly implemented on the 2° × 2.5° grid; subsequently, the obtained fused XCO data at the 2° × 2.5° resolution were compared

with the fused product obtained by aggregating the fused XCO data at the 0.25° resolution to the same 2° × 2.5° grid using the area-weighted averaging method. (Lines 262–269)

Figure S12 shows that the fused XCO data generated directly at the native 2° × 2.5° resolution (Fig. S12a) and the 0.25° fused XCO data aggregated to the same 2° × 2.5° grid (Fig. S12b) present relatively similar global spatial distributions, with relatively high XCO over East Asia, central Africa, and parts of South America. The quantitative comparison results in Figure S13 also show high consistency, with R^2 = 0.926, RMSE = 6.93 ppb, and mean bias = 0.15 ppb. These results indicate that the 0.25° fused XCO data preserve the coarse-scale spatial structure, and that the resolution mismatch does not have a substantial influence on the coarse-scale reconstruction fidelity of XCO. (Lines 820–827)

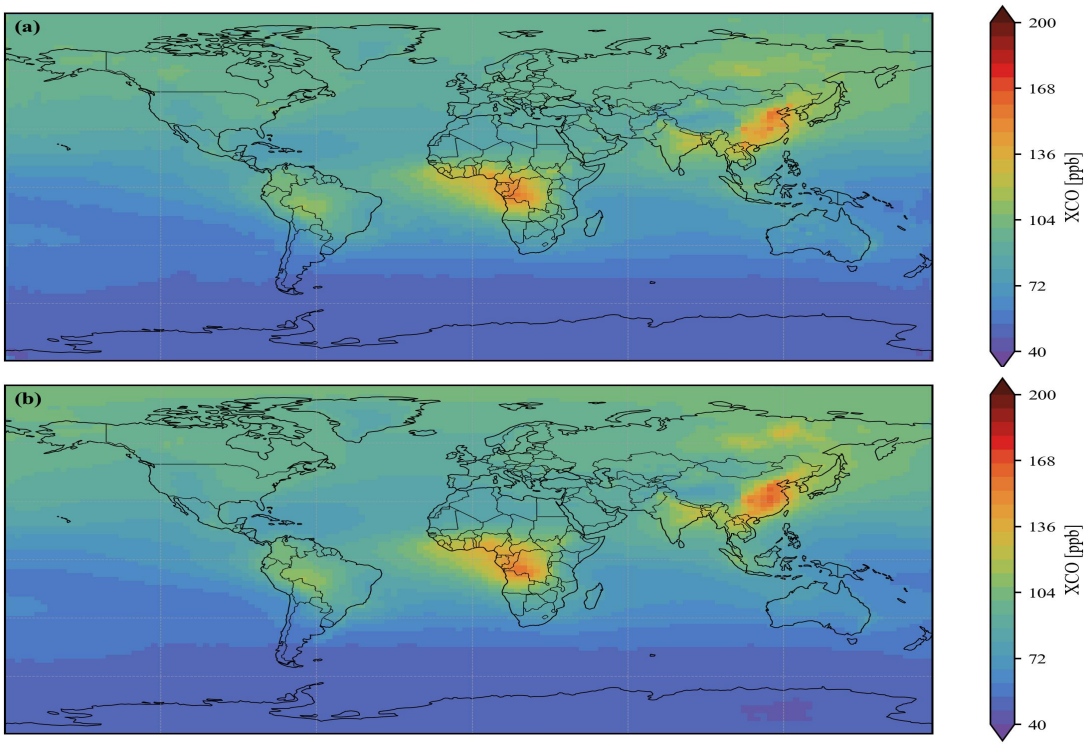


Figure S12. Spatial comparison of the 2019 XCO resolution-sensitivity experiment. (a) Fused XCO data generated directly at the native 2° × 2.5° GEOS-Chem resolution. (b) Fused XCO data obtained by aggregating the 0.25° resolution fused XCO data to the same 2° × 2.5° grid using the area-weighted averaging method.

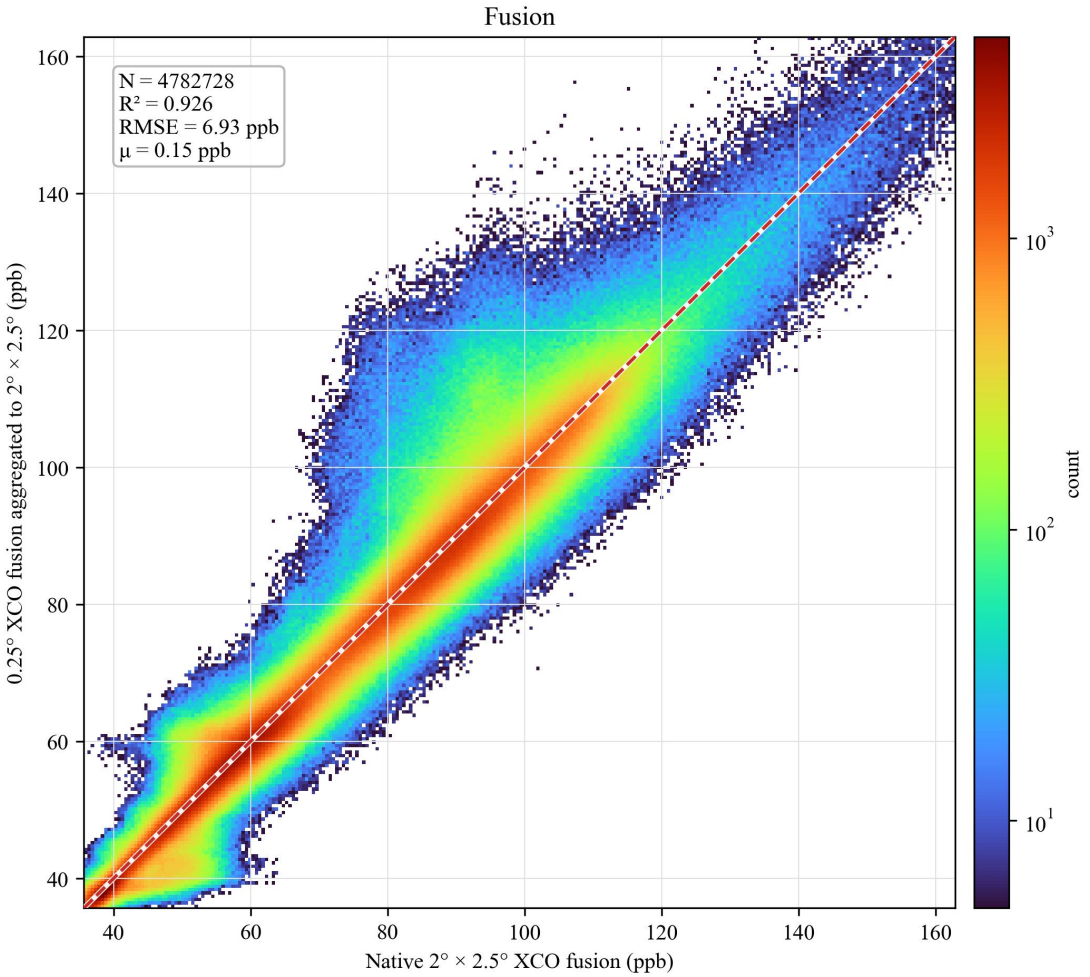


Figure S13. Quantitative comparison of the 2019 XCO resolution-sensitivity experiment.

Comment 7: *The manuscript requires significant language polishing. There are grammatical errors, duplicated statements, imprecise terminology, and ambiguous phrasing. The overall readability would benefit from careful editing by a fluent academic English speaker.*

Response 7:

We sincerely apologize for the linguistic deficiencies in the initial submission. We agree that the readability and precision of the manuscript needed improvement. We have carefully revised the full manuscript to improve the language quality, clarity, and consistency. The main revisions include the following:

(1) Grammatical and stylistic corrections: We corrected grammatical errors, improved sentence structure, and revised awkward or unclear expressions throughout the manuscript to make the text more fluent and suitable for academic writing.

(2) Removal of duplicated statements: We removed repeated descriptions, especially in the data and method sections, where some information about GEOS-Chem, TROPOMI, and validation procedures had been partially repeated.

(3) Standardization of terminology: We standardized key terms throughout the manuscript. For example, we now consistently use “fused dataset(s)” or “fused data” instead of mixed expressions such as “integrated data,” “amalgamated data,” and “fusion-generated data.”
(4) Clarification of ambiguous methodological descriptions: We revised several unclear descriptions, including the definitions of missing rate, product versions, DCT/SVD reconstruction, masked loss, bias correction, trend calculation, and site-selection criteria, to make the workflow more reproducible.
(5) More cautious interpretation of results: We revised statements such as “outperforms TROPOMI” to more careful wording, such as “shows improved agreement with TCCON” (Lines 622–623) or “show that the fused products perform comparably to original TROPOMI retrievals under low and moderate missing-rate conditions ,” (Lines 29–30) to avoid overstatement.
(6) Revision of figures, captions, and supplementary descriptions: We checked figure captions, supplementary tables, and related descriptions for consistency, missing punctuation, and clearer interpretation.
The revised manuscript has been carefully checked again to improve overall readability and academic expression.

Specific comments

Comment 8: *Line 16: Specify whether “continuity” refers to temporal continuity, spatial coverage, or both.*

Response 8:

We thank you for pointing out this ambiguity. In the revised manuscript, we have modified this statement and clarified it as “spatiotemporal continuity.” This revision is intended to clarify that the limitations of existing methods concern both temporal continuity, namely the ability to generate sequences without missing daily data, and spatial coverage, namely completeness over the target grid. The relevant text has been added to Lines 18–20 of the revised manuscript.

Revised text:

However, satellite observations from instruments such as the TROPOspheric Monitoring Instrument (TROPOMI) are often spatially and temporally incomplete, while existing fusion methods still struggle to achieve both high accuracy and spatiotemporal continuity.

Comment 9: *Line 22: The phrase “superior to TROPOMI” requires clarification. Superior in what metric? What datasets are used to compute R2? Please define explicitly.*

Response 9:

We sincerely thank you for this valuable comment. We agree that the expression “superior to TROPOMI” was not entirely accurate. What we originally intended to express was that, compared with the original TROPOMI retrievals, the fused results show a higher correlation with ground-based TCCON measurements. In the revised manuscript, we have reorganized this part; specifically, we refined the wording and clarified the comparison criterion. The relevant text has

been added to Lines 29–32 of the revised manuscript.

Revised text:

the fused products perform comparably to original TROPOMI retrievals under low and moderate missing-rate conditions and show improved performance under sparse TROPOMI coverage (MR > 0.5), with R^2 values of 0.91 for XCO and 0.83 for XCH_4, along with reduced biases and standard deviations.

Comment 10: *Line 27–28: Overestimation/underestimation relative to which reference dataset?*

Response 10:

The overestimation and underestimation mentioned in the original manuscript refer to those of our fused dataset relative to the TROPOMI dataset.

Comment 11: *Line 47–49: Add supporting references.*

Response 11:

We thank you for your suggestion. We have added the reference Lelieveld et al. (2016) to support this statement. The relevant text has been added to Lines 54–56 of the revised manuscript.

Revised text:

CH_4 plays a crucial role in atmospheric chemistry, influencing the oxidative capacity of the atmosphere and contributing to tropospheric ozone formation. Its rising concentration is a major driver of climate change (Lelieveld et al., 2016).

Comment 12: *Line 57–58: Clarify whether the listed satellites observe both CO and CH4 simultaneously or individually.*

Response 12:

We thank you for pointing out this ambiguity. To make this clarification more explicit, the observation capabilities of the listed satellite instruments are summarized below. The relevant text has been added to lines 73–81 of the revised manuscript.

Instrument	CO observe	CH_4 observe
MOPITT	Yes	No
AIRS	Yes	Yes
TES	Yes	Yes
IASI	Yes	Yes

Revised text:

Global satellite monitoring enables effective tracking and analysis of the sources, transport, and removal processes of atmospheric pollutants. At present, several satellite instruments provide observations of CO and/or CH_4, including the Measurements of Pollution in the Troposphere (MOPITT), which primarily focuses on CO retrievals (Deeter et al., 2003), as well as the Atmospheric Infrared Sounder (AIRS) (McMillan et al., 2005), Tropospheric Emission Spectrometer (TES) (Rinsland et al., 2006), and Interferometer for the Exploration of the Atmosphere in the Infrared (IASI) (Turquety et al., 2004), which are capable of retrieving multiple trace gases, including CO and CH_4, using different spectral channels and retrieval algorithms. These instruments offer critical data support for the global distribution of CO and CH_4 in the atmosphere.

Comment 13: *Line 65: TROPOMI provides tropospheric column retrievals, not total column in all cases. Please clarify.*

Response 13:

We thank you for your attention to the description of the TROPOMI data products. We checked the official Sentinel-5P/TROPOMI Level-2 product table (https://sentiwiki.copernicus.eu/web/s5p-products) and confirmed that the CO and CH_4 products used in this study are both listed as total column products. In the revised manuscript, we further clarified that XCO and XCH_4 refer to the column-averaged dry-air mole fractions of atmospheric CO and CH_4, respectively, derived from the corresponding TROPOMI total column products. This clarification has been added to Lines 84–87 of the revised manuscript.

Revised text:

According to the official Sentinel-5P/TROPOMI Level-2 product table, the CO and CH_4 products are provided as total column products. In this study, XCO and XCH_4 refer to the column-averaged dry-air mole fractions of atmospheric CO and CH_4, respectively, derived from the corresponding TROPOMI total column products.

Comment 14: *Line 71–72: TROPOMI does not measure surface concentrations. Revise accordingly.*

Response 14:

We thank you for pointing out this inaccuracy. We agree that TROPOMI provides column-averaged observations rather than direct surface measurements. Accordingly, we have revised the text to remove the reference to "surface" observations.

Comment 15: *Line 73–74: Provide specific regional examples with references.*

Response 15:

Thank you for this helpful suggestion. We have revised Lines 73–74 to provide specific regional

examples and supporting references. The relevant text has been added to Lines 90–94 of the revised manuscript.

Revised text:

However, the continuity of TROPOMI atmospheric column retrievals is affected by clouds and aerosols, which can attenuate or scatter the reflected solar radiation measured by satellite sensors. As a result, observations may be incomplete, particularly in regions with high cloud cover and aerosol load—such as the humid tropics and high-latitude zones (Gao et al., 2023), Southeast Asia (Valerio et al., 2025), and eastern China during intense pollution episodes (Wang et al., 2021).

Comment 16: *Line 75–92: The categorization of gap-filling methods into three groups is unclear and potentially incomplete. The logical connectors "on the one hand" and "on the other hand" are misused. Moreover, hybrid methods combining machine learning and model constraints already exist. Please clarify how your classification is defined and whether it is exhaustive.*

Response 16:

Thank you for this helpful comment. We agree that the original wording in Lines 75–92 was not sufficiently clear. In particular, the use of "on the one hand" and "on the other hand" could incorrectly imply a binary classification, whereas the existing gap-filling methods actually involve multiple methodological pathways. In the revised manuscript, we no longer present the classification as an exhaustive or mutually exclusive taxonomy. Instead, we clarify that the three groups are a practical, literature-based summary of the major approaches commonly used for improving the continuity of TROPOMI XCO and XCH_4 products. Specifically, we reorganized the discussion into: (1) machine learning-based interpolation methods, which reconstruct missing observations by learning nonlinear spatiotemporal relationships; (2) enhanced spectral fitting algorithms, which improve retrieval continuity by refining the spectral inversion process; and (3) physics–data–driven coupling approaches, including model fusion, data assimilation, and hybrid machine learning frameworks, which integrate satellite retrievals with chemical transport models or reanalysis datasets.

We also explicitly acknowledge that hybrid methods combining machine learning and physical/model constraints already exist, and we have incorporated these studies into the third category rather than treating model fusion and machine learning as completely separate or opposing approaches. This revision better reflects the current state of the field and clarifies that our proposed framework builds upon, but differs from, existing hybrid approaches by jointly combining GEOS-Chem model priors, signal-domain low-rank reconstruction using 3D DCT–SVD, and residual U-Net refinement under masked observational constraints.The relevant text has been added to Lines 95–117 of the revised manuscript.

Revised text:

The generation of continuous TROPOMI XCO and XCH_4 products has been addressed through a range of gap-filling approaches. First, machine learning-based interpolation methods (X. Chen et al., 2022; Hu et al., 2022; Valerio et al., 2025; Wei et al., 2025) reconstruct missing observations by learning nonlinear spatiotemporal relationships. However, their performance strongly depends

on the quality and representativeness of the training data, and biased or insufficient samples may lead to unreliable predictions. Second, enhanced spectral fitting algorithms (Borsdorff et al., 2019; Guanter et al., 2015; Schneising et al., 2023; Wang et al., 2020) improve retrieval continuity by refining the spectral inversion process, for example through higher-order polynomial fitting to reduce reflectance-induced biases, albeit at increased computational cost. Third, physics–data–driven coupling approaches, including model fusion, data assimilation, and hybrid machine learning frameworks, integrate satellite retrievals with chemical transport models or reanalysis datasets (Fritz et al., 2022; Schneising et al., 2023; Sicard et al., 2021). By combining physical constraints with data-driven learning or observational corrections, these methods improve spatial consistency and temporal continuity while reducing retrieval bias. However, their performance remains dependent on uncertainties in prior model simulations and the representativeness of assimilated observations (Inness et al., 2022; Wang et al., 2023).

Despite these advances, current methods face challenges in balancing computational efficiency with physical consistency, particularly in regions with persistent cloud cover. To overcome these limitations, this study proposes a novel signal-domain guided spatio-temporal fusion framework. Our core objective is to generate daily global and regional continuous XCO and XCH_4 products (2019–2023) at high resolution (0.25° globally and 0.05° over China) by effectively leveraging complementary information from chemical transport modeling and frequency-domain representations.

Comment 17: *Line 105: Clarify what "reconstructed fields" refers to.*

Response 17:

We thank you for pointing out that the term "reconstructed fields" was not sufficiently clear. In the revised manuscript, we have clarified that "reconstructed fields" refers to the signal-domain fields generated by the DCT/SVD-based low-rank reconstruction step, which are used to approximate missing TROPOMI values and provide a prior for the subsequent fusion procedure. The relevant revision has been added to Lines 128-130 of the revised manuscript.

Revised text:

This masked learning strategy ensures that the model is supervised exclusively by valid observations, thereby preserving the physical consistency of the DCT/SVD-derived signal-domain reconstructed fields;

Comment 18: *Line 106: Citation error ("Tang, n.d.").*

Response 18:

We thank the reviewer for identifying the citation error ("Tang, n.d.") in the original manuscript.The specific revision is provided below.

Revised text:

(Ronneberger et al., 2015; Tang, 2022)

Comment 19: *Line 107–109: The sentence is unclear. If TROPOMI data serve as spatial constraints rather than direct inputs, this needs clearer explanation.*

Response 19:

We thank you for pointing out this ambiguity. We have revised the text to clarify the "Masked Learning" strategy used in our Residual U-Net framework. Specifically, TROPOMI observations are not directly fed into the network as input features, since the satellite fields contain substantial spatial gaps caused by cloud coverage and retrieval limitations. Instead, the network takes spatially continuous fields (GEOS-Chem simulations and preliminary DCT-based fusion results) as inputs. The valid TROPOMI observations are incorporated as a weighting mask in the loss function during training, such that only pixels with valid satellite observations contribute to the optimization process. Invalid or missing regions are excluded from gradient computation. The relevant revision has been added to Lines 126-130 of the revised manuscript.

Revised text:

Crucially, instead of using the TROPOMI data mask as a direct input feature, we incorporate it as a weighting mask in the loss function (Wei et al., 2022). This masked learning strategy ensures that the model is supervised exclusively by valid observations, thereby preserving the physical consistency of the DCT/SVD-derived signal-domain reconstructed fields;

Comment 20: *Line 121–122: Provide the full name of TROPOMI at first appearance.*

Response 20:

We thank you for pointing out this oversight. In the revised manuscript, following standard academic writing conventions, we have provided the full name of "TROPOMI" ("TROPOspheric Monitoring Instrument") at its first occurrence in both the Abstract and the Introduction. Specifically, the full name has been added in Lines 18–19 and Line 82 of the revised manuscript.

Revised text:

The TROPOspheric Monitoring Instrument (TROPOMI)

Comment 21: *Line 121–123: Remove duplicated meaning in consecutive sentences.*

Response 21:

We thank you for pointing out the duplicated meaning in the original manuscript. In the revised manuscript, we have streamlined the text by removing repetitive statements regarding the Sentinel-5 Precursor (S5P) satellite and the TROPOMI instrument. The revised text can be found in Lines 143–145 of the revised manuscript. The specific revisions are as follows.

Revised text:

The TROPOMI instrument onboard the Sentinel-5 Precursor (S5P) satellite operates in a polar

sun-synchronous orbit, and provides daily global XCO and XCH_4 observations at approximately 13:30 local solar time, thereby enabling near-global daily coverage.

Comment 22: *Line 123: Clarify whether the discussion applies only to XCO or also XCH4.*

Response 22:

We thank you for pointing out this ambiguity. We have reviewed the text and ensured that all technical descriptions explicitly specify whether they refer to XCO, XCH_4, or both. In cases where the characteristics (such as the 13:30 overpass time and 7 × 5.5 km^2 spatial resolution) are identical for both products, we now use the combined term "XCO and XCH_4". Where differences exist (such as spectral retrieval bands), we have separated the descriptions to avoid confusion.The relevant text has been added to Lines 143–145 of the revised manuscript.

Revised text:

The TROPOMI instrument onboard the Sentinel-5 Precursor (S5P) satellite operates in a polar sun-synchronous orbit, and provides daily global XCO and XCH_4 observations at approximately 13:30 local solar time, thereby enabling near-global daily coverage.

Comment 23: *Line 127–132: Discuss cloud and aerosol sensitivity for XCH4 retrievals as well.*

Response 23:

We thank you for this important suggestion to discuss cloud and aerosol sensitivity for XCH_4 retrievals as well. In the revised manuscript, we have expanded the description of the XCH_4 retrieval algorithm and added a discussion of its sensitivity to cloud and aerosol effects. The relevant text has been added to Lines 152–155 of the revised manuscript.

Revised text:

For XCH_4, the operational retrieval algorithm simultaneously estimates column-averaged dry-air mole fractions of CH_4 together with atmospheric scattering parameters. High retrieval accuracy and precision are generally achieved under clear-sky conditions over land after strict filtering of scenes affected by clouds, aerosols, or high surface albedo variability (Hu et al., 2016).

Comment 24: *Line 133: Specify exact data version.*

Response 24:

We thank you for the suggestion to specify the exact data version. In the revised manuscript, we have explicitly stated the precise product names, product versions and the corresponding Digital Object Identifiers (DOIs) for the TROPOMI Level-2 datasets used in our study. The relevant text has been added to Lines 156–160 of the revised manuscript.

Revised text:

The TROPOMI Level-2 CO and CH_4 products used in this study were obtained from the NASA GES DISC Sentinel-5P/TROPOMI archive for the period 2019–2023. We used the OFFL (Offline) datasets. Specifically, the CO product used is S5P_L2__CO_____HiR, Version 02 (DOI: 10.5270/S5P-bj3nry0), while the CH_4 product used is S5P_L2__CH4____HiR, Version 02 (DOI: 10.5270/S5P-3lcdqiv).

Comment 25: *Line 134: "Quality Value" or "Quality Flag"? Use consistent terminology.*

Response 25:

We thank you for pointing out the inconsistency in our terminology. Following the official TROPOMI product documentation and to ensure technical precision, we have standardized the terminology throughout the manuscript to "quality value (qa_value)", which represents the quality assurance value. We have replaced all previous instances of "Quality Value" and "Quality Flag" with this standard term. The relevant revisions have been made in the revised manuscript.

Revised text:

(1) we applied product-specific quality screening and retained only TROPOMI retrievals with a quality value (qa_value) greater than 0.5 for subsequent analysis. (Lines 164–166)

(2) TROPOMI XCO and XCH_4 retrievals with qa_value below 0.5 were excluded. (Lines 255–256)

(3) where N_{total} is the total number of 0.25° grid cells within the 2° radius and N_{valid} is the number of grid cells with valid TROPOMI retrievals after applying the qa_value > 0.5 criterion. Grid cells without satellite retrievals or with qa_value $\leqslant$ 0.5 were treated as missing. We selected site-day observation cases with MR > 0.5, corresponding to less than 50% valid TROPOMI coverage within the collocation window, to evaluate the performance of the fused XCO and XCH_4 datasets under sparse TROPOMI coverage conditions. (Lines 508–513)

(4) For each gas, we first screened TCCON stations that retained a relatively large number of valid spatially collocated data pairs after applying the 13:30 ± 1 h temporal window, the 2° spatial matching radius, and the TROPOMI quality filter (qa_value > 0.5). (Lines 543–546)

Comment 26: *Figure 1 / Table 1: Since TCCON data are not produced in this study, consider moving these to Supplementary Information. Main figures should prioritize new results.*

Response 26:

We agree with your suggestion. To prioritize the presentation of our original fusion results and the methodological framework, we have moved Figure 1 (Global Distribution of TCCON Sites) and Table 1 (Site details) to the Supplementary Information as Figure S1 and Table S1. This allows the main manuscript to focus more directly on the proposed signal-domain fusion architecture and the resulting gap-free datasets. The detailed revisions are as follows.

Revised content:

The TCCON sites used in this study, shown in Figure S1 and listed in Table S1, were selected based on data availability during 2019–2023. (Lines 219–221)

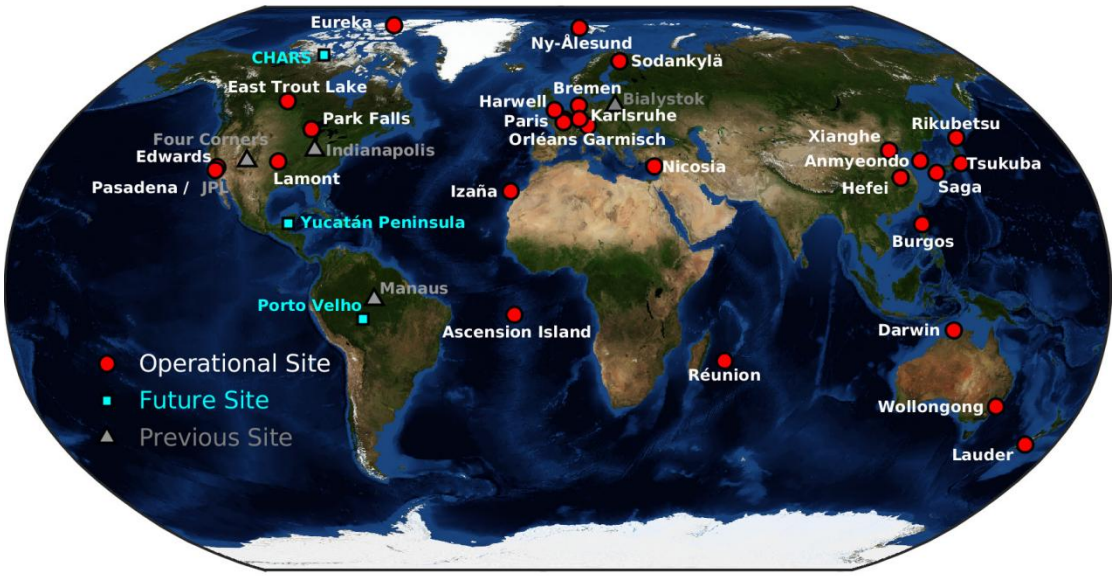


Figure S1. Global distribution of TCCON sites, including operational sites, future sites, and previous sites, respectively.

No.	Site name	Latitude	Longitude	Location	Start date	End date
1	bremen01	53.104	8.85	Bremen, Germany	2009/1/6	2021/6/24
2	burgos01	18.5325	120.6496	Burgos, Philippines	2017/3/3	2023/6/23
3	darwin01	-12.425	130.891	Darwin, Australia	2013/1/1	2022/12/27
4	easttroutlake01	54.354	-104.987	East Trout Lake, Canada	2016/10/3	2024/2/15
5	edwards01	34.9599	-117.881	AFRC, Edwards, CA, USA	2013/7/20	2024/2/22
6	garmisch01	47.476	11.063	Garmisch, Germany	2007/7/18	2023/5/4
7	hefei01	31.91	117.17	Hefei, China	2015/11/2	2023/12/25
8	izana01	28.31	-16.5	Izana, Tenerife, Spain	2014/1/2	2023/8/30
9	karlsruhe01	49.103	8.44	Karlsruhe, Germany	2014/1/15	2023/6/26
10	lamont01	36.604	-97.486	Lamont, Oklahoma, USA	2011/4/16	2024/2/25
11	lauder03	-45.038	169.684	Lauder, New Zealand	2018/10/2	2023/12/28
12	nicosia01	35.141	33.381	Nicosia, Cyprus	2019/9/1	2023/5/10
13	orleans01	47.97	2.113	Orleans, France	2009/9/6	2023/6/23
14	parkfalls01	45.945	-90.273	Park Falls, Wisconsin, USA	2004/6/2	2024/2/25
15	pasadena01	34.136	-118.127	Pasadena, California, USA	2012/9/20	2024/2/25
16	reunion01	-20.901	55.485	Reunion Island, France	2015/3/1	2020/7/18
17	rikubetsu01	43.4567	143.7661	Rikubetsu, Hokkaido, Japan	2014/6/24	2021/6/30
18	sodankyla01	67.367	26.631	Sodankylä, Finland	2009/5/16	2023/5/30
19	tsukuba02	36.0513	140.1215	Tsukuba, Ibaraki, Japan, 125HR	2014/3/28	2021/3/31
20	wollongong01	-34.406	150.891	Wollongong, Australia	2013/1/4	2023/6/27
21	xianghe01	39.8	116.96	Xianghe, China	2018/6/14	2023/5/29

Table S1. Details of the TCCON sites used in this study, No. is the serial number.

Comment 27: *Figure 2: Unify font type across figures (e.g., Arial).*

Response 27:

We thank you for your suggestion to unify the font type in Figure 2. We have redesigned Figure 2 (now Figure 1 due to figure reordering) to ensure that a consistent font type (Arial) is used throughout the figure. In addition, following the suggestions of other reviewers, we increased the font size in the technical schematic of the U-Net architecture to improve readability.

Revised figure:

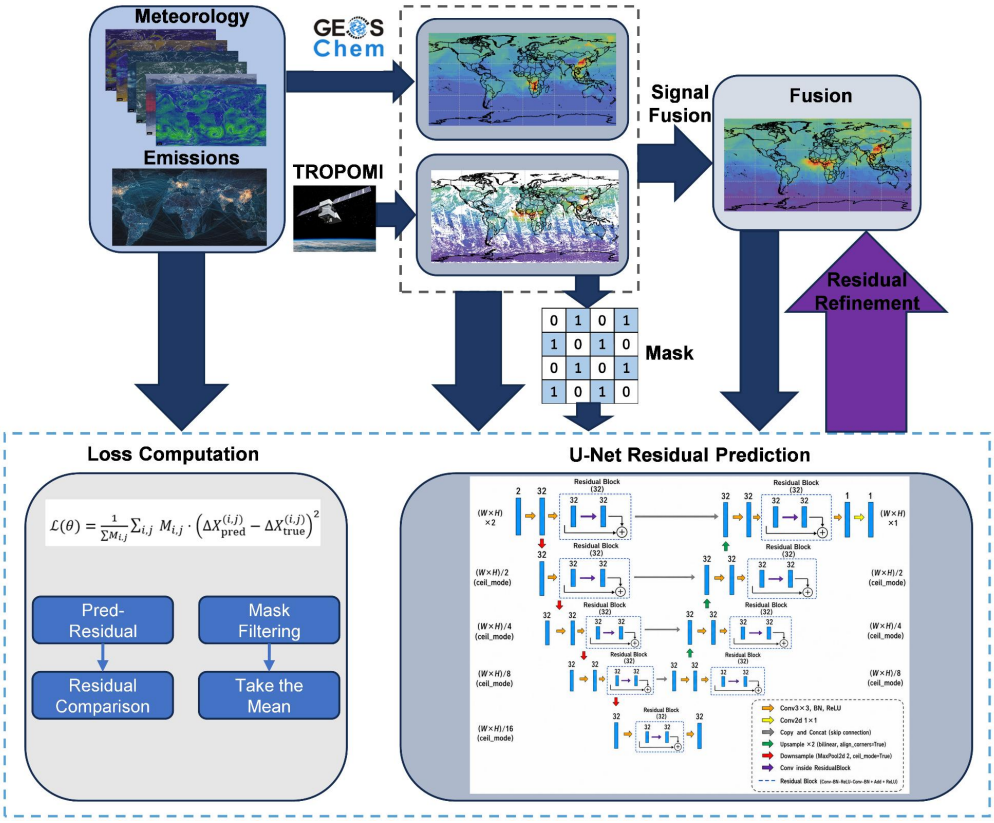


Figure 1. Overview of the proposed fusion framework combining physical modeling, signal-domain reconstruction, and deep learning-based residual correction. The process includes data preprocessing, DCT&SVD-based spatio-temporal fusion, and residual refinement via a lightweight residual U-Net architecture. Arrows indicate the information flow across different modules.

Comment 28: *Line 217–220 & 224–227: Remove duplicated expressions.*

Response 28:

We thank you for pointing out the duplicated descriptions in Section 2.2.1. In the revised manuscript, we streamlined the data preprocessing description by removing repetitive statements regarding quality control, regridding, and GEOS-Chem interpolation. The revised text now presents the preprocessing workflow in a more concise and logically organized manner. The revised text can be found in Lines 254–261 of the revised manuscript.

Revised text:

Data preprocessing is essential for ensuring the reliability of the fusion results. To reconcile the spatial resolution differences between satellite retrievals and model simulations, TROPOMI XCO and XCH_4 retrievals with qa_value below 0.5 were excluded, and the remaining observations were aggregated onto a global 0.25° × 0.25° grid (720 × 1440) using area-weighted averaging (Wang et al., 2018). GEOS-Chem outputs were subsequently interpolated to the same grid using inverse distance weighting (Setianto and Triandini, 2013). The 0.25° spatial resolution was selected to balance the representation of mesoscale spatial variability with computational efficiency for multi-year global fusion analyses (Hu et al., 2024; Wang et al., 2023).

Comment 29: *Line 232–238: Rewrite for clarity.*

Response 29:

We thank you for your suggestion to improve the clarity of this paragraph. We have rephrased the original text to more clearly explain that, although magnitude biases exist in the GEOS-Chem simulations, the GEOS-Chem results remain consistent with TROPOMI observations in terms of spatiotemporal variability. The relevant text has been added to Lines 282–290 of the revised manuscript.

Revised text:

Previous studies have shown that emission inventories typically underestimate GEOS-Chem simulation results (Hu et al., 2018; Liang et al., 2023). Despite these magnitude biases, GEOS-Chem simulations generally remain consistent with TROPOMI observations in terms of spatio-temporal variability, exhibiting similar temporal evolution patterns, including concurrent increases and decreases. This consistency suggests that the spatio-temporal correlation between GEOS-Chem simulations and TROPOMI observations can be effectively exploited. By leveraging this relationship, a data fusion framework based on signal-domain reconstruction can be used to generate datasets with improved spatial continuity and more complete coverage (J. Chen et al., 2022; He et al., 2022; Wang et al., 2021).

Comment 30: *Line 362: Typo in "Scheme".*

Response 30:

We apologize for this spelling error. The heading "Evaluation Schem" has been corrected to "Evaluation Scheme" in Section 2.2.6.

Revised text:

2.2.6. Evaluation scheme

Comment 31: *Line 378: TCCON sites are not uniformly distributed globally. Revise.*

Response 31:

We thank you for pointing out this inaccurate description. In the revised manuscript, we have corrected the statement that TCCON sites are uniformly distributed worldwide. The relevant text has been added to Lines 490–492 of the revised manuscript.

Revised text:

The TCCON sites are sparsely and unevenly distributed across the globe, with higher concentrations in North America, Europe, East Asia, and a limited number of remote oceanic and high-altitude regions.

Comment 32: *Line 383: Define "missing rate".*

Response 32:

We thank you for pointing out that the definition of "missing rate" was unclear in the original manuscript. In the revised manuscript, we have added an explicit definition of the TROPOMI missing rate used in the TCCON validation procedure. The relevant text has been added to lines 504–513 of the revised manuscript.

Revised text:

For each site-day collocation, the missing rate was defined as the fraction of grid cells within the 2° radius that had no valid TROPOMI retrieval after quality filtering:

$$MR = 1 - \frac{N_{valid}}{N_{total}} \tag{14}$$

where N_{total} is the total number of 0.25° grid cells within the 2° radius and N_{valid} is the number of grid cells with valid TROPOMI retrievals after applying the qa_value > 0.5 criterion. Grid cells without satellite retrievals or with qa_value $\leqslant$ 0.5 were treated as missing. We selected site-day observation cases with MR > 0.5, corresponding to less than 50% valid TROPOMI coverage within the collocation window, to evaluate the performance of the fused XCO and XCH_4 datasets under sparse TROPOMI coverage conditions.

Comment 33: *Line 386–388: Justify why these sites are selected.*

Response 33:

We thank you for this valuable comment. We agree that the rationale for selecting these sites for the time-series comparison was not sufficiently explained in the original manuscript. In the revised manuscript, we have clarified that these sites were selected to illustrate the temporal variation characteristics of the fused data, TROPOMI, GEOS-Chem, and TCCON observations. Because the purpose of these figures is to show time-series consistency, we selected stations with relatively continuous collocated records and a relatively large number of valid spatially collocated data pairs. The relevant text has been added to lines 543–548 of the revised manuscript.

Revised text:

For each gas, we first screened TCCON stations that retained a relatively large number of valid spatially collocated data pairs after applying the 13:30 ± 1 h temporal window, the 2° spatial matching radius, and the TROPOMI quality filter (qa_value > 0.5). Among these qualified stations, we further selected those with relatively continuous multi-year records and limited long data gaps to illustrate the temporal consistency among GEOS-Chem, TROPOMI, the fused data, and TCCON observations.

Comment 34: *Line 521: The focus on rice cultivation regions should be introduced earlier before interpreting Figures 9–10.*

Response 34:

We thank you for this valuable suggestion. In the revised manuscript, before interpreting the original Figures 9–10 (now Figures 8–9 due to figure reordering), we introduced the relevance of rice cultivation regions to CH_4 and CO variability. The relevant text has been added to Lines 745–749 of the revised manuscript.

Revised text:

Rice paddies are an important source of CH_4 emissions, while associated agricultural activities, such as crop-residue burning, may also affect regional CO variability. Therefore, we further analyzed the short-term temporal changes in XCO and XCH_4 over the major rice-growing regions of Northeast and Southeast China to evaluate the applicability of the high-resolution fused datasets in agricultural emission-related regions.

Comment 35: *Line 564: Attribution of model underestimation solely to inventory accuracy is oversimplified. Chemical mechanisms, transport, and parameterizations should also be considered.*

Response 35:

Thank you for this important comment. We agree that attributing the GEOS-Chem underestimation solely to emission inventory uncertainties is oversimplified. We have revised the conclusion to clarify that the systematic underestimation may arise from multiple sources, including uncertainties in emission inventories, chemical mechanisms, atmospheric transport, boundary conditions, vertical mixing, and model parameterizations. The relevant text has been added to Lines 923–928 of the revised manuscript.

Revised text:

The fusion process also mitigates the systematic underestimation observed in GEOS-Chem simulations. Such model biases may arise from multiple sources, including uncertainties in emission inventories, chemical mechanisms, atmospheric transport, boundary conditions, vertical mixing, and other model parameterizations. Therefore, the fused products can serve as an improved basis for analyzing long-term spatial patterns, seasonal variability, and regional changes in atmospheric CO and CH_4.

Response to Reviewer 2:

General comments

Comment 1: *This manuscript presents a two-stage framework that combines signal-domain reconstruction (DCT + SVD) with a residual U-Net to generate gap-free global datasets of XCO and XCH_4 from TROPOMI and GEOS-Chem simulations. The topic is relevant for ESSD, and the proposed framework is interesting because it integrates signal processing and deep learning for spatiotemporal data fusion. However, the dataset, particularly the CH_4 record, covers only five years (2019–2023), which limits its value for atmospheric and climate research. Moreover, several methodological details and validation aspects require further analysis before the dataset and method can be fully assessed.*

Response 1:

We sincerely thank you for your positive evaluation of our study and for recognizing the potential value of the proposed two-stage framework that combines signal-domain reconstruction with residual learning. We also appreciate your concern regarding the relatively short temporal coverage of this study and the need for clearer methodological and validation details.

We agree that the 2019–2023 record, especially for XCH_4, is not sufficient for robust long-term climate-trend analysis. Therefore, in the revised manuscript, we have avoided overinterpreting the dataset as a long-term climate record. We revised the relevant wording from "long-term trends" or "annual trends" to more cautious terms such as "five-year rate-of-change analysis," "short-term linear changes," or "fitted annual rates of change during 2019–2023."

In response to your comments on methodological and validation details, we have substantially revised the manuscript to improve transparency and reproducibility. In particular, we clarified the satellite product versions and data sources, the missing-rate definition and site-selection criteria, the DCT/SVD and residual-learning procedures, the masked-loss strategy, the treatment of averaging kernels, the implications of the GEOS-Chem/fused-dataset resolution mismatch, and the calculation of regional rates and event-related enhancements. We also revised the language and terminology throughout the manuscript to improve readability and avoid overstatement.

Although the current dataset is limited to 2019–2023, we believe the revised manuscript now presents it more appropriately as a daily gap-free XCO and XCH_4 dataset for analyzing short-term spatiotemporal variability, assessing regional events, and demonstrating a reproducible fusion framework that can be extended to longer records in future work.

Specific comments

Comment 2: *The manuscript adopts a signal-domain reconstruction approach using 3-D DCT followed by SVD truncation. However, the rationale for choosing this specific combination is not sufficiently explained. In particular, the manuscript states that 80% of the singular value energy is retained, but no sensitivity analysis is provided.*

Response 2:

Thank you for your valuable comment. We agree that the rationale for using the 3-D DCT followed by SVD truncation, as well as the choice of the 80% cumulative singular-value energy threshold, should be more clearly justified. We have revised the Methods section accordingly and added a sensitivity analysis.

1. Why did we choose the DCT + SVD method?

The 3-D DCT is employed to transform the spatiotemporal data into the frequency domain, effectively isolating the low-frequency "background" signals (large-scale atmospheric patterns) from high-frequency noise. However, DCT alone can sometimes retain redundant cross-dimensional correlations. By applying SVD subsequently, we can further compress these coefficients and extract the most dominant spatiotemporal structures (the primary singular vectors). This combination helps the preliminary fusion stage capture robust regional and seasonal XCO and XCH_4 patterns while reducing random retrieval noise. We have added the relevant description in Lines 306–314 of the revised manuscript.

The parameter matrix ρ is assumed to be smooth in space and time, inspired by previous studies on filling missing values in spatio-temporal data and enhancing data smoothness using multidimensional discrete cosine transforms (Elharar et al., 2007; Garcia, 2010, 2010; J. Robinson and V. Kecman, 2003; Okolie and Smit, 2022; Peng et al., 2005; Rao and Yip, 2014; Wang et al., 2023). We therefore propose a spatio-temporal 3D matrix smoothing algorithm based on the discrete DCT and SVD to enhance data smoothness and fill missing values in spatio-temporal data. The method effectively handles spatio-temporal data with missing values while maintaining spatio-temporal correlation by combining spatio-temporal nearest-neighbor interpolation and regularized optimization techniques.

2. Why was 80% of the singular value energy retained?

To evaluate whether this empirical threshold materially affects the reconstruction, we conducted an additional threshold-sensitivity experiment for the 2019 daily global XCO reconstruction(Due to the limited revision time and available computational resources, we only performed the analysis for CO). Daily GEOS-Chem XCO simulations and collocated TROPOMI CO retrievals on the 0.25° × 0.25° grid were used to construct the TROPOMI/GEOS-Chem ratio fields. The same DCT-SVD reconstruction procedure was repeated using cumulative singular-value energy thresholds of 70%, 75%, 80%, 85%, 90%, and 95%, with the 80% case used as the reference. The results are summarized in Table S8 and Figure S2. (Lines 370–376)

We added Sect. 3.1.2 to describe the results of the sensitivity test.

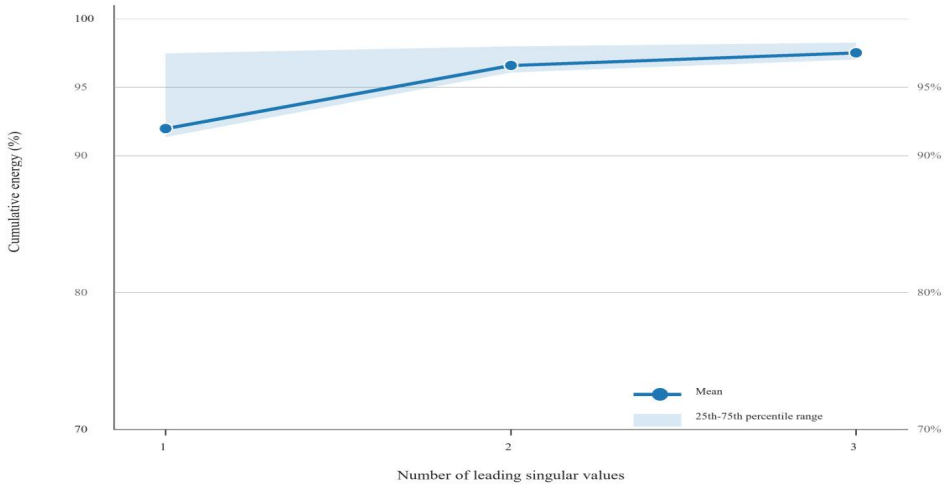
The sensitivity experiment showed that the reconstructed XCO fields were highly stable across the tested SVD energy thresholds. Relative to the 80% reference case, all threshold experiments yielded spatial correlation coefficients higher than 0.9998, RMSE differences below 0.62 ppb, spatial standard deviation ratios close to unity, and hotspot overlap ratios greater than 0.97 (Table S8). The cumulative singular-value energy analysis further showed that the first, first two, and first three singular values explained approximately 92.0%, 96.6%, and 97.5% of the total energy, respectively (Fig. S2). These results indicate that the XCO reconstruction is not sensitive to the exact SVD truncation threshold within the tested range of 70%-95%, supporting the use of 80% as a stable and computationally efficient threshold.

Table S8. Sensitivity of the 2019 daily XCO reconstruction to the SVD cumulative-energy threshold.

SVD energy threshold (%)	Spatial Rvs. 80%	RMSE vs. 80% (ppb)	μ vs. 80% (ppb)	SD ratio vs. 80%	Hotspot overlap
70	0.99998	0.206	-0.088	0.99989	0.983
75	0.99999	0.160	-0.086	0.99990	0.984
80	1.00000	0.000	0.000	1.00000	1.000
85	0.99996	0.271	0.000	1.00000	1.000
90	0.99992	0.382	0.003	0.99965	0.999
95	0.99984	0.616	0.285	1.00019	0.975

Note: The 80% threshold was used as the reference case. Spatial R denotes the Pearson correlation coefficient calculated across valid grid cells between each threshold experiment and the 80% case. RMSE, μ, and SD ratio were also calculated relative to the 80% case, where SD denotes the spatial standard deviation. Hotspot overlap was calculated using the upper 10% of reconstructed XCO grid cells. All statistics were derived from the 2019 global daily 0.25° × 0.25° CO reconstruction experiments.

Figure S2. Cumulative singular-value energy explained by the leading singular values in the 2019 daily global XCO reconstruction.



Comment 3: *The proposed method contains two key stages:(1) signal-domain reconstruction using DCT and SVD, and (2) residual correction using a deep neural network. However, the current experiments do not clearly isolate the contribution of each stage. The authors compare several machine learning models, but an explicit ablation study would help clarify the role of each component.*

Response 3:

Thank you for your valuable comment. We agree that the contribution of each component in the

proposed two-stage framework should be more clearly distinguished. In the revised manuscript, we added two subsections, “Sect. 2.2.5 Ablation study design” and “Sect. 3.1.1 TCCON-based validation setup and stage-wise ablation,” to clarify the tested configurations and evaluation strategy.

Specifically, we conducted site-based validation across all TCCON stations in 2021 using the same TCCON collocation strategy as that used for the final product validation. Five configurations were compared: (1) DCT-only fusion, (2) DCT/SVD mixed-signal reconstruction, (3) DCT/SVD + residual CNN, (4) DCT/SVD + residual U-Net, and (5) DCT/SVD + residual XGBoost. The comparison between “Only DCT” and “DCT/SVD mixed-signal reconstruction” isolates the contribution of the SVD-based low-rank constraint, while the comparison between “DCT/SVD mixed-signal reconstruction” and the residual-learning configurations isolates the contribution of the residual-correction stage. The corresponding site-based validation results are summarized in Tables S2 and S3.

The ablation results show that both stages make measurable contributions to the final fused datasets. For XCO, incorporating SVD into the DCT-based reconstruction increased R^2 from 0.8227 to 0.8335 and reduced RMSE from 10.4390 to 10.1166 ppb; the residual U-Net further improved R^2 to 0.8450 and reduced RMSE to 9.7616 ppb. For XCH_4, incorporating SVD increased R^2 from 0.7056 to 0.7320 and reduced RMSE from 17.5156 to 16.7111 ppb; the residual U-Net further improved R^2 to 0.7598 and reduced RMSE to 15.8212 ppb. These results demonstrate that the DCT/SVD signal-domain reconstruction stage and the residual U-Net correction stage both improve the validation performance of the fused XCO and XCH_4 datasets.

2.2.5. Ablation study design

To explicitly isolate the contribution of each component in the proposed two-stage framework, we conducted an ablation study using site-based validation across all TCCON stations for the year 2021. Five configurations were compared: (1) only DCT fusion(Garcia, 2010), (2) DCT and SVD mixed-signal reconstruction(M. Bengherabi et al., 2008; Majhi and Pal, 2021), (3) residual CNN(Y. Jiang et al., 2024), (4) residual U-Net (Ronneberger et al., 2015; Yan et al., 2022), and (5) residual XGBoost(Naseem et al., 2024). The performance was assessed using the coefficient of determination (R^2), root mean square error (RMSE), and mean bias (μ). This design allows the contribution of the DCT/SVD signal-domain reconstruction stage and the residual-correction stage to be evaluated separately. The site-based validation results are summarized in Tables S2 and S3.

3.1.1. TCCON-based validation setup and stage-wise ablation

Based on the same TCCON collocation strategy, we first performed an ablation analysis to clarify the contribution of each component in the proposed two-stage framework before evaluating the final fused datasets. As summarized in Tables S2 and S3, the comparison between “Only DCT” and “DCT/SVD mixed-signal reconstruction” isolates the contribution of the SVD-based low-rank constraint, while the comparison between “DCT/SVD mixed-signal reconstruction” and the residual-learning configurations isolates the contribution of the residual-correction stage. For XCO, adding SVD to DCT increased R^2 from 0.8227 to 0.8335 and reduced RMSE from 10.439 to 10.1166 ppb. The residual U-Net further improved R^2 to 0.8450 and reduced RMSE to 9.7616 ppb. For XCH_4, adding SVD increased R^2 from 0.7056 to 0.7320 and reduced RMSE from

17.5156 to 16.7111 ppb, while the residual U-Net further improved R^2 to 0.7598 and reduced RMSE to 15.8212 ppb. These results indicate that both the DCT/SVD signal-domain reconstruction stage and the residual U-Net correction stage make measurable contributions to the final fused datasets.

After confirming the contribution of each component, we further evaluated the final fused datasets against TCCON observations at representative stations and across all qualified station-day collocations. (Lines 527–529)

Table S2. Ablation study of site-based XCO validation using different reconstruction configurations. Metrics include coefficient of determination (R^2), root mean square error (RMSE, ppb), and mean bias (μ, ppb).

Configuration	R^2	RMSE	μ
DCT-only	0.8227	10.439	-3.815
DCT/SVD	0.8335	10.1166	-3.7682
DCT/SVD + residual CNN	0.8385	9.9619	-3.7677
DCT/SVD + residual U-Net	**0.845**	**9.7616**	**-3.686**
DCT/SVD + residual XGBoost	0.8392	9.9406	-3.7435

Table S3. Ablation study of site-based XCH_4 validation using different reconstruction configurations. Metrics include coefficient of determination (R^2), root mean square error (RMSE, ppb), and mean bias (μ, ppb).

Column	R^2	RMSE	μ
DCT-only	0.7056	17.5156	-6.9345
DCT/SVD	0.732	16.7111	-6.9635
DCT/SVD + residual CNN	0.7469	16.2404	-6.9162
DCT/SVD + residual U-Net	**0.7598**	**15.8212**	**-6.7858**
DCT/SVD + residual XGBoost	0.7476	16.2188	-6.9563

Comment 4: *The model is trained using a masked loss that only considers locations with valid satellite observations. However, the final dataset is expected to fill large areas where satellite observations are missing due to clouds or aerosols. The method is the same as in previous studies, for example: https://doi.org/10.1016/j.atmosres.2025.108411 or https://doi.org/10.1038/s41467-023-43862-3.*

Response 4:

Thank you for your valuable comment, and for providing literature support for our masked loss function method. We have cited this literature in the revised manuscript. At the same time, we also believe that the use of the masked loss function needs further clarification, especially because the final fused dataset is expected to cover large areas where satellite observations are missing due to the effects of clouds or aerosols. In our method, the model is trained using only valid observations, i.e., locations where M=1; the masked loss function is adopted because reliable TROPOMI observational targets exist only at valid retrieval pixels; therefore, invalid or missing retrieval

results should not affect the calculation of the training loss. Although the loss is computed only at valid pixels, the trained residual model is applied to the full spatial domain. This is possible because the model inputs, including GEOS-Chem simulations and the preliminary DCT/SVD reconstruction, are spatially complete or reconstructed before residual learning. Thus, the model transfers residual relationships learned from valid observations to missing regions with similar spatiotemporal patterns represented by the two input fields. We have added the relevant discussion in Lines 413–423 of the revised manuscript.

Revised text:

The model is trained using only the valid observations, i.e., locations where M=1. The masked loss was used because reliable TROPOMI targets are available only at valid retrieval pixels; pixels affected by clouds, aerosols, or retrieval failures do not provide trustworthy reference values. The mask therefore avoids penalizing the network at invalid or missing satellite pixels. Although the loss is computed only at valid pixels, the trained residual model is applied to the full spatial domain because its inputs, including GEOS-Chem simulations and the preliminary DCT/SVD reconstruction, are spatially complete or reconstructed before residual learning. Thus, the model transfers residual relationships learned from valid observations to missing regions with similar spatiotemporal patterns represented by the two input fields, following the logic of recent masked-learning atmospheric gap-filling studies (Wei et al., 2023; Sun et al., 2026). The loss function is designed to minimize the residual error at these observed locations:

Comment 5: *The validation against TCCON uses an averaging radius of ±2°. Since TCCON provides point-scale measurements while the fused dataset has a grid resolution (0.25°), the spatial representativeness mismatch should be discussed in more detail.*

Response 5:

Thank you for this important comment. We agree that spatial representativeness mismatch exists when comparing point-scale TCCON measurements with gridded satellite/model-based products. We have added the relevant discussion in Lines 496–502 of the revised manuscript.

Revised text:

We acknowledge that TCCON provides near-point-scale column measurements, whereas the TROPOMI and fused datasets represent spatially averaged grid-cell values at 0.25° resolution. Therefore, a spatial representativeness mismatch is unavoidable, especially in regions with strong horizontal gradients, complex terrain, coastal influences, or localized emission sources. The 2° radius was used as a compromise to include a sufficient number of satellite and gridded data samples around each station while limiting the influence of distant air masses.

Comment 6: *The manuscript mainly compares the fused results with TROPOMI and GEOS-Chem simulations. However, other atmospheric composition datasets or assimilation products exist (e.g., CAMS atmospheric composition products or other fused greenhouse-gas datasets). Including comparisons with at least one existing product would help better position the proposed dataset*

within the current landscape of atmospheric composition data products.

Response 6:

Thank you for this valuable suggestion. We agree that comparing the dataset developed in this study with at least one existing product would help better position it within the current landscape of atmospheric composition data products. In the revised manuscript, we added a literature-based comparison with representative existing products, including CAMS-EGG4 and the seamless daily global XCO_2/XCH_4 fusion product proposed by Wang et al. (2023).

Because these products differ in target species, input observations, assimilation or fusion methods, spatial/temporal resolutions, and validation protocols, a direct pixel-level comparison would be difficult to interpret. Therefore, we compared our TCCON-based validation statistics with the reported TCCON-based accuracies of these existing products. This contextual comparison shows that our fused XCO and XCH_4 datasets achieve competitive validation performance while providing daily gap-free TROPOMI–GEOS-Chem fused products at 0.25° global resolution and 0.05° resolution over China. We have clarified that this comparison is intended for contextual positioning rather than direct product ranking.We added this comparison in Lines 663–682 of the revised manuscript. The specific comparison is as follows:

Product	Target variables	Main input data	Reported validation performance
CAMS-EGG4	XCO_2, XCH_4	Model simulations and assimilated satellite observations	Monthly systematic and random errors generally within about 1% for CO_2 and CH_4 during 2003–2020
Wang et al. (2023)	XCO_2, XCH_4	GOSAT, OCO-2, and CAMS-EGG4	TCCON validation: $\sigma \approx 1.18$ ppm for XCO_2 and $\sigma \approx 11.3$ ppb for XCH_4; R^2 = 0.91–0.95 for XCO_2 and $R^2 = 0.90$ for XCH_4
This study	XCO, XCH_4	TROPOMI, GEOS-Chem,	TCCON validation: $R^2 = 0.91$ for XCO and $R^2 = 0.83$ for XCH_4

Revised text:

To further position the proposed fused datasets within the current landscape of atmospheric composition products, we compared our TCCON-based validation results with reported accuracies of representative existing products in the literature, rather than conducting a direct pixel-level comparison. Direct product-to-product comparison is not straightforward because existing products differ in target species, input observations, assimilation or fusion methods, spatial and temporal resolutions, and validation protocols. For example, CAMS-EGG4 provides global greenhouse-gas reanalysis fields for CO_2 and CH_4 by combining model and satellite information within the ECMWF Integrated Forecasting System, but it does not provide a TROPOMI-based XCO product directly comparable to the fused XCO dataset developed here (Agustí-Panareda et al., 2023). Agustí-Panareda et al. (2023) reported that CAMS-EGG4 generally shows monthly systematic and random errors for CO_2 and CH_4 within about 1% during 2003–2020, while also noting caveats related to changes in assimilated observations and fixed underlying emissions. Wang et al. (2023) developed a daily seamless 0.25° global XCO_2 and XCH_4 fusion product from GOSAT, OCO-2, and CAMS-EGG4, and reported TCCON validation statistics of $\sigma \approx 1.18$ ppm

for XCO_2 and σ ≈ 11.3 ppb for XCH_4, with R^2 values of 0.91–0.95 for XCO_2 and 0.90 for XCH_4. In comparison, our fused datasets achieve R^2 values of 0.91 for XCO and 0.83 for XCH_4 against TCCON. Although these values are not strictly comparable because of differences in target gases, input data, and validation settings, they indicate that the proposed TROPOMI–GEOS-Chem fusion framework provides competitive TCCON-based validation performance while extending daily gap-free mapping to both XCO and XCH_4 at global 0.25° resolution and China-specific 0.05° resolution.

Comment 7: *In several places the manuscript suggests that the fused dataset "outperforms TROPOMI" This statement should be phrased more carefully. Since satellite observations are direct measurements, the improvement likely reflects reduced random noise or improved spatiotemporal consistency rather than strictly higher accuracy than the observations themselves.*

Response 7:

Thank you for raising this important issue. We should not simply describe the fused dataset as "superior to" TROPOMI, because TROPOMI provides the direct satellite observations used to constrain the fusion process. We have revised the manuscript to avoid overstatement. Specifically, we replaced expressions such as "superior to TROPOMI," "exceeds TROPOMI," and "outperforms TROPOMI" with more cautious wording, such as "shows improved agreement with TCCON compared with the original TROPOMI retrievals," "reduces random scatter," and "enhances spatiotemporal consistency." We also added a clarification that the fused dataset is not intended to replace the original TROPOMI observations; rather, it provides a gap-free and spatiotemporally consistent reconstruction constrained by TROPOMI, GEOS-Chem, and meteorological information. Therefore, the improved validation statistics relative to TROPOMI should be interpreted as improved consistency with TCCON and reduced random variability under the adopted collocation protocol, rather than as evidence that the fused product is intrinsically more accurate than direct satellite measurements.

Revised text:

shows improved agreement with TCCON compared with the original TROPOMI retrievals for several validation metrics. (Lines 622–624)

Comment 8: *Finally, can robust CO and CH_4 trends be derived from the data without applying gap-filling?*

Response 8:

We thank you for this fundamental question regarding the necessity of gap-filling. In principle, trends or rates of change can be derived directly from raw TROPOMI observations. However, such estimates may be strongly affected by spatiotemporal sampling biases because satellite data gaps are not randomly distributed. Missing observations are often concentrated in specific seasons or regions with persistent cloud cover, aerosol contamination, high surface albedo variability, or low-quality retrievals. As a result, statistics derived from raw TROPOMI data may be biased

toward clear-sky conditions and may not consistently represent regional mean conditions.

In the revised manuscript, we therefore avoid overinterpreting the 2019–2023 record as a robust long-term climate trend. Instead, we describe the analysis as a five-year rate-of-change or short-term linear change assessment. The purpose of gap-filling in this context is to provide spatially and temporally consistent daily fields, which allows area-weighted regional averages to be calculated more consistently than when using irregular raw satellite sampling alone.

Our comparison over rice-growing regions illustrates this issue: raw TROPOMI and the fused product give different fitted rates of change, suggesting that observational gaps can affect regional estimates. Therefore, while rates of change can be calculated from raw TROPOMI observations, the fused dataset provides a more complete and internally consistent basis for short-term spatiotemporal analysis. The following is our detailed comparative analysis for different rice-growing regions.

Therefore, we further analyzed the short-term temporal changes in XCO and XCH_4 over the major rice-growing regions of Northeast and Southeast China to evaluate the applicability of the high-resolution fused datasets in agricultural emission-related regions. Figures 8 and 9 show the temporal variations in XCO and XCH_4 derived from TROPOMI and the fused data over these regions, respectively. To reduce the influence of the seasonal cycle, the short-term rates of change were derived using harmonic regression with annual and semi-annual terms, and the corresponding 95% confidence intervals are provided in Tables S9 and S10. (In lines 746–753 of the revised manuscript)

The results show that XCO exhibits negative short-term rates of change in both the satellite and fused datasets across the rice cultivation areas of Northeast and Southeast China. In Northeast China, the XCO rate is −2.36 ppb/year for TROPOMI and −2.28 ppb/year for the fused data. In Southeast China, the XCO rate is −1.29 ppb/year for TROPOMI and −1.76 ppb/year for the fused data. The consistent negative rates indicate that the fused dataset preserves the main decreasing XCO signal observed by TROPOMI. Differences between the two datasets may be partly related to the uneven temporal sampling of TROPOMI caused by clouds, aerosols, and retrieval filtering, whereas the fused dataset provides spatially and temporally continuous fields. (In lines 764–771 of the revised manuscript)

In contrast, XCH_4 exhibits positive short-term rates of change in both the satellite and fused datasets. In Northeast China, the XCH_4 rate is 12.08 ppb/year for TROPOMI and 12.16 ppb/year for the fused data. In Southeast China, the XCH_4 rate is 12.16 ppb/year for TROPOMI and 11.78 ppb/year for the fused data. The close agreement between the two datasets indicates that the fused dataset preserves the main increasing XCH_4 signal observed by TROPOMI while providing improved temporal continuity. (In lines 772–777 of the revised manuscript)

Figure 8. Daily time series of TROPOMI and fused XCO data over rice-growing regions in (a, b) Northeast China and (c, d) Southeast China. Red dashed lines indicate short-term rates of change derived from harmonic regression after accounting for the seasonal cycle, with values shown in parentheses. Unit: parts per billion per year.

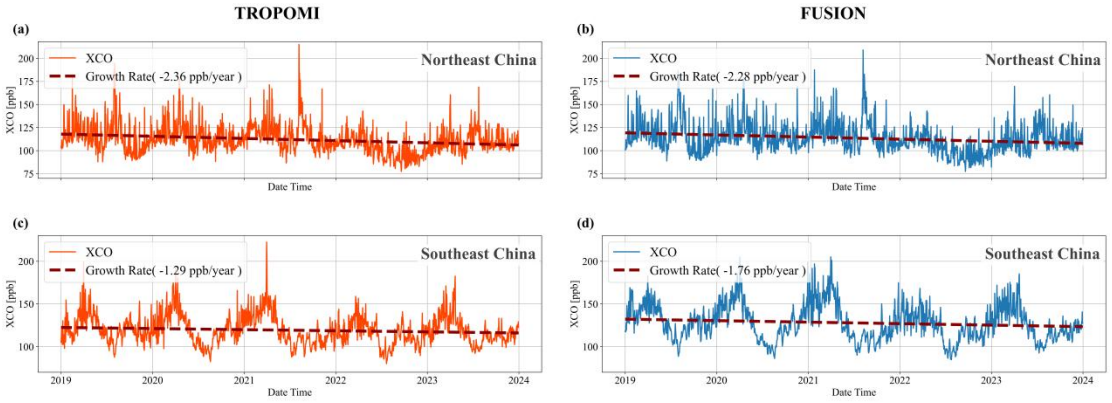


Figure 9. Daily time series of TROPOMI and fused XCH_4 data over rice-growing regions in (a, b) Northeast China and (c, d) Southeast China. Red dashed lines indicate short-term rates of change derived from harmonic regression after accounting for the seasonal cycle, with values shown in parentheses. Unit: parts per billion per year.

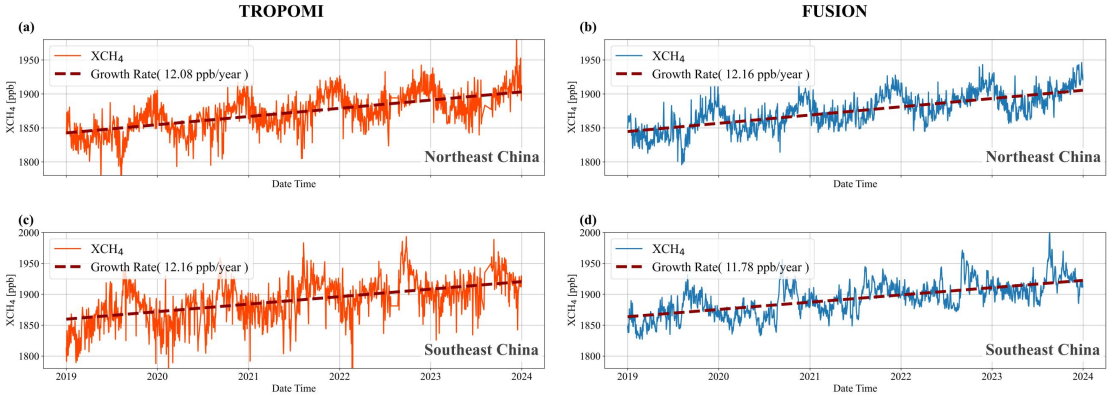


Table S9. Seasonally adjusted short-term rates of change and 95% confidence intervals for XCO over rice-growing regions during 2019–2023. Unit: ppb/year.

Region	Rate of change		95% confidence interval	
	TROPOMI	Fusion	TROPOMI	Fusion
Northeast China	−2.36	−2.28	−3.30 to −1.43	−3.30 to −1.27
Southeast China	−1.29	−1.76	−2.15 to −0.43	−2.79 to −0.72

Table S10. Seasonally adjusted short-term rates of change and 95% confidence intervals for XCH_4 over rice-growing regions during 2019–2023. Unit: ppb/year.

Region	Rate of change		95% confidence interval	
	TROPOMI	Fusion	TROPOMI	Fusion
Northeast China	12.08	12.16	11.24 to 12.92	11.27 to 13.04
Southeast China	12.16	11.78	10.83 to 13.49	9.72 to 13.83

Technical corrections:

Comment 9: *Figure 1: missing ".".*

Response 9:

We thank you for pointing out the missing period in Figure 1. Please note that following the suggestion of another reviewer, Figure 1 has been moved to the Supplementary Information as Figure S1. We have added the missing period at the end of the caption of Figure S1 in the revised Supplementary Material.

Revised text:

Figure S1. Global distribution of TCCON sites, including operational sites, future sites, and previous sites, respectively.

Comment 10: *Lines 87-88, Your model is also limited by the accuracy of prior model simulations and the representativeness of assimilated observations (GEOS- Chem).*

Response 10:

We agree with the reviewer. The performance of our fusion framework is inherently coupled with the quality of the GEOS-Chem prior and its underlying assimilated data. To address this dependency, we have explicitly discussed this limitation in the newly added Section 3.4.

Revised text:

Another concern relates to the dependence of the fused datasets on the accuracy of the GEOS-Chem prior and the representativeness of valid TROPOMI observations. GEOS-Chem provides physically consistent large-scale background information, but it carries uncertainties from emission inventories, transport, chemistry, boundary conditions, vertical mixing, and other model parameterizations. In regions or periods where valid TROPOMI retrievals are persistently sparse—such as the Intertropical Convergence Zone, high-latitude zones, or heavily polluted areas with dense cloud or aerosol cover—the reconstruction relies more strongly on the GEOS-Chem prior and the DCT/SVD spatiotemporal structure. Consequently, systematic biases in the prior model may propagate into the fused fields. Users are therefore advised to consult the obs_mask variable provided in the dataset to assess observational support and to interpret the reconstructed values cautiously in persistently missing regions.

Comment 11: *Line 122: avoid redundant definition (e.g., S5P).*

Response 11:

We thank you for pointing out the redundant definition of S5P. In the revised manuscript, we have removed the repeated definition and retained the abbreviation only at its first occurrence.The revised wording is provided in Lines 143–145 of the revised manuscript.

Revised text:

The TROPOMI instrument onboard the Sentinel-5 Precursor (S5P) satellite operates in a polar sun-synchronous orbit, and provides daily global XCO and XCH4 observations at approximately 13:30 local solar time, thereby enabling near-global daily coverage.

Response to Reviewer 3:

General (major) comments

Comment 1:

Dataset

I checked the linked repository and evaluated selected files. The NetCDF files lack basic CF-compliant metadata. Time is not defined in physical units, variable attributes are missing, and units are not provided. The dataset in its current form is not self-describing and, in my opinion, does not meet ESSD data standards.

1、 *Data content description states that there is Global 0.25x0.25° data as well as Regional (China) data at 0.05x0.05° resolution. In the repository, I could find only 0.25x0.25°. data.*

2、 *The files contain monthly aggregated data, yet there is no time variable, so analyzing a large set of files would be inconvenient.*

3、 *There is no CH_4 variable in the files.*

4、 *The variable named CO should be XCO.*

5、 *There are no error/uncertainty/ML_confidence /reconstruction_flag/ observation_fraction and any similar variables. As a user, I would expect at least some of this information in the file. Also, without this information, the dataset risks being interpreted as observation-based while large regions may be model-driven.*

Response 1:

Thank you for your careful inspection of our repository and NetCDF files. We agree that the previous NetCDF files did not provide sufficient CF-style metadata expected for ESSD data reuse. Therefore, we reorganized the dataset and revised the NetCDF file structure. These improvements have improved the readability, reusability, and long-term maintainability of the dataset. The specific changes are shown below.

1. Locations of the global and regional (China) products in the data repository

The revised dataset includes XCO and XCH_4 products at two spatial resolutions: global 0.25° × 0.25° and China regional 0.05° × 0.05°. XCO and XCH_4 are stored in separate NetCDF files. Due to repository limits on file size and file number, the files are organized by month and uploaded as annual compressed archives. The final 2019-2023 five-year dataset consists of 20 compressed packages, corresponding to 5 years × 2 gases × 2 spatial resolutions.

2. Addition of a time variable

For file-size and upload-management reasons, these files are organized by month, but the data are daily. We added a time coordinate to each monthly NetCDF file. The time coordinate uses CF-style physical time units, such as days since 2019-01-01 00:00:00. Here, time = 0 represents the first day of the month. Users can obtain daily fused XCO or XCH4 fields by selecting the corresponding time index within each monthly file.

3. Variable names and units

We corrected the variable names. The former variable CO was renamed XCO. The methane product is now provided as XCH_4. Both variables use ppb as the unit. Their attributes state that they represent column-averaged dry-air mole fractions.

4. Metadata and CF-style attributes

We added essential metadata to the NetCDF files. These include coordinate attributes for latitude, longitude, and time; variable long names; units; missing-value definitions; and global attributes. The global attributes describe the product, spatial resolution, temporal coverage, product region, data source, authors, institution, and contact information. The revised files also include CF-style metadata, such as Conventions = CF-1.8.

5. Observation and reconstruction indicators

We added an obs_mask variable. This variable helps users distinguish satellite-observed values from reconstructed values. The obs_mask variable is a unitless binary mask. When obs_mask = 1, the value in the corresponding grid cell is based on a TROPOMI observation. When obs_mask = 0, the value is generated by the proposed reconstruction method.

We also added an observation_fraction variable. This variable is calculated as the mean of obs_mask over the time dimension in each monthly file. It represents the fraction of days with valid TROPOMI observations at each grid cell during that month. These two variables help prevent users from interpreting the full gap-free product as purely observation-based. This is especially important in regions or periods where reconstruction contributes substantially.

We have added the relevant description in Lines 853–892 of the revised manuscript.

Revised text:

The final data products generated in this study consist of daily gap-free fused XCO and XCH_4 datasets for the period 2019–2023. Two spatial scales are provided. The global product contains daily XCO and XCH_4 fields at 0.25° × 0.25° resolution, covering the global domain. The regional product provides higher-resolution daily XCO and XCH_4 fields over China at 0.05° × 0.05° resolution. Both products are reported as column-averaged dry-air mole fractions in parts per billion (ppb).

The datasets are distributed in NetCDF4 (.nc/.nc4) format. Each monthly file contains gridded daily fields with the dimensions `time`, `lat`, and `lon`. The `time` coordinate is defined using CF-style physical time units, for example, `days since 2019-01-01 00:00:00`, with `time = 0` corresponding to the first day of the month. Users can therefore retrieve a specific daily field by selecting the corresponding `time` index in the monthly NetCDF file. The coordinate variables `lat` and `lon` define the spatial grid of each product.

The main data variables are `XCO` and `XCH4`, representing the fused column-averaged dry-air mole fractions of CO and CH_4, respectively. In addition, an obs_mask variable is provided to indicate whether the value at each grid cell is a TROPOMI observation or reconstructed by the fusion framework. The obs_mask variable is a dimensionless binary mask, where obs_mask = 1 indicates that the value at that grid cell and day is a TROPOMI observation, while obs_mask = 0 indicates that the value was generated by the proposed reconstruction method.

An `observation_fraction` variable is also included in each monthly file. This variable is calculated as the mean of `obs_mask` over the `time` dimension and represents the fraction of days in that month for which valid TROPOMI observations are available at each grid cell. For example, `observation_fraction = 1` indicates that valid satellite observations are available for all days in the month at that grid cell, whereas `observation_fraction = 0` indicates that all daily values at that grid cell are reconstructed by the fusion framework. This information is provided to prevent users from interpreting the entire gap-free product as purely observation-based.

Because of file-size and repository upload limitations, the daily data are organized into monthly NetCDF files rather than being provided as a single file for the entire study period. This monthly organization does not indicate temporal averaging. Each monthly file contains all daily fields for that calendar month.

The revised NetCDF files include CF-style metadata, including coordinate attributes, variable long names, units, missing-value definitions, temporal coverage, spatial resolution, product region, data source, author, institution, and contact information. The main variables included in the released NetCDF files are summarized as follows: `XCO` (ppb), fused column-averaged dry-air mole fraction of carbon monoxide; `XCH4` (ppb), fused column-averaged dry-air mole fraction of methane; `obs_mask` (unitless binary mask), indicating whether the grid-cell value is derived from a valid satellite observation (`1`) or reconstructed by the proposed fusion method (`0`); `observation_fraction` (unitless fraction from 0 to 1), indicating the monthly fraction of days with valid satellite observations at each grid cell; `time`, the daily time coordinate within each monthly file, defined using CF-style units; and `lat` and `lon`, the latitude and longitude coordinates of the gridded product.

Manuscript

Comment 2: *Language. Multiple grammatical and style issues, typos ("satelite"), missing spaces before references, etc. Professional English editing required.*

Response 2:

We sincerely apologize for the linguistic errors and typographical oversights in the original manuscript. We have carefully conducted a line-by-line professional English proofreading of the entire text.

Specifically, we have:
1. Corrected all identified typos (e.g., changing "satelite" to "satellite") and ensured consistent terminology throughout.
2. We have corrected the citation-related issues, including incomplete reference information and other reference formatting errors.
3. Ensured all acronyms are defined at their first use and used consistently thereafter.

Specifically, we corrected the spelling error in Line 57 of the original manuscript. The reference originally listed at Line 713 has been removed. The reference previously at Line 847 has been moved to Line 1221 of the revised manuscript, and the reference previously at Line 875 has been moved to Line 1268 of the revised manuscript. Regarding the revision of abbreviations,

representative examples are provided below; these examples appear in Lines 346 and 351 of the revised manuscript.

Revised text:

1. changing "satelite" to "satellite".

2. Wang, Y., Yuan, Q., Li, T., Yang, Y., Zhou, S., Zhang, L., 2023. Seamless mapping of long-term (2010–2020) daily global XCO2 and XCH4 from the Greenhouse Gases Observing Satellite (GOSAT), Orbiting Carbon Observatory 2 (OCO-2), and CAMS global greenhouse gas reanalysis (CAMS-EGG4) with a spatiotemporally self-supervised fusion method. Earth Syst. Sci. Data 15, 3597-3622. https://doi.org/10.5194/essd-15-3597-2023

Yang, Y., Zhou, M., Langerock, B., Sha, M.K., Hermans, C., Wang, T., Ji, D., Vigouroux, C., Kumps, N., Wang, G., 2020. New ground-based Fourier-transform near-infrared solar absorption measurements of XCO_2, XCH_4 and XCO at Xianghe, China. Earth Syst. Sci. Data 12, 1679-1696. https://doi.org/10.5194/essd-12-1679-2020

3. "Three-dimensional discrete cosine signal transform (DCT_3) rule," was revised to "The 3D DCT, denoted as DCT_3, is defined as:", and "Three-dimensional discrete cosine inverse transform ($IDCT_3$) rule," was revised to "The inverse 3D DCT, denoted as $IDCT_3$, is defined as:".

Comment 3: *Lack of the actual data product description. Useful would be a section e.g. Data product summary that covers:*

- *clear dataset specs: variables, domains, temporal coverage, file formats, QC flags,*
- *Table of key validation metrics across all TCCON sites,*
- *User guidance: strengths + limitations,*
- *File naming conventions and Zenodo structure*

Response 3:

Thank you for this helpful suggestion. We agree that the manuscript should include a clearer description of the actual released data products. We have therefore added a new "Fused dataset summary" subsection to the revised manuscript. This subsection describes the dataset specifications, including variables, spatial domains, temporal coverage, file format, coordinate variables, quality/observation indicators, and CF-style metadata.

We also clarified the Zenodo file organization and naming convention. The daily data are stored in monthly NetCDF files and uploaded as annual compressed archives, corresponding to 5 years × 2 gases × 2 spatial resolutions. The file names indicate the year/month, gas species, and spatial resolution. In addition, we added descriptions of the obs_mask and observation_fraction variables to help users distinguish values directly constrained by valid TROPOMI observations from reconstructed values.

The key validation metrics across all qualified TCCON sites are reported in Tables S4 and S5 and summarized in the validation section. User guidance on the strengths and limitations of the dataset has also been added in the uncertainty and limitation discussion, including cautions for

persistently cloudy or data-sparse regions. We have added the relevant description in Lines 853-892 of the revised manuscript.

Revised tables:

Table S4. (GEOS) Single-site validation results for GEOS-Chem, TROPOMI, and fused XCO data, with RMSE & μ in parts per billion (ppb).

SITE NAME	R^2			RMSE			μ		
	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION
burgos01	0.237	0.919	0.935	16.909	5.769	5.246	9.667	-2.799	-2.425
darwin01	-0.808	0.370	0.646	13.896	8.546	6.330	-9.241	-4.196	-2.848
easttroutlake01	-0.487	0.827	0.903	19.017	10.070	7.372	4.540	0.688	1.092
garmisch01	0.664	0.795	0.806	5.903	3.897	4.444	-1.453	0.737	1.113
hefei01	0.583	0.930	0.980	16.501	7.362	3.421	-15.897	-1.450	1.444
izana01	0.093	0.185	0.513	6.943	7.120	5.482	-3.692	5.790	3.909
karlsruhe01	0.521	0.411	0.596	7.055	6.892	6.441	4.002	-3.137	-4.337
lamont01	-0.134	0.534	0.580	9.297	6.447	6.013	5.209	-4.803	-3.650
lauder03	-0.524	0.849	0.663	11.990	3.893	5.666	-8.828	-2.228	-2.424
nicosia01	-0.251	0.868	0.847	8.114	2.791	3.008	5.978	0.557	-1.147
orleans01	0.581	0.791	0.537	6.704	6.656	8.947	3.012	-2.665	-2.786
parkfalls01	-0.468	0.899	0.811	19.445	6.597	8.677	7.525	-2.116	-2.106
reunion01	0.109	0.893	0.896	8.529	3.113	3.061	2.327	0.448	-0.859
rikubetsu01	0.635	0.961	0.899	7.674	2.893	4.545	1.122	0.064	0.438
sodankyla01	0.388	0.267	0.917	6.255	6.203	2.230	-0.775	2.807	1.381
tsukuba02	0.955	0.825	0.958	4.446	9.713	4.031	-2.817	-6.798	-1.684
wollongong01	-0.319	0.776	0.745	14.709	6.562	6.969	-5.625	-4.501	-4.418
xianghe01	0.279	0.269	0.553	11.965	10.749	9.469	-2.922	6.306	2.628

Table S5. (GEOS) Single-site validation results for GEOS-Chem, TROPOMI, and fused XCH_4 data, with RMSE & μ in parts per billion (ppb).

SITE NAME	R^2			RMSE			μ		
	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION
bremen01	-0.140	0.134	0.631	13.602	10.634	8.139	-0.024	-6.005	-2.375
burgos01	-0.280	0.741	0.877	24.716	13.217	8.892	-11.397	-7.712	-5.010
easttroutlake01	-0.560	0.371	0.322	27.858	18.152	20.471	-7.800	-10.457	-12.565
edwards01	-1.666	0.938	0.908	33.363	5.655	6.750	-20.863	2.314	4.679
garmisch01	-0.539	0.693	0.831	23.574	11.552	8.763	-4.177	-4.162	-3.266
hefei01	-1.130	0.882	0.770	38.129	8.789	12.575	-26.749	-3.298	-5.167
karlsruhe01	-1.466	0.546	0.561	31.590	13.937	14.156	-17.057	-8.526	-11.191
lamont01	-2.086	0.838	0.830	34.635	9.308	9.313	-22.403	-3.808	-4.785
lauder03	-1.178	0.725	0.742	33.162	12.366	12.116	-16.148	-7.392	-9.347
nicosia01	-2.566	0.670	0.893	32.223	10.568	5.954	-23.891	2.413	-1.740
orleans01	-1.780	0.551	0.705	30.640	13.366	11.070	-16.560	-6.472	-6.654
parkfalls01	-0.923	0.137	0.221	29.718	21.329	20.668	-12.079	-14.704	-16.099
pasadena01	-3.177	0.955	0.910	41.561	4.659	6.603	-32.970	-0.942	-3.476

rikubetsu01	0.230	0.740	0.715	15.539	9.189	9.892	6.588	-2.146	-4.784
tsukuba02	-0.582	0.777	0.712	15.420	6.238	7.147	-1.677	-2.308	-3.793
wollongong01	-1.396	0.195	0.347	29.981	19.964	17.320	-14.676	-12.975	-13.672
xianghe01	0.033	0.855	0.888	25.528	10.349	9.120	-9.444	-0.354	-1.247

Comment 4: *Make it clear what is novel versus previous fusion methods (e.g. Wang et al., 2023 https://doi.org/10.5194/essd-15-3597-2023).*

Response 4:

We thank you for pointing out this important reference. Wang et al. (2023) provided a valuable long-term global XCO_2 and XCH_4 dataset using a spatiotemporally self-supervised fusion method. Our study builds upon this existing body of work but differs in both methodological design and data characteristics. We have revised the manuscript to more clearly distinguish the novelty of our approach from previous fusion methods. The main innovations of our study are as follows:

1. Unlike previous methods that typically rely on a single reconstruction strategy, our framework introduces a two-stage hybrid approach: signal-domain reconstruction using DCT and SVD, followed by deep-learning refinement using a residual U-Net. This design separates the reconstruction of large-scale spatiotemporal structures from the correction of local residual patterns.

2. While Wang et al. (2023) focused on long-term global XCO_2 and XCH_4 products at 0.25° resolution, our study provides simultaneous gap-free XCO and XCH_4 datasets, including a 0.25° global product and a 0.05° regional product for China. The regional product provides enhanced spatial detail for analyzing local-scale variations over urban, fire-affected, and agricultural regions.

3. The target variables were extended from XCO_2/XCH_4 to XCO/XCH_4. In particular, CO and CH_4 share common sources and are chemically coupled through OH, making their joint analysis more environmentally meaningful.

We highlighted the novelty of this study in Lines 118–138 of the revised Introduction. The specific revisions are provided below.

Revised text:

The proposed method specifically addresses the shortcomings of existing approaches in three key aspects: (1) Addressing Data Sparsity and Non-linearity: Unlike pure machine learning interpolation that may struggle with unbalanced datasets, we employ a two-stage strategy. We first use 3D Discrete Cosine Transform (DCT) and Singular Value Decomposition (SVD) to perform low-rank signal-domain reconstruction (Rao and Yip, 2014; Wall et al., 2003). This step exploits the shared spatiotemporal structure between GEOS-Chem simulations and TROPOMI observations to approximate missing values in the frequency domain, providing a robust prior; (2) Enhancing Physical Consistency: To correct residual biases without overfitting, we implement a lightweight residual U-Net for pixel-level refinement (Ronneberger et al., 2015; Tang et al., 2022). Crucially, instead of using the TROPOMI data mask as a direct input feature, we incorporate it as

a weighting mask in the loss function (Wei et al., 2022). This masked learning strategy ensures that the model is supervised exclusively by valid observations, thereby preserving the physical consistency of the DCT/SVD-derived signal-domain reconstructed fields; (3) Improving Efficiency and Accuracy: By fusing model-driven priors and observational constraints in both frequency and spatial domains, our method significantly reduces computational load compared to enhanced spectral fitting while avoiding the "black-box" limitation of pure data-driven methods. Validation against independent TCCON observations demonstrates that the fused outputs outperform GEOS-Chem simulations alone and maintain accuracy comparable to or better than original TROPOMI retrievals in cloud-covered regions. This framework provides an efficient and interpretable solution for large-scale trace gas monitoring and offers new opportunities for atmospheric data assimilation and climate analysis.

Comment 5: *Reproducibility Gaps - Methods must be sufficiently described to allow reproduction.*

Response 5:

Thank you for raising this important comment. We agree that the methodology must be described in sufficient detail to ensure the reproducibility of the experimental results. In the revised manuscript, we have substantially expanded and clarified the data, method, model training, validation, and released-dataset descriptions.

Specifically, we expanded the TROPOMI data description by adding the data source, product IDs, processor versions, data streams, DOIs, study period, and the qa_value > 0.5 quality-screening criterion. We revised the GEOS-Chem description by specifying the model version, meteorological driver, simulation resolutions, model time steps, emission inventories, regridding procedure, and the post-processing used to derive XCO and $XCH_4$from vertical profiles. We also clarified the TCCON data release, site-selection criteria, and the distinction between representative sites and all qualified validation samples.

In addition, we revised the preprocessing and fusion workflow by detailing TROPOMI aggregation, GEOS-Chem regridding, resolution mismatch interpretation, DCT/SVD parameter settings, smoothing-parameter selection, retained singular-value threshold, and the threshold-sensitivity test. We further expanded the residual U-Net description by reporting the network structure, input/output channels, optimizer, learning rate, batch size, epochs, dropout setting, normalization, residual truncation, and masked-loss design.

Finally, we clarified the validation protocol by defining the temporal and spatial collocation criteria, missing-rate calculation, representative-site selection, diagnostic bias correction, and the use of descriptive validation metrics. We also added descriptions of the released NetCDF variables, including obs_mask and observation_fraction, to help users distinguish observation-constrained values from reconstructed values. These revisions improve the transparency and reproducibility of the proposed workflow. The main revisions are reflected in Lines 153-164 for the TROPOMI data description, Lines 166-195 for the GEOS-Chem configuration, Lines 197-233 for the TCCON data release and site-selection criteria, Lines 278-459 for the fusion workflow and residual U-Net settings, Lines 538-679 for the validation protocol, and Lines 850-889 for the NetCDF variables and obs_mask description.

Comment 6: *Validation description is insufficient. Contradictory descriptions of TCCON selection criteria (missing rate >0.5 vs. "highest availability" stations). "Manual bias correction" of GEOS-Chem is vaguely described – was this a single global offset, site-specific, applied before or after fusion? No uncertainty quantification: R^2/RMSE values given without confidence intervals or cross-validation across years/sites.*

Response 6:

We thank you for this important comment. We agree that our description of the site-selection criteria was not sufficiently clear, that the description of the GEOS-Chem correction was incomplete, and that uncertainty quantification was not included. We have revised the manuscript to address these issues.

First, we clarified the ambiguity caused by the inaccurate wording regarding "missing rate > 0.5" and "stations with the highest availability". The actual criterion for site selection was as follows: under sparse TROPOMI coverage conditions, site-day collocations with MR > 0.5 were retained. Then, based on these MR > 0.5 site-day collocations, stations with a relatively large number of valid collocated samples were selected to facilitate visualization.

Second, we clarified the treatment of GEOS-Chem in the validation. The raw, uncorrected GEOS-Chem simulations were used as the physical prior in the fusion and reconstruction framework. The diagnostic mean-bias adjustment was applied only for the GEOS-Chem–TCCON comparison and was not used to generate the fused datasets. This adjustment was site-specific: for each TCCON station and each gas, the mean difference between GEOS-Chem simulations and TCCON observations was calculated over all valid collocated days, and this station-specific mean bias was then removed from the GEOS-Chem time series at that station. This procedure was performed after the TCCON collocation step and only for diagnostic comparison, so it did not affect the fused products.

Third, to better evaluate the performance of the fused datasets, we added an analysis of the fused dataset performance under different levels of TROPOMI coverage. Three MR intervals were used: MR < 0.3, 0.3 ≤ MR ≤ 0.5, and MR > 0.5, corresponding to different levels of TROPOMI coverage. The resulting validation statistics are listed in Tables S6 and S7. These results show that, under low and moderate missing-rate conditions (i.e., high or moderate TROPOMI coverage), the fused XCO and XCH_4 datasets perform comparably to the TROPOMI datasets. Under high-missing-rate conditions (i.e., low TROPOMI coverage), the fused XCO dataset shows a higher R^2 and a lower RMSE than the TROPOMI dataset, while the fused XCH_4 dataset also shows slight improvements in R^2 and RMSE compared with TROPOMI.

These revisions are reflected in Lines 581-590, Lines 570-580 and Lines 644-662 of the revised manuscript.

we further evaluated site-level statistical agreement using scatterplots at selected representative stations under sparse TROPOMI coverage conditions in Figs. 4 and 5. This selection was based on the number of valid collocated samples after applying the missing-rate criterion, rather than on validation performance. Specifically, for each TCCON station, we first applied the qa_value > 0.5

quality filter, the 13:30 ± 1 h temporal window, the 2° spatial collocation radius, and the site-day missing-rate criterion (MR > 0.5). Subsequently, stations with a relatively large number of valid collocated samples after applying the missing-rate screening criterion were selected for visualization, whereas the overall validation statistics were calculated using all qualified high-missing-rate site-day collocations rather than only these representative stations.

GEOS-Chem simulations show a systematic low bias relative to TCCON observations. The raw, uncorrected GEOS-Chem simulations are used as the physical prior in the fusion and reconstruction framework. The mean-bias adjustment is applied only for the diagnostic GEOS-Chem–TCCON comparison and is not used in generating the fused datasets. For each TCCON station and each gas, we calculate the mean difference between GEOS-Chem simulations and TCCON observations over all valid collocated days, and then subtract this mean bias from the GEOS-Chem time series at that station. This procedure removes the station-mean offset while preserving the model's day-to-day variability, allowing a clearer assessment of whether GEOS-Chem captures the temporal variations observed by TCCON. All figures and tables involving this bias-adjusted GEOS-Chem diagnostic comparison are explicitly labeled to indicate that the GEOS-Chem results were bias-corrected before validation.

To demonstrate the performance of our fused datasets under different TROPOMI coverage, we conducted experiments using different MR thresholds. Tables S6 and S7 further quantify how the performance of the fused datasets varies with TROPOMI coverage. For XCO, when the missing rate is low or moderate (MR < 0.3 and 0.3 ≤ MR ≤ 0.5), TROPOMI shows slightly higher R^2 values and lower RMSE values than the fused dataset, indicating the high reliability of TROPOMI retrievals under relatively good coverage conditions. However, under sparse TROPOMI coverage conditions (MR > 0.5), the performance of TROPOMI decreases, with its R^2 decreasing to 0.89 and RMSE increasing to 6.16 ppb. In contrast, the fused XCO data maintain a higher R^2 value of 0.91 and a lower RMSE value of 5.65 ppb, indicating greater robustness under high-missing-rate conditions. For XCH_4, a similar pattern is observed. TROPOMI performs better than the fused product under low and moderate missing-rate conditions, with higher R^2 and lower RMSE. However, when MR > 0.5, TROPOMI performance decreases markedly, with R^2 declining to 0.81 and RMSE increasing to 15.84 ppb. The fused XCH_4 data show a slightly higher R^2 of 0.83 and a lower RMSE of 15.40 ppb under the same high-missing-rate condition. Although the fused XCH_4 bias is larger than that of TROPOMI, the higher R^2 and lower RMSE suggest that the fused dataset reduces overall errors under sparse-observation conditions. Overall, under conditions of good TROPOMI coverage, the fused XCO and XCH_4 datasets perform comparably to the TROPOMI datasets. When TROPOMI retrievals are sparse, the fused datasets show relatively high reliability. These results indicate that our fusion framework is effective for gap-filling applications.

Table S6. Validation statistics of XCO under different TROPOMI missing-rate intervals. GEOS-Chem results were bias-corrected before validation, and RMSE and μ are reported in parts per billion (ppb).

MR	R^2			RMSE			μ		
	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION
MR < 0.3	0.65	0.97	0.94	16.60	2.04	5.69	-0.22	-0.43	-1.10
0.3 ≤ MR ≤ 0.5	0.63	0.96	0.93	13.44	2.96	5.16	0.02	-0.88	-1.21

MR > 0.5	0.48	0.89	0.91	11.99	6.16	5.65	0.79	-0.87	-1.40

Table S7. Validation statistics of XCH_4 under different TROPOMI missing-rate intervals. GEOS-Chem results were bias-corrected before validation, and RMSE and μ are reported in parts per billion (ppb).

MR	R^2			RMSE			μ		
	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION
MR < 0.3	-0.15	0.95	0.87	29.10	3.05	9.66	-2.87	-0.57	-4.37
0.3 ≤ MR ≤ 0.5	-0.17	0.94	0.85	28.39	3.84	10.78	-3.19	-0.65	-2.17
MR > 0.5	0.34	0.81	0.83	29.84	15.84	15.40	-0.80	-5.48	-7.30

Comment 7: *There are a lot of very strong, overconfident conclusions/statements in the manuscript. Please tone them down or provide supporting data.*

Response 7:

We sincerely thank you for reminding us to maintain a cautious and scientifically rigorous interpretation of the results. We have carefully re-evaluated the manuscript and revised statements that may have sounded overly strong or insufficiently qualified. The main changes are summarized below.

1. More cautious wording when comparing the fused datasets with TROPOMI. We replaced overly strong expressions such as "outperforms TROPOMI," "exceeds TROPOMI," and "superior to TROPOMI" with more precise descriptions, such as "shows improved agreement with TCCON relative to the original TROPOMI retrievals," "shows comparable or improved statistical consistency," and "reduces random scatter and improves spatiotemporal continuity." These revisions avoid implying that the fused datasets are intrinsically more accurate than direct satellite observations.

2. More careful interpretation of temporal changes. Because the dataset covers only 2019–2023, we revised terms such as "long-term trends" and "annual trends" to more cautious wording, including "short-term rates of change," "five-year rate-of-change analysis," and "fitted annual rates of change during 2019–2023."

3. Clearer discussion of methodological limitations. We added a dedicated "Discussion of uncertainties and limitations" section to acknowledge user-relevant uncertainties, including AK-related uncertainty in satellite–model comparisons, possible propagation of GEOS-Chem prior biases in regions with persistent TROPOMI data gaps, uncertainty in fine-scale structures below the native GEOS-Chem resolution, and extrapolation uncertainty associated with the masked residual U-Net.

4. More precise description of validation and case analyses. We clarified that representative TCCON stations were used for visualization, whereas the overall validation statistics were calculated using all qualified site-day collocations. We also refined the definitions of the missing-rate criterion, fire-affected area, reference period, and rate-of-change calculations to make the event and regional analyses more reproducible.

5. Revision of the conclusion and terminology. We revised the Conclusion section to emphasize that the fused datasets are most suitable for gap-free daily mapping, daily-to-seasonal spatiotemporal variability analysis, and regional event assessment during 2019–2023, rather than for robust long-term climate-trend attribution. We also standardized the terminology throughout the manuscript by consistently using "fused dataset(s)" or "fused data."

These revisions were made throughout the manuscript to provide a more neutral, transparent, and scientifically cautious presentation of the dataset, results, and limitations.

Comment 8: *I would be very careful about drawing any conclusions on trends calculated over such a short period; I would rather refer to them as short-term changes. The associated uncertainty and statistical significance should be assessed. One option is to perform a regression on deseasonalized data (e.g., with the annual cycle removed) and report the confidence interval of the trend, and optionally the trend significance using the Mann–Kendall test and Sen's slope estimator.*

Response 8:

Thank you for raising this important suggestion. We also believe that the period of 2019–2023 is relatively short. To avoid overinterpretation, we revised the original manuscript by replacing terms such as "trend" and "long-term trend" with the more cautious expression "short-term change". Also following your suggestion related to the analysis method, we reanalyzed the short-term rates of change separately for XCO and XCH_4 using both the TROPOMI and fused datasets over rice-growing regions in Northeast and Southeast China during 2019–2023, and reported the 95% confidence intervals.

Specifically, to reduce the influence of seasonal cycles, we applied harmonic regression to the daily regional mean XCO and XCH_4 time series. The regression model includes a linear term, an annual periodic term, and a semiannual periodic term. The coefficient of the linear term represents the short-term rate of change after removing the seasonal cycle, and the 95% confidence interval was calculated from the regression standard error. The relevant methods and results have been added to the revised manuscript. This content can be found in Lines 746–777 of the revised manuscript.

Revised text:

Therefore, we further analyzed the short-term temporal changes in XCO and XCH_4 over the major rice-growing regions of Northeast and Southeast China to evaluate the applicability of the high-resolution fused datasets in agricultural emission-related regions. Figures 8 and 9 show the temporal variations in XCO and XCH_4 derived from TROPOMI and the fused data over these regions, respectively. To reduce the influence of the seasonal cycle, the short-term rates of change were derived using harmonic regression with annual and semiannual terms, and the corresponding 95% confidence intervals are provided in Tables S9 and S10.

The results show that XCO exhibits negative short-term rates of change in both the satellite and fused datasets across the rice cultivation areas of Northeast and Southeast China. In Northeast China, the XCO rate is −2.36 ppb/year for TROPOMI and −2.28 ppb/year for the fused data. In

Southeast China, the XCO rate is −1.29 ppb/year for TROPOMI and −1.76 ppb/year for the fused data. The consistent negative rates indicate that the fused dataset preserves the main decreasing XCO signal observed by TROPOMI. Differences between the two datasets may be partly related to the uneven temporal sampling of TROPOMI caused by clouds, aerosols, and retrieval filtering, whereas the fused dataset provides spatially and temporally continuous fields.

In contrast, XCH_4 exhibits positive short-term rates of change in both the satellite and fused datasets. In Northeast China, the XCH_4 rate is 12.08 ppb/year for TROPOMI and 12.16 ppb/year for the fused data. In Southeast China, the XCH_4 rate is 12.16 ppb/year for TROPOMI and 11.78 ppb/year for the fused data. The close agreement between the two datasets indicates that the fused dataset preserves the main increasing XCH_4 signal observed by TROPOMI while providing improved temporal continuity.

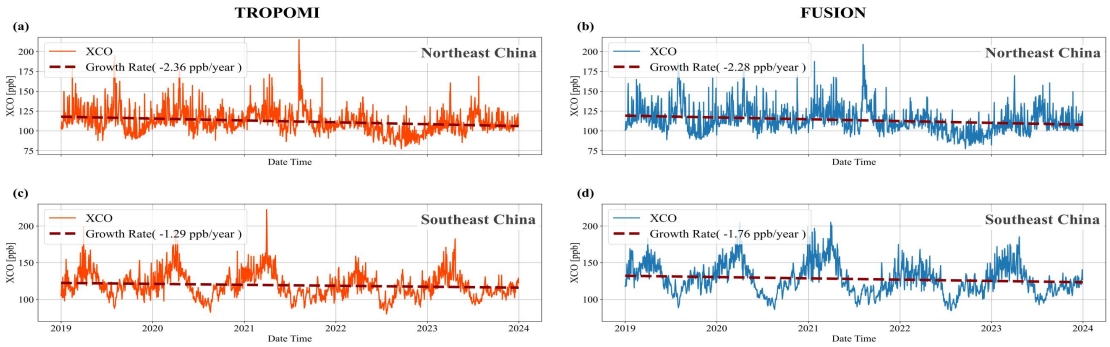


Figure 8. Daily time series of TROPOMI and fused XCO data over rice-growing regions in (a, b) Northeast China and (c, d) Southeast China. Red dashed lines indicate short-term rates of change derived from harmonic regression after accounting for the seasonal cycle, with values shown in parentheses. Unit: parts per billion per year.

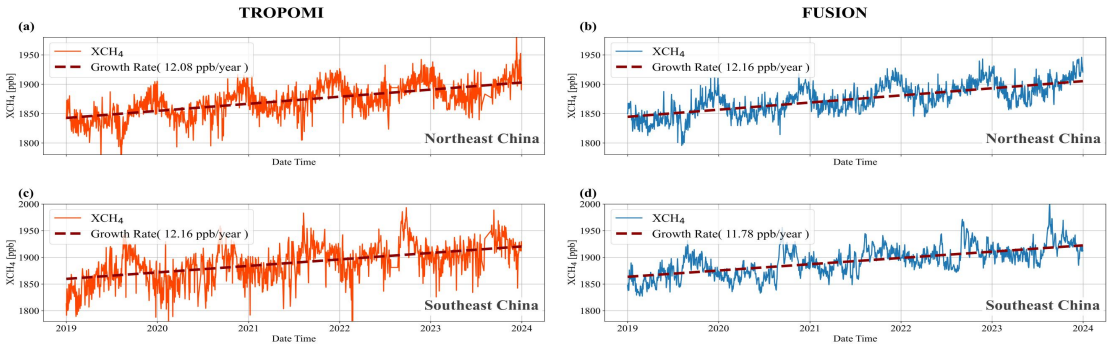


Figure 9. Daily time series of TROPOMI and fused XCH_4 data over rice-growing regions in (a, b) Northeast China and (c, d) Southeast China. Red dashed lines indicate short-term rates of change derived from harmonic regression after accounting for the seasonal cycle, with values shown in parentheses. Unit: parts per billion per year.

Table S9. Seasonally adjusted short-term rates of change and 95% confidence intervals for XCO over rice-growing regions during 2019–2023. Unit: ppb/year.

Region	Rate of change		95% confidence interval	
	TROPOMI	Fusion	TROPOMI	Fusion
Northeast China	−2.36	−2.28	−3.30 to −1.43	−3.30 to −1.27

Region	Rate of change		95% confidence interval	
	TROPOMI	Fusion	TROPOMI	Fusion
Southeast China	−1.29	−1.76	−2.15 to −0.43	−2.79 to −0.72

Table S10. Seasonally adjusted short-term rates of change and 95% confidence intervals for XCH_4 over rice-growing regions during 2019–2023. Unit: ppb/year.

Region	Rate of change		95% confidence interval	
	TROPOMI	Fusion	TROPOMI	Fusion
Northeast China	12.08	12.16	11.24 to 12.92	11.27 to 13.04
Southeast China	12.16	11.78	10.83 to 13.49	9.72 to 13.83

Comment 9: *Acronyms should be defined at first use (plus in the abstract) and their expansions should not be repeated later.*

Response 9:

We apologize for the inconsistent use of acronyms in the original manuscript. We have reviewed the manuscript to improve acronym usage and reduce redundant expansions. Acronyms such as TROPOMI, DCT, SVD, and TCCON are defined at their first mention in the Abstract and again at their first mention in the main text, where appropriate. After definition, the abbreviated forms are used throughout the remainder of the manuscript.

We also corrected inconsistent notation, including the mixed use of XCH4 and XCH_4, and removed redundant expansions in later sections to improve conciseness and readability. Reference titles were kept unchanged to preserve the original citation information. Some examples of the revisions are provided below, appearing in Line 346 and Line 351 of the revised manuscript.

Examples of revisions:

"Three-dimensional discrete cosine signal transform (DCT_3) rule," was revised to "The 3D DCT, denoted as DCT_3, is defined as:", and "Three-dimensional discrete cosine inverse transform ($IDCT_3$) rule," was revised to "The inverse 3D DCT, denoted as $IDCT_3$, is defined as:".

Comment 10: *Gap-free sounds great. But does it mean that f.e. oceanic Intertropical Convergence Zone cloudy regions are 90% ML?*

Response 10:

Thank you for raising this important point. We agree that the term "gap-free" should not be interpreted as meaning that every grid-cell value is directly constrained by valid TROPOMI observations. In persistently cloudy regions, such as the oceanic Intertropical Convergence Zone (ITCZ), valid TROPOMI retrievals can be sparse, and the fused values may rely more strongly on the reconstruction framework and the GEOS-Chem prior.

To make this distinction transparent, we added an obs_mask variable and an observation_fraction variable to the released NetCDF files. The obs_mask variable is provided to indicate whether the value at each grid cell is a TROPOMI observation or reconstructed by the fusion framework. The

observation_fraction variable summarizes the fraction of days within each monthly file for which valid TROPOMI observations are available at each grid cell. Thus, users can identify regions such as the ITCZ where the fraction of direct satellite observations is low and interpret the reconstructed values accordingly.

We have also expanded the uncertainty and limitation discussion to clarify that, in persistently cloudy or data-sparse regions, the reconstruction relies more strongly on the GEOS-Chem prior and the DCT/SVD spatiotemporal structure, and residual relationships learned from valid-observation regions. We also note that reconstructed values in such regions should be interpreted with caution. These revisions are reflected in Lines 813-819 and Lines 828-837 of the revised manuscript.

Revised text:

In regions or periods where valid TROPOMI retrievals are persistently sparse—such as the Intertropical Convergence Zone, high-latitude zones, or heavily polluted areas with dense cloud or aerosol cover—the reconstruction relies more strongly on the GEOS-Chem prior and the DCT/SVD spatiotemporal structure. Consequently, systematic biases in the prior model may propagate into the fused fields. Users are therefore advised to consult the obs_mask variable provided in the dataset to assess observational support and to interpret the reconstructed values cautiously in persistently missing regions.

The masked residual-learning strategy itself introduces uncertainty in persistently missing regions. Because the residual U-Net is trained only at valid TROPOMI pixels, the model learns the statistical relationship between the preliminary reconstruction error and the input predictors under observed conditions. When applied to unobserved regions—especially those with persistent cloud cover or aerosol contamination—the model infers the most likely residual pattern given similar meteorological and spatiotemporal conditions, rather than the exact unobserved atmospheric state on a specific day. This does not invalidate the gap-filling framework, but it implies that reconstructed values in long-term data gaps should be understood as observation-constrained estimates rather than direct measurements. Users should be aware of this extrapolation uncertainty when analyzing local short-lived anomalies in persistently cloudy regions.

Specific comments

Comment 11: *Introduction. I would suggest a shorter general background and a clearer focus on: (a) the current state of satellite/model CO and CH_4 products, (b) identified data gaps, and (c) how your new dataset addresses them.*

Response 11:

Thank you for this constructive suggestion. We agree that the Introduction should be more concise and more clearly focused on the data-product motivation of the study. Accordingly, we revised the Introduction to reduce the general background and to emphasize three key aspects: the current state of satellite- and model-based CO and CH_4 products, the remaining data gaps, and how the dataset developed in this study addresses these gaps.

Specifically, we shortened and reorganized the background on CO and CH_4, retaining only the information necessary to explain their atmospheric relevance and the rationale for joint reconstruction. We then clarified the current status of satellite observations, including MOPITT, AIRS, TES, IASI, and TROPOMI, as well as the role of model-based information and fusion approaches in improving spatial and temporal continuity.

We also revised the Introduction to explicitly identify the limitations of current XCO and XCH_4 products. Although TROPOMI provides high-resolution daily observations, its retrievals remain incomplete in regions affected by clouds, aerosols, and limited valid observations. Existing gap-filling and fusion approaches can improve data coverage, but they still face challenges related to data sparsity, dependence on prior simulations, computational cost, and insufficient integration of physical constraints with observational information.

Finally, we added a clearer statement of how our new dataset addresses these gaps. The revised Introduction now emphasizes that our signal-domain-guided spatio-temporal fusion framework integrates GEOS-Chem simulations, TROPOMI observations, frequency-domain low-rank reconstruction, and residual U-Net refinement to generate continuous daily XCO and XCH_4 products for 2019-2023 at 0.25° global resolution and 0.05° resolution over China. These revisions are reflected in Lines 42-138 of the revised manuscript.

Revised text:

Carbon monoxide (CO) and methane (CH_4) are two critical atmospheric gases that play a key role in air quality monitoring and climate change research. CO is primarily produced through incomplete combustion of fossil fuels and biomass, as well as the oxidation of CH_4 and other hydrocarbons. Its main sink is the reaction with hydroxyl radicals (OH), which contributes to the formation of tropospheric ozone (O_3) and greenhouse gases such as carbon dioxide (CO_2) (Lelieveld et al., 2016; Spivakovsky et al., 2000). Due to its atmospheric lifetime of weeks to months, CO serves as an effective tracer for pollutant transport, facilitating the study of both horizontal and vertical atmospheric movements (Heald et al., 2003). CH_4 is a powerful greenhouse gas with a significantly longer atmospheric lifespan—around ten years—and a greater global warming potential compared to numerous other gases (Filonchyk et al., 2024; Heilig, 1994). It originates from both natural sources, such as wetlands, permafrost, and wildfires, and anthropogenic activities, including agriculture, livestock digestion, landfills, and fossil fuel extraction (Chai et al., 2016; Jackson et al., 2020). CH_4 plays a crucial role in atmospheric chemistry, influencing the oxidative capacity of the atmosphere and contributing to tropospheric ozone formation. Its rising concentration is a major driver of climate change (Lelieveld et al., 2016).

Although CO and CH_4 are often discussed separately in terms of their sources and impacts, they are chemically and dynamically coupled in the atmosphere. Both gases are predominantly removed by the hydroxyl radical (OH), meaning that changes in one can affect the atmospheric lifetime and concentration of the other through OH competition (Lelieveld et al., 2016). Moreover, they share major emission sources, particularly biomass burning and fossil-fuel-related activities such as incomplete combustion and flaring, where they can be emitted simultaneously (Andreae and Merlet, 2001; Akagi et al., 2011; Andreae, 2019). Their co-emission and co-transport imply that the spatial patterns of CO and CH_4 may be correlated near fire hotspots, industrial regions,

and other combustion-influenced areas. Therefore, joint reconstruction of XCO and XCH_4 offers two key advantages: (1) shared spatial structure and common meteorological drivers can provide complementary constraints for deep learning models, improving generalization in data-sparse regions; (2) simultaneous monitoring can support more consistent estimation of emission ratios and source attribution, as demonstrated by studies using CH_4/CO or related trace-gas ratios to constrain biomass-burning and fossil-fuel methane emissions (Worden et al., 2013; Worden et al., 2017; Ramsden et al., 2022). A comprehensive understanding of these coupled gases is fundamental to improving air quality predictions and formulating climate mitigation policies.

Global satellite monitoring enables effective tracking and analysis of the sources, transport, and removal processes of atmospheric pollutants. At present, several satellite instruments provide observations of CO and/or CH_4, including the Measurements of Pollution in the Troposphere (MOPITT), which primarily focuses on CO retrievals (Deeter et al., 2003), as well as the Atmospheric Infrared Sounder (AIRS) (McMillan et al., 2005), Tropospheric Emission Spectrometer (TES) (Rinsland et al., 2006), and Interferometer for the Exploration of the Atmosphere in the Infrared (IASI) (Turquety et al., 2004), which are capable of retrieving multiple trace gases, including CO and CH_4, using different spectral channels and retrieval algorithms. These instruments offer critical data support for the global distribution of CO and CH_4 in the atmosphere.

The TROPOspheric Monitoring Instrument (TROPOMI) onboard the European Space Agency's Sentinel-5P mission provides higher daily global coverage and finer spatial resolution than previous observing missions. TROPOMI provides observations of atmospheric CO and CH_4 in the form of column-averaged dry-air mole fractions (denoted as XCO and XCH_4, respectively). TROPOMI has already been employed in numerous pertinent applications for XCO and XCH_4 investigations, including the calculation of emissions from biomass combustion (Goudar et al., 2023; Griffin et al., 2024). However, the continuity of TROPOMI atmospheric column retrievals is affected by clouds and aerosols, which can attenuate or scatter the reflected solar radiation measured by satellite sensors. As a result, observations may be incomplete, particularly in regions with high cloud cover and aerosol load—such as the humid tropics and high-latitude zones (Gao et al., 2023), Southeast Asia (Valerio et al., 2025), and eastern China during intense pollution episodes (Wang et al., 2021).

The generation of continuous TROPOMI XCO and XCH_4 products has been addressed through a range of gap-filling approaches. First, machine learning-based interpolation methods (X. Chen et al., 2022; Hu et al., 2022; Valerio et al., 2025; Wei et al., 2025) reconstruct missing observations by learning nonlinear spatiotemporal relationships. However, their performance strongly depends on the quality and representativeness of the training data, and biased or insufficient samples may lead to unreliable predictions. Second, enhanced spectral fitting algorithms (Borsdorff et al., 2019; Guanter et al., 2015; Schneising et al., 2023; Wang et al., 2020) improve retrieval continuity by refining the spectral inversion process, for example through higher-order polynomial fitting to reduce reflectance-induced biases, albeit at increased computational cost. Third, physics–data–driven coupling approaches, including model fusion, data assimilation, and hybrid machine learning frameworks, integrate satellite retrievals with chemical transport models or reanalysis datasets (Fritz et al., 2022; Schneising et al., 2023; Sicard et al., 2021). By combining physical constraints with data-driven learning or observational corrections, these methods improve

spatial consistency and temporal continuity while reducing retrieval bias. However, their performance remains dependent on uncertainties in prior model simulations and the representativeness of assimilated observations (Inness et al., 2022; Wang et al., 2023).

Despite these advances, current methods face challenges in balancing computational efficiency with physical consistency, particularly in regions with persistent cloud cover. To overcome these limitations, this study proposes a novel signal-domain guided spatio-temporal fusion framework. Our core objective is to generate daily global and regional continuous XCO and XCH_4 products (2019–2023) at high resolution (0.25° globally and 0.05° over China) by effectively leveraging complementary information from chemical transport modeling and frequency-domain representations.

The proposed method specifically addresses the shortcomings of existing approaches in three key aspects: (1) Addressing Data Sparsity and Non-linearity: Unlike pure machine learning interpolation that may struggle with unbalanced datasets, we employ a two-stage strategy. We first use 3D Discrete Cosine Transform (DCT) and Singular Value Decomposition (SVD) to perform low-rank signal-domain reconstruction (Rao and Yip, 2014; Wall et al., 2003). This step exploits the shared spatiotemporal structure between GEOS-Chem simulations and TROPOMI observations to approximate missing values in the frequency domain, providing a robust prior; (2) Enhancing Physical Consistency: To correct residual biases without overfitting, we implement a lightweight residual U-Net for pixel-level refinement (Ronneberger et al., 2015; Tang et al., 2022). Crucially, instead of using the TROPOMI data mask as a direct input feature, we incorporate it as a weighting mask in the loss function (Wei et al., 2022). This masked learning strategy ensures that the model is supervised exclusively by valid observations, thereby preserving the physical consistency of the DCT/SVD-derived signal-domain reconstructed fields; (3) Improving Efficiency and Accuracy: By fusing model-driven priors and observational constraints in both frequency and spatial domains, our method significantly reduces computational load compared to enhanced spectral fitting while avoiding the "black-box" limitation of pure data-driven methods.

Validation against independent TCCON observations demonstrates that the fused outputs outperform GEOS-Chem simulations alone and maintain accuracy comparable to or better than original TROPOMI retrievals in cloud-covered regions. This framework provides an efficient and interpretable solution for large-scale trace gas monitoring and offers new opportunities for atmospheric data assimilation and climate analysis.

Comment 12: *line 22. "The fused results outperform GEOS-Chem and are comparable or superior to TROPOMI." Those are very strong conclusions, yet with unclear supporting data (R^2 values - how calculated?)*

Response 12:

We thank you for your attention to this question. In the revised manuscript, we have weakened the relevant statement and provided the calculation method for R^2. The revised text is provided below, corresponding to Lines 28–32 and Lines 492–496 of the revised manuscript.
revised statement:

Missing-rate-threshold experiments further show that the fused products perform comparably to original TROPOMI retrievals under low and moderate missing-rate conditions and show improved performance under sparse TROPOMI coverage (MR > 0.5), with R^2 values of 0.91 for XCO and 0.83 for XCH_4, along with reduced biases and standard deviations.

Clarification of R^2 Calculation :

The TROPOMI satellite overpasses each region at approximately 13:30 local time; therefore, the mean column concentrations from TCCON sites within 13:30 ± 1 h local time are used as ground-based validation data. For spatial consistency, GEOS-Chem simulations, TROPOMI observations, and the fused datasets were extracted within a 2°radius around each TCCON site for comparison.

Comment 13: *line 63. Acronym, first use, please expand.*

Response 13:

We sincerely apologize for this oversight. We have now spelled out the abbreviation at its first occurrence. The specific revision is as follows. We have added the relevant description in Lines 82–84 of the revised manuscript.

Revised text:

The TROPOspheric Monitoring Instrument (TROPOMI) onboard the European Space Agency's Sentinel-5P mission provides higher daily global coverage and finer spatial resolution than previous observing missions.

Comment 14: *lines 63-74. TROPOMI description appears later again in Sec. 2.1.1 (lines 119–126), with redundant information. I suggest keeping a concise overview in the Introduction, and moving the detailed orbit/spectral descriptions entirely to Sec. 2.1.1 to avoid repetition.*

Response 14:

Thank you for this helpful suggestion. We agree that the original Introduction contained technical details of TROPOMI that were repeated later in Section 2.1.1, which could reduce the clarity and conciseness of the manuscript. In response to your suggestion, we revised both sections to avoid this redundancy. We have added the relevant description in Lines 82–94 and 156–161 of the revised manuscript.

Revised text:

The TROPOspheric Monitoring Instrument (TROPOMI) onboard the European Space Agency's Sentinel-5P mission provides higher daily global coverage and finer spatial resolution than previous observing missions. According to the official Sentinel-5P/TROPOMI Level-2 product table, the CO and CH_4 products are provided as total column products. In this study, XCO and XCH_4 refer to the column-averaged dry-air mole fractions of atmospheric CO and CH_4, respectively, derived from the corresponding TROPOMI total column products. TROPOMI has already been employed in numerous pertinent applications for XCO and XCH_4 investigations, including the calculation of emissions from biomass combustion (Goudar et al., 2023; Griffin et

al., 2024). However, the continuity of TROPOMI atmospheric column retrievals is affected by clouds and aerosols, which can attenuate or scatter the reflected solar radiation measured by satellite sensors. As a result, observations may be incomplete, particularly in regions with high cloud cover and aerosol load—such as the humid tropics and high-latitude zones (Gao et al., 2023), Southeast Asia (Valerio et al., 2025), and eastern China during intense pollution episodes (Wang et al., 2021).

The TROPOMI Level-2 CO and CH_4 products used in this study were obtained from the NASA GES DISC Sentinel-5P/TROPOMI archive for the period 2019–2023. We used the OFFL (Offline) datasets. Specifically, the CO product used is S5P_L2__CO_____HiR, Version 02 (DOI: 10.5270/S5P-bj3nry0), while the CH_4 product used is S5P_L2__CH4____HiR, Version 02 (DOI: 10.5270/S5P-3lcdqiv). These Version 02 products include updated spectroscopic parameters and improved treatments for non-scattering clouds compared with earlier releases.

Comment 15: *lines 71-72. You give information that TROPOMI measures "XCO and XCH_4 at the surface". Please correct this sentence (also, the wording of the latter part of this sentence should be improved).*

Response 15:

Thank you for pointing out this inaccurate wording. We have revised the sentence to clarify that TROPOMI retrieves atmospheric column-averaged XCO and XCH_4, rather than measuring these quantities “at the surface”. We also improved the wording of the latter part of the sentence by explaining that clouds and aerosols can attenuate or scatter the reflected solar radiation measured by the satellite sensor, leading to incomplete retrievals.

In addition, we added specific regional examples where such retrieval gaps are more likely to occur, including the humid tropics and high-latitude regions (Gao et al., 2023), Southeast Asia (Valerio et al., 2025), and eastern China during intense pollution episodes (Wang et al., 2021). We have added the relevant description in Lines 85–94 of the revised manuscript.

Revised text:

In this study, XCO and XCH_4 refer to the column-averaged dry-air mole fractions of atmospheric CO and CH_4, respectively, derived from the corresponding TROPOMI total column products. TROPOMI has already been employed in numerous pertinent applications for XCO and XCH_4 investigations, including the calculation of emissions from biomass combustion (Goudar et al., 2023; Griffin et al., 2024). However, the continuity of TROPOMI atmospheric column retrievals is affected by clouds and aerosols, which can attenuate or scatter the reflected solar radiation measured by satellite sensors. As a result, observations may be incomplete, particularly in regions with high cloud cover and aerosol load—such as the humid tropics and high-latitude zones (Gao et al., 2023), Southeast Asia (Valerio et al., 2025), and eastern China during intense pollution episodes (Wang et al., 2021).

Comment 16: *lines 107-109. Add one explanatory sentence on why this is advantageous, and clarify whether mask information influences extrapolation behaviour in heavily cloudy regions.*

Response 16:

Thank you for this helpful suggestion. We have revised Lines 107-109 to clarify the role of the TROPOMI observation mask in the learning framework. In our framework, the TROPOMI mask is used only to define which pixels contribute to the training loss. Specifically, residual errors are calculated only at pixels with valid TROPOMI retrievals, while pixels without valid TROPOMI observations because of clouds, aerosols, or retrieval failure are excluded from the loss calculation. The mask itself is not provided to the residual U-Net as an input variable.

This means that, during inference, the residual U-Net does not use cloud/missing-data information to directly determine the prediction at a given pixel. Instead, it predicts residual corrections for all grid cells from the available input fields, including GEOS-Chem simulations and the DCT/SVD-reconstructed fields. Therefore, predictions in heavily cloudy or persistently missing regions are not directly controlled by the mask; they are inferred from the learned relationships between valid-observation regions and the physically based input fields. We have added the relevant description in Lines 124–130 of the revised manuscript.

Revised text:

Enhancing Physical Consistency: To correct residual biases without overfitting, we implement a lightweight residual U-Net for pixel-level refinement (Ronneberger et al., 2015; Tang et al., 2022). Crucially, instead of using the TROPOMI data mask as a direct input feature, we incorporate it as a weighting mask in the loss function (Wei et al., 2022). This masked learning strategy ensures that the model is supervised exclusively by valid observations, thereby preserving the physical consistency of the reconstructed fields;

Comment 17: *lines 121-123. "Sentinel-5 Precursor" and "equipped with TROPOMI" repeated twice. Please rewrite those sentences.*

Response 17:

Thank you for pointing out the repetition regarding the Sentinel-5 Precursor satellite and TROPOMI. We have revised the manuscript and removed the redundant statement. The revised text has been placed in lines 143–145 of the revised manuscript.

Revised text:

The TROPOMI instrument onboard the Sentinel-5 Precursor (S5P) satellite operates in a polar sun-synchronous orbit, and provides daily global XCO and XCH_4 observations at approximately 13:30 local solar time, thereby enabling near-global daily coverage.

Comment 18: *Sec. 2.1.1. Specify version numbers and product IDs explicitly.*

Response 18:

Thank you for this important comment. We have revised Section 2.1.1 to explicitly specify the TROPOMI product IDs, product versions, data streams, and DOIs used in this study. We have

added the relevant description in Lines 156–161 of the revised manuscript.

Revised text:

The TROPOMI Level-2 CO and CH_4 products used in this study were obtained from the NASA GES DISC Sentinel-5P/TROPOMI archive for the period 2019–2023. We used the OFFL (Offline) datasets. Specifically, the CO product used is S5P_L2__CO_____HiR, Version 02 (DOI: 10.5270/S5P-bj3nry0), while the CH_4 product used is S5P_L2__CH4____HiR, Version 02 (DOI: 10.5270/S5P-3lcdqiv). These Version 02 products include updated spectroscopic parameters and improved treatments for non-scattering clouds compared with earlier releases.

Comment 19: *line 134. You select QV following Kawka et al. (2021), but that work focuses on NO_2. Can it be applied and why to the CO/CH_4.*

Response 19:

Thank you for pointing this out. We agree that Kawka et al. (2021), which focuses on NO_2, is not the most appropriate reference for the quality screening of TROPOMI XCO and XCH_4 products. We have revised the manuscript by removing this citation from the quality-screening description. In the revised manuscript, the qa_value > 0.5 threshold is justified using product-specific guidance for the Sentinel-5P/TROPOMI CO and CH_4 Level-2 products. Specifically, the Sentinel-5P CO Product Readme recommends using TROPOMI CO retrievals with qa_value > 0.5 (*S5P MPC Product Readme Carbon Monoxide*), and the Sentinel-5P CH_4 Product Readme similarly recommends retaining CH_4 pixels with qa_value > 0.5 to avoid misinterpretation of data quality (*S5P MPC Product Readme Methane*). We have added the relevant description in Lines 163–167 of the revised manuscript.

Revised text:

To ensure the quality and reproducibility of the satellite observations, we applied product-specific quality screening and retained only TROPOMI retrievals with a quality value (qa_value) greater than 0.5 for subsequent analysis, following the recommended quality-screening threshold for the Sentinel-5P CO and CH_4 Level-2 products.

Comment 20: *line 141-144. Repetition.*

Response 20:

Thank you for pointing out this repetition. We have revised the sentence in Section 2.1.2 to remove the duplicated phrase. We have added the relevant description in Lines 169–171 of the revised manuscript.

Revised text:

This study employs the GEOS-Chem model, a global three-dimensional atmospheric chemistry model driven by meteorological input from the Goddard Earth Observing System (GEOS) of the NASA Global Modeling and Assimilation Office (GMAO).

Comment 21: *lines 150-152. What's the implication of the spatial resolution of GEOS-Chem on the spatial resolution of the output dataset? How was regridding done?*

Response 21:

Thank you for this important comment. We have revised Section 2.1.2 to clarify the relationship between the native resolution of GEOS-Chem and the target resolution of the fused datasets. The global GEOS-Chem simulation was conducted at a resolution of 2° × 2.5° and regridded to the 0.25° × 0.25° target grid, and the TROPOMI data were also gridded to the 0.25° × 0.25° target grid. The China regional simulation was conducted at a resolution of 0.25° × 0.3125° and regridded to the 0.05° × 0.05° target grid, and the TROPOMI data were also gridded to the 0.05° × 0.05° target grid. The resolution of the output datasets is determined by the target grid. Regridding was performed using inverse distance weighting (IDW). We have added the relevant description in Lines 194–198 of the revised manuscript.

Revised text:

In the proposed fusion framework, these GEOS-Chem simulations serve as large-scale physical priors. To match the spatial resolution of satellite observations, the global 2° × 2.5° GEOS-Chem fields were regridded to a 0.25° × 0.25° target grid, while the China 0.25° × 0.3125° fields were regridded to a 0.05° × 0.05° target grid. Regridding was performed using inverse distance weighting (IDW) to interpolate GEOS-Chem fields onto the target grids.

Comment 22: *line 154. Clarify whether XCO/XCH_4 are model native outputs or diagnosed from species, and how column averaging (dry-air) is computed.*

Response 22:

Thank you for pointing out this ambiguity. We have revised Section 2.1.2 to clarify that XCO and XCH_4 are not direct native outputs of GEOS-Chem. GEOS-Chem simulates three-dimensional CO and CH_4 fields based on emissions, transport, chemistry, and meteorological processes. The column-averaged dry-air mole fractions, XCO and XCH_4, were therefore diagnosed during post-processing from the simulated vertical profiles.

Specifically, the simulated CO and CH_4 dry-air mole fraction profiles were vertically averaged using the pressure thickness of each model layer as the weighting factor. This pressure-weighted vertical averaging produced model-derived column-averaged dry-air mole fractions that are directly comparable in units and physical meaning to the XCO and XCH_4 quantities reported by TROPOMI and TCCON, although no satellite averaging kernels or prior profiles were applied in this diagnostic calculation. The resulting GEOS-Chem-derived XCO and XCH_4 fields were then converted to parts per billion (ppb) for use in the fusion framework. We have added the relevant description in Lines 188–193 of the revised manuscript.

Revised text:

Based on these inputs and physical processes, GEOS-Chem simulates three-dimensional CO and CH_4 fields. However, XCO and XCH_4 are not direct native outputs; they were diagnosed during post-processing from the simulated vertical profiles. Specifically, the dry-air mole fraction profiles were vertically averaged using pressure-layer thickness as weighting, producing column-averaged dry-air mole fractions consistent with the definitions used by TROPOMI and TCCON. The resulting fields were converted to parts per billion (ppb) for use in the fusion framework.

Comment 23: *Sec. 2.1.3. Explain the site selection criteria in detail. Please clarify: whether all stations in Tab. 1 are used in global validation, which are used only for examples, and any filtering by cloud cover or data gaps beyond what is stated later in Sec. 3.1.*

Response 23:

We thank you for this valuable comment. We agree that the description of the site-selection criteria in the original manuscript may have caused ambiguity. In the revised manuscript, we have clarified the roles of the different TCCON stations used in this study.

All TCCON stations listed in Table 1 were first considered for the TCCON-based validation. For each gas, stations were retained for quantitative validation if they had valid site-day collocations after applying the 13:30 ± 1 h local-time window, the 2° spatial collocation radius, and the TROPOMI quality filter (qa_value > 0.5). The representative stations shown in the time-series figures and scatterplots were selected only for visualization purposes. Specifically, the time-series examples were selected from stations with relatively continuous multi-year collocated records and limited long data gaps, whereas the scatterplot examples under sparse-coverage conditions were selected from stations with a relatively large number of valid high-missing-rate collocations after applying the MR > 0.5 criterion.

No additional filtering based directly on cloud cover was applied beyond the TROPOMI qa_value filter and the missing-rate criterion described in Sec. 3.1. The missing-rate criterion indirectly accounts for cloud/aerosol-related data loss by quantifying the fraction of missing or invalid TROPOMI grid cells within the 2° collocation window. We have added this clarification to Sec. 3.1 and Sec. 3.1.3 of the revised manuscript.

Comment 24: *Fig. 2. Check spelling (TROPOIMI). The font on the U-Net Prediction Residual part is too small.*

Response 24:

Thank you for your careful review of Figure 2. In the revised manuscript, this figure has been renumbered as Figure 1. We have redrawn the figure to ensure that all labels and text are accurate and clearly readable. The specific revisions are as follows.

Revised figure:

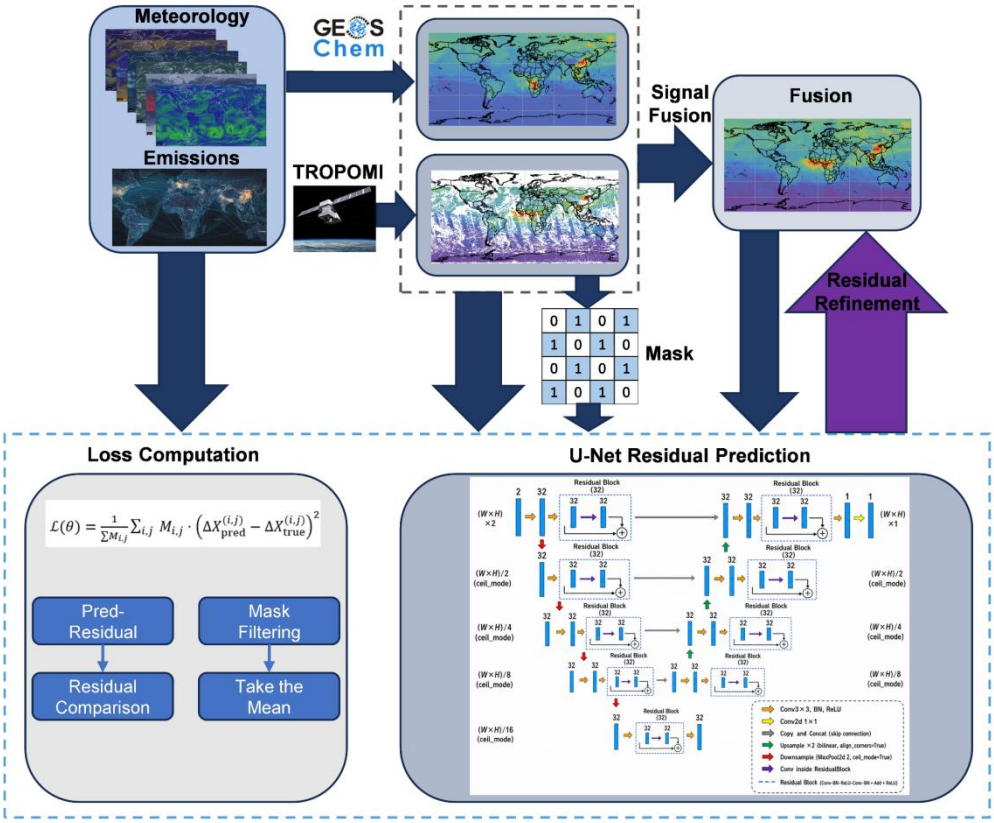


Figure 1. Overview of the proposed fusion framework combining physical modeling, signal-domain reconstruction, and deep learning-based residual correction. The process includes data preprocessing, DCT&SVD-based spatio-temporal fusion, and residual refinement via a lightweight residual U-Net architecture. Arrows indicate the information flow across different modules.

Comment 25: *Lines 214–221 and 223–229. Some information is partially repeated from the previous section (e.g., QV).*

Response 25:

Thank you for pointing out the duplicated content in Lines 214–221 and 223–229. We have removed the repeated description. The corresponding revision appears in Lines 254–261 of the revised manuscript.

Revised text:

Data preprocessing is essential for ensuring the reliability of the fusion results. To reconcile the spatial resolution differences between satellite retrievals and model simulations, TROPOMI XCO and XCH_4 retrievals with qa_value below 0.5 were excluded, and the remaining observations were aggregated onto a global 0.25° × 0.25° grid (720 × 1440) using area-weighted averaging (Wang et al., 2018). GEOS-Chem outputs were subsequently interpolated to the same grid using inverse distance weighting (Setianto and Triandini, 2013). The 0.25° spatial resolution was selected to balance the representation of mesoscale spatial variability with computational efficiency for multi-year global fusion analyses (Hu et al., 2024; Wang et al., 2023).

Comment 26: *Line 228. Mention how sensitive your method is to this choice (e.g., would results change significantly at 0.5° or 0.1°?) or note that you did not explore other resolutions.*

Response 26:

Thank you for this helpful comment. We agree that the choice of target-grid resolution may affect the fusion results. In the revised manuscript, we clarified the rationale for using the 0.25° grid and added a one-year resolution-sensitivity experiment using the 2019 CO data.

The 0.25° spatial resolution was selected as a practical compromise between preserving satellite-observed mesoscale spatial variability and maintaining computational feasibility for multi-year global fusion analyses. A coarser grid, such as 0.5°, would likely smooth local spatial variability, whereas a finer grid, such as 0.1°, would substantially increase computational cost and may increase uncertainty in regions with sparse valid TROPOMI retrievals.

To further assess the influence of resolution choice, we conducted a controlled experiment at the native 2° × 2.5° GEOS-Chem resolution for 2019. The 0.25° fused XCO data were aggregated to the same 2° × 2.5° grid using area-weighted averaging and compared with the fused XCO data generated directly at 2° × 2.5°. The comparison showed high consistency, with $R^2 = 0.926$, RMSE = 6.93 ppb, and mean bias = 0.15 ppb. These results suggest that the selected 0.25° resolution preserves the coarse-scale spatial structure of XCO.

We note that we did not perform additional full-year experiments at 0.5° or 0.1° because of the computational cost of multi-year global reconstruction. Therefore, the sensitivity test mainly evaluates whether the 0.25° product preserves coarse-scale consistency when compared with the native GEOS-Chem resolution. We have added this clarification to the revised manuscript. We have added the relevant description in Lines 259–269 and Lines 820–827 of the revised manuscript.

Revised text:

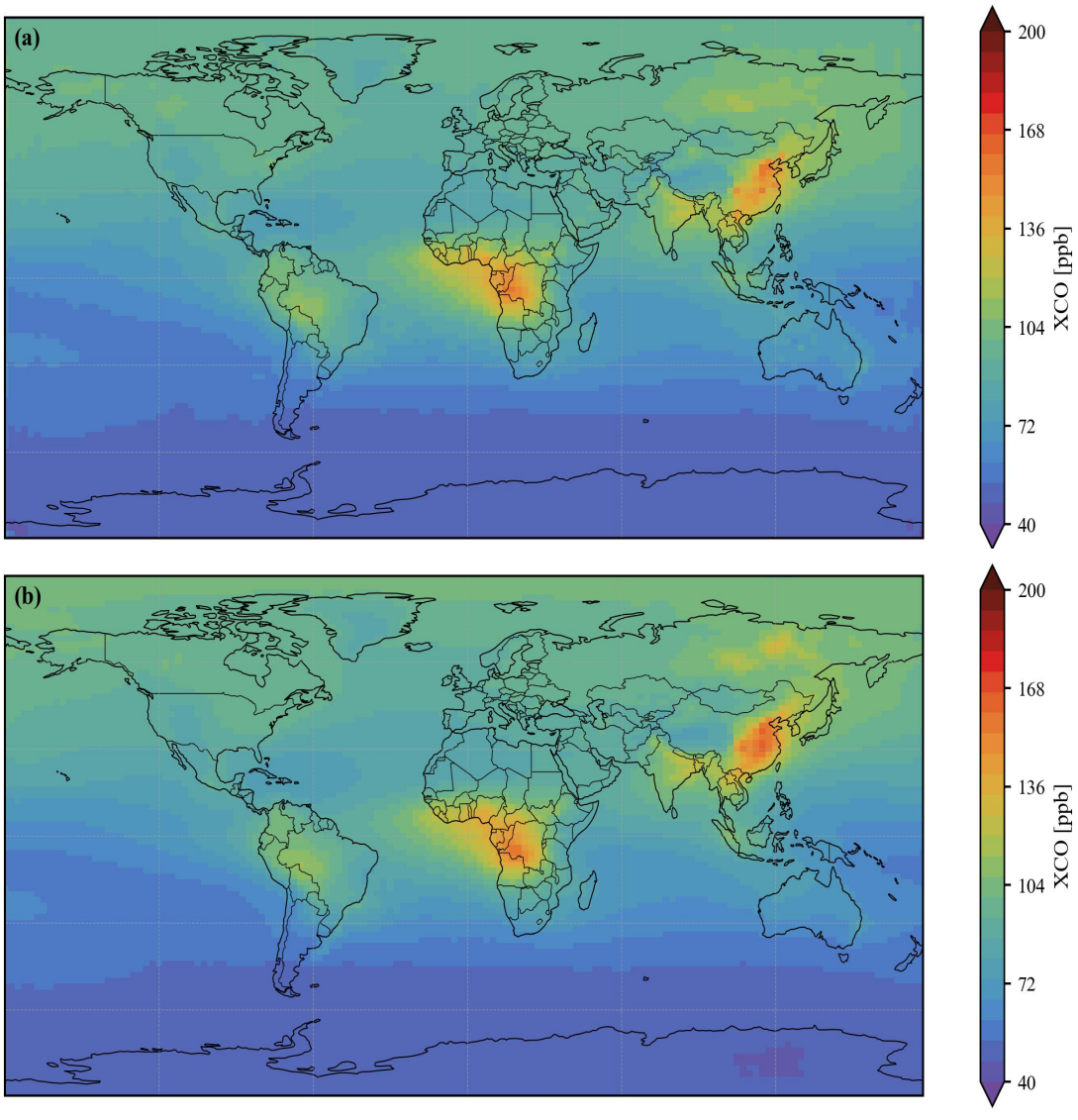
The 0.25° spatial resolution was selected to balance the representation of mesoscale spatial variability with computational efficiency for multi-year global fusion analyses (Hu et al., 2024; Wang et al., 2023).

To evaluate the influence of the resolution mismatch between GEOS-Chem simulation results and the final fused product, we used the 2019 CO data to conduct a one-year resolution sensitivity experiment. In this experiment, we used GEOS-Chem simulation data at the native 2° × 2.5° resolution and aggregated TROPOMI XCO to the 2° × 2.5° resolution using the area-weighted averaging method. The related fusion calculation framework was directly implemented on the 2° × 2.5° grid; subsequently, the obtained fused XCO data at the 2° × 2.5° resolution were compared with the fused product obtained by aggregating the fused XCO data at the 0.25° resolution to the same 2° × 2.5° grid using the area-weighted averaging method.

Figure S12 shows that the fused XCO data generated directly at the native 2° × 2.5° resolution (Fig. S12a) and the 0.25° fused XCO data aggregated to the same 2° × 2.5° grid (Fig. S12b) present relatively similar global spatial distributions, with relatively high XCO over East Asia,

central Africa, and parts of South America. The quantitative comparison results in Figure S13 also show high consistency, with R^2 = 0.926, RMSE = 6.93 ppb, and mean bias = 0.15 ppb. These results indicate that the 0.25° fused XCO data preserve the coarse-scale spatial structure, and that the resolution mismatch does not have a substantial influence on the coarse-scale reconstruction fidelity of XCO.

Figure S12. Spatial comparison of the 2019 XCO resolution-sensitivity experiment. (a) Fused XCO data generated directly at the native 2° × 2.5° GEOS-Chem resolution. (b) Fused XCO data obtained by aggregating the 0.25° resolution fused XCO data to the same 2° × 2.5° grid using the area-weighted averaging method.



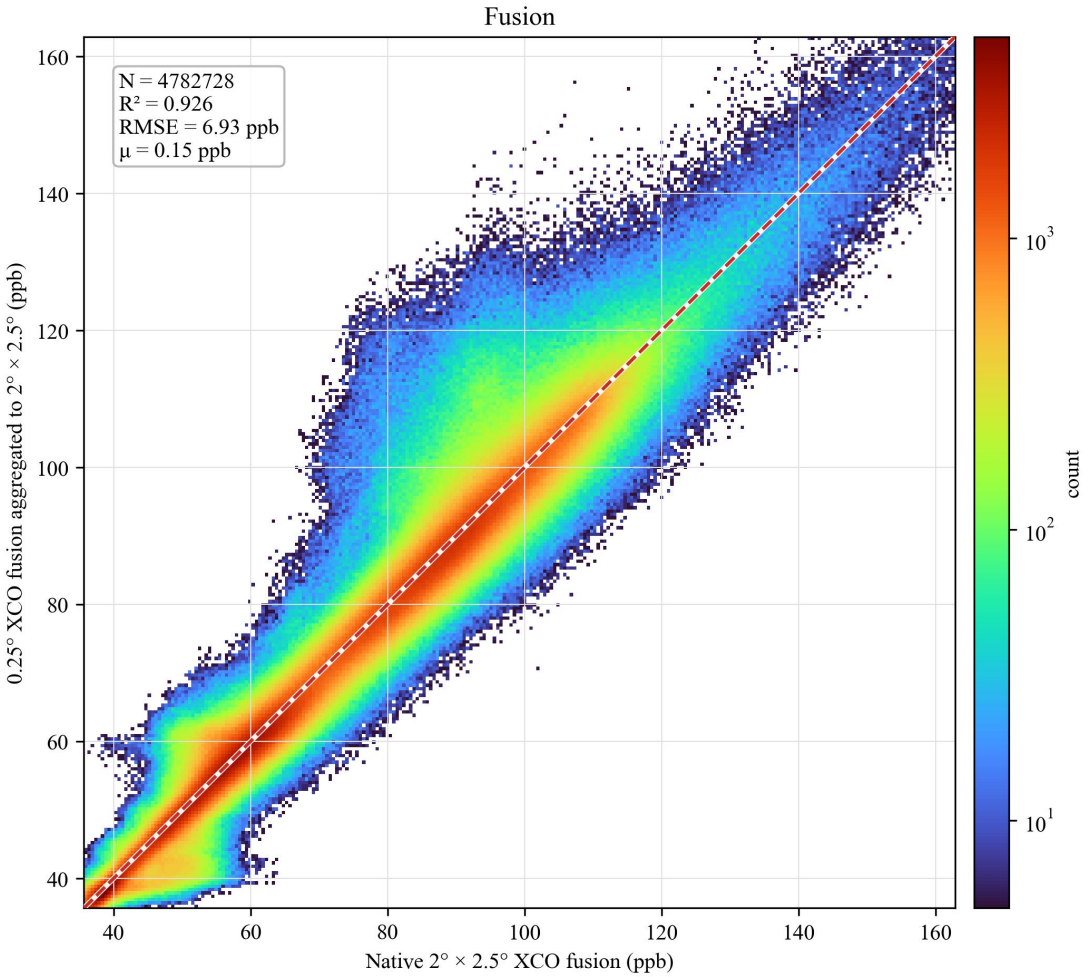


Figure S13. Quantitative comparison of the 2019 XCO resolution-sensitivity experiment.

Comment 27: *line 248. Eq. 2 Meaning of this eq. depends on which type of multiplication the author intends (asterix used).*

Response 27:

Thank you for pointing out this ambiguity. We have revised Equation (2) to clarify the intended mathematical operation. In the revised manuscript, the asterisk symbol has been replaced by the Hadamard product symbol ∘ . We explicitly state that ∘ denotes element-wise multiplication. We also clarified that the spatio-temporal transformation matrix ρ has the same dimensions as XT and XG , and that each element of ρ represents the local relationship between GEOS-Chem and TROPOMI at the corresponding grid cell and time step. The revised Equation (2) is as follows.

Revised equation:

$$XT = XG \circ \rho \qquad (2)$$

Comment 28: *line 257. Those acronyms were explained before.*

Response 28:

We appreciate your correction regarding the inconsistent use of abbreviations. We have carefully reviewed the entire manuscript to ensure that abbreviations already defined earlier in the text are subsequently used only in their abbreviated forms, adhering to the journal's formatting standards and improving the conciseness of the manuscript. The following is the revision to Line 257, which has been incorporated into Line 310–311 of the revised manuscript.

Revised text:

We therefore propose a spatio-temporal 3D matrix smoothing algorithm based on the discrete DCT and SVD to enhance data smoothness and fill missing values in spatio-temporal data.

Comment 29: *line 283. Parameter ε, stated to be "in the middle of the range from 10^{3} to 10^{-1}", which is ambiguous. Please give explicit values used in production and a brief justification (e.g., grid search examples, sensitivity tests, or literature-based choice). Also specify whether ε is the same for XCO and XCH_4 and for global vs. China domains.*

Response 29:

We thank you for the request to clarify the parameter ε. We have revised the manuscript to explicitly state that ε is determined through an automated iterative optimization process rather than being a fixed constant. We employed an optimization function in Python to search for the optimal ε value within the range of 10^{-3} to 10^{-1}. This iterative approach identifies the value that minimizes the reconstruction error by balancing the trade-off between spatial smoothness and the preservation of observational details. We have also clarified that this optimization is performed independently for XCO and XCH_4 to account for their distinct spatial characteristics, ensuring that the resulting fused products are both physically consistent and statistically robust across the global and China domains. We have added the relevant description in Lines 335–338 of the revised manuscript.

Revised text:

In this study, the total number of iterations is empirically set to 100, α is set to 0.75, and ε is selected through an automated search within 10^{-3} to 10^{-1}, using the value that minimizes the reconstruction error on valid observations. This optimization is performed independently for XCO and XCH_4 and for the global and China domains.

Comment 30: *line 285. Eq. 6 and 7 are just placed there without linking them to adjacent text.*

Response 30:

Thank you for pointing this out. We have revised Section 2.2.3 to better connect Eqs. (6) and (7). We have added the relevant description in Lines 339–345 of the revised manuscript.

Revised text:

Equations (6) and (7) define the forward and inverse three-dimensional discrete cosine transforms used in the iterative update in Eq. (4). Specifically, DCT_3 transforms the spatio-temporal parameter matrix ρ from the latitude–longitude–time domain into the frequency domain, where the dominant large-scale and slowly varying structures are represented by low-frequency coefficients. After the smoothing filter Γ^3 in Eq. (5) is applied to these coefficients, $IDCT_3$ transforms the filtered signal back to the original spatio-temporal domain. Therefore, Eqs. (6) and (7) provide the mathematical basis for the frequency-domain smoothing and reconstruction of ρ.

Comment 31: *line 296. "80% singular value energy". Please indicate whether you tested other thresholds (e.g. 90%, 95%) and whether these materially change RMSE/R².*

Response 31:

Thank you for your comment. We agree that the original manuscript did not provide sufficient justification for retaining 80% of the singular-value energy. Therefore, we conducted a threshold sensitivity analysis to evaluate whether the 80% threshold setting has a substantial impact on the reconstruction results. Due to the limited revision time and computational resources, this sensitivity experiment was performed only using the 2019 XCO dataset.

Specifically, daily GEOS-Chem XCO simulations and collocated TROPOMI CO retrievals on the 0.25° × 0.25° grid were used to construct the TROPOMI/GEOS-Chem ratio fields. The same DCT-SVD reconstruction procedure was repeated using cumulative singular-value energy thresholds of 70%, 75%, 80%, 85%, 90%, and 95%, with the 80% case used as the reference. The results are summarized in Table S8 and Figure S2.

The results show that changing the threshold within this range did not materially alter the reconstructed XCO fields. Relative to the 80% case, the spatial correlation coefficients were all higher than 0.9998, the RMSE differences were below 0.62 ppb, the spatial standard deviation ratios remained close to 1.0, and the hotspot overlap ratios were greater than 0.97. The validation statistics against available TROPOMI retrievals also changed only slightly among thresholds. In addition, the cumulative singular-value energy analysis showed that the first, first two, and first three singular values explained approximately 92.0%, 96.6%, and 97.5% of the cumulative energy, respectively, indicating that most of the reconstructed signal is captured by a small number of leading modes. We have added the relevant description in Lines 370–376 and Lines 531–539 of the revised manuscript.

Revised text:

To evaluate whether this empirical threshold materially affects the reconstruction, we conducted an additional threshold-sensitivity experiment for the 2019 daily global XCO reconstruction. Daily GEOS-Chem XCO simulations and collocated TROPOMI CO retrievals on the 0.25° × 0.25° grid were used to construct the TROPOMI/GEOS-Chem ratio fields. The same DCT-SVD reconstruction procedure was repeated using cumulative singular-value energy thresholds of 70%, 75%, 80%, 85%, 90%, and 95%, with the 80% case used as the reference. The results are summarized in Table S8 and Figure S2.

The sensitivity experiment showed that the reconstructed XCO fields were highly stable across the tested SVD energy thresholds. Relative to the 80% reference case, all threshold experiments yielded spatial correlation coefficients higher than 0.9998, RMSE differences below 0.62 ppb, spatial standard deviation ratios close to unity, and hotspot overlap ratios greater than 0.97 (Table S8). The cumulative singular-value energy analysis further showed that the first, first two, and first three singular values explained approximately 92.0%, 96.6%, and 97.5% of the total energy, respectively (Fig. S2). These results indicate that the XCO reconstruction is not sensitive to the exact SVD truncation threshold within the tested range of 70%-95%, supporting the use of 80% as a stable and computationally efficient threshold.

Table S8. Sensitivity of the 2019 daily XCO reconstruction to the SVD cumulative-energy threshold.

SVD energy threshold (%)	Spatial R vs. 80%	RMSE vs. 80% (ppb)	μ vs. 80% (ppb)	SD ratio vs. 80%	Hotspot overlap
70	0.99998	0.206	-0.088	0.99989	0.983
75	0.99999	0.160	-0.086	0.99990	0.984
80	1.00000	0.000	0.000	1.00000	1.000
85	0.99996	0.271	0.000	1.00000	1.000
90	0.99992	0.382	0.003	0.99965	0.999
95	0.99984	0.616	0.285	1.00019	0.975

Note: The 80% threshold was used as the reference case. Spatial R denotes the Pearson correlation coefficient calculated across valid grid cells between each threshold experiment and the 80% case. RMSE, μ, and SD ratio were also calculated relative to the 80% case, where SD denotes the spatial standard deviation. Hotspot overlap was calculated using the upper 10% of reconstructed XCO grid cells. All statistics were derived from the 2019 global daily 0.25° × 0.25° CO reconstruction experiments.

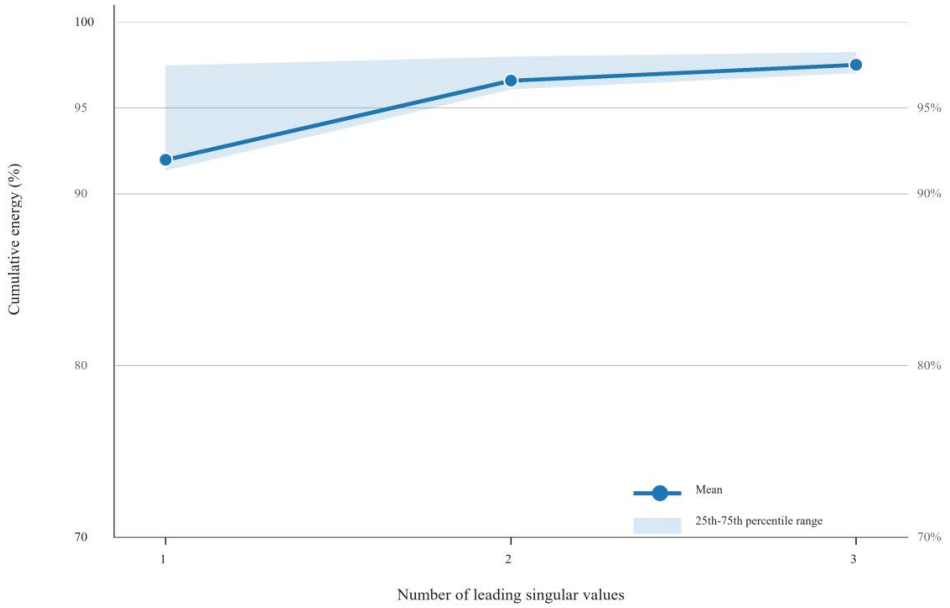


Figure S2. Cumulative singular-value energy explained by the leading singular values in the 2019 daily global XCO reconstruction.

Comment 32: *line 310. "Tables S1 and S2" compare five methods using 2021 data. Please clarify: (i) whether the residual U-Net architecture and hyperparameters were selected based on these 2021 results (risk of tuning to that year); (ii) whether cross-validation across years or sites was performed to avoid overfitting to specific conditions.*

Response 32:

Thank you for your comments regarding the model architecture, hyperparameter selection, and the need to clarify the contribution of each component. We have revised the manuscript by separating the ablation experiment into an independent subsection titled "Ablation Study Design."

In the revised Section 2.2.5, we clarify that the ablation study was primarily conducted using 2021 data because this year provides relatively complete TROPOMI–TCCON collocations for site-based validation. To avoid tuning the framework only to a single year, the final architecture and hyperparameter settings were further checked using multi-year validation samples from the 2019–2023 study period.

We did not perform a formal cross-validation experiment across years or sites. Instead, overfitting risk was assessed by applying the same fixed architecture and hyperparameters to all years and all available TCCON stations during 2019-2023, without site-specific or year-specific retuning. We have added the relevant description in Lines 464–472 of the revised manuscript.

Revised text:

To explicitly isolate the contribution of each component in the proposed two-stage framework, we conducted an ablation study using site-based validation across all TCCON stations for the year 2021. Five configurations were compared: (1) only DCT fusion(Garcia, 2010), (2) DCT and SVD mixed-signal reconstruction(M. Bengherabi et al., 2008; Majhi and Pal, 2021), (3) residual CNN(Y. Jiang et al., 2024), (4) residual U-Net (Ronneberger et al., 2015; Yan et al., 2022), and (5) residual XGBoost(Naseem et al., 2024). The performance was assessed using the coefficient of determination (R^2), root mean square error (RMSE), and mean bias (μ). This design allows the contribution of the DCT/SVD signal-domain reconstruction stage and the residual-correction stage to be evaluated separately. The site-based validation results are summarized in Tables S2 and S3.

Revised tables:

Table S2. Ablation study of site-based XCO validation using different reconstruction configurations. Metrics include coefficient of determination (R^2), root mean square error (RMSE, ppb), and mean bias (μ, ppb).

Configuration	R^2	RMSE	μ
DCT-only	0.8227	10.439	-3.815
DCT/SVD	0.8335	10.1166	-3.7682
DCT/SVD + residual CNN	0.8385	9.9619	-3.7677
DCT/SVD + residual U-Net	**0.845**	**9.7616**	**-3.686**

DCT/SVD + residual XGBoost	0.8392	9.9406	-3.7435

Table S3. Ablation study of site-based XCH_4 validation using different reconstruction configurations. Metrics include coefficient of determination (R^2), root mean square error (RMSE, ppb), and mean bias (μ, ppb).

Configuration	R^2	RMSE	μ
DCT-only	0.7056	17.5156	-6.9345
DCT/SVD	0.732	16.7111	-6.9635
DCT/SVD + residual CNN	0.7469	16.2404	-6.9162
DCT/SVD + residual U-Net	**0.7598**	**15.8212**	**-6.7858**
DCT/SVD + residual XGBoost	0.7476	16.2188	-6.9563

Comment 33: *Eq.12. Masked MSE only over M = 1 pixels. Comment briefly on the risk of biased training if M = 1 is spatially clustered (e.g. cloud-free regions).*

Response 33:

Thank you for this important comment. We agree that computing the masked MSE only over valid TROPOMI pixels may introduce sampling bias if these valid pixels are spatially or temporally clustered, for example in cloud-free regions. We have revised the Loss Function Design subsection to explicitly acknowledge this limitation. We have added the relevant description in Lines 425–432 of the revised manuscript.

Revised text:

This masked mean squared error ensures that the learning focuses on valid TROPOMI retrievals. However, because valid retrievals may be spatially or temporally clustered in cloud-free regions, the masked loss may introduce sampling bias and affect extrapolation in persistently cloudy or aerosol-affected areas where valid observations are sparse. This risk is partly mitigated by the use of spatially complete predictors, including GEOS-Chem simulations and the preliminary DCT/SVD reconstruction, which provide physical and spatiotemporal constraints across the full domain. Nevertheless, fused data in regions with persistent data gaps should be interpreted with caution.

Comment 34: *line 345. lightweight residual U-Net architecture. Information on hyperparameters such as levels, filters, dropout rate, epochs, etc., should be added. Do you train separate networks for XCO and XCH_4 or a joint multi-output model?*

Response 34:

Thank you for pointing out that the model architecture and training hyperparameters were not sufficiently described. We have revised the "Model Architecture and Training" subsection to provide these details explicitly. We have also clarified that we train separate, specialized networks for XCO and XCH_4 to account for their distinct atmospheric lifetimes, transport patterns, and spatial characteristics, rather than using a joint multi-output model. We have added the relevant

description in Lines 440–448 of the revised manuscript.

Revised text:

The residual U-Net consists of an encoder–decoder structure with skip connections and lightweight residual blocks. The network contains 10 convolutional layers in total, including the convolutional layers inside the residual blocks. All hidden convolutional layers use 32 filters, with two input channels and one output channel for the predicted residual field. The input and output spatial dimensions remain unchanged (i.e., 720×1440 at 0.25° resolution), enabling pixel-wise learning of spatial residuals. The model was trained using daily global maps from 2019 to 2023. Training was performed for 5 epochs with a batch size of 4. We used the Adam optimizer with a fixed learning rate of 1×10^{-4}, and no learning-rate scheduler was applied. No dropout layer or data augmentation was used in the final configuration.

Comment 35: *line 355. Missing details on optimizer, learning rate schedule, batch size, number of epochs, train/validation split, and whether any data augmentation was used.*

Response 35:

Thank you for pointing out that the training settings were insufficiently described. We have revised the "Model Architecture and Training" subsection to explicitly report the missing hyperparameters. In the revised manuscript, we now state that the residual U-Net was trained using the Adam optimizer with a fixed learning rate of 1×10^{-4} . No learning-rate scheduler was applied. The batch size was set to 4, and the model was trained for 5 epochs. No dropout layer or data augmentation was used in the final configuration.

Regarding the data partitioning, we have clarified that our framework employs a masked-learning strategy. The model is trained to reconstruct residuals only in regions where satellite observations are available, using the observation mask M = 1. Because the objective is to learn spatiotemporal residual structures for gap filling over continuous global fields, the full available training period was used to preserve geographic and temporal continuity rather than applying a random pixel-level split. No separate random train/validation split was used for model selection. Model performance was instead evaluated against valid TROPOMI observations and through the comparison experiments reported in the validation section. We have added the relevant description in Lines 445–448 of the revised manuscript.

Revised text:

The model was trained using daily global maps from 2019 to 2023. Training was performed for 5 epochs with a batch size of 4. We used the Adam optimizer with a fixed learning rate of 1×10^{-4}, and no learning-rate scheduler was applied. No dropout layer or data augmentation was used in the final configuration.

Comment 36: *line 359. (γ) value? Can his truncation bias extremes?*

Response 36:

Thank you for this comment. We have revised the Model Architecture and Training subsection to explicitly specify the value of γ, which was set to 3.

We agree that residual truncation may partly dampen extremely large residual corrections and could therefore reduce the magnitude of rare local extremes. We have added this limitation to the revised manuscript. We also clarified that the truncation is applied only to the neural-network-predicted residual correction, rather than directly to the final XCO or XCH_4 concentration fields. We have added the relevant description in Lines 451–462 of the revised manuscript.

Revised text:

$$|\Delta X_{pred}| < \gamma \cdot \sigma_{TROPOMI}, \gamma = 3 \quad (13)$$

where $\sigma_{TROPOMI}$ denotes the standard deviation of observed valid values, and γ is a hyperparameter controlling confidence bounds. This truncation is applied only to the neural-network-predicted residual correction, not directly to the final XCO or XCH_4 concentration fields. Its purpose is to prevent unphysical residual corrections arising from unstable extrapolation in regions with sparse valid TROPOMI observations. We acknowledge that such truncation may partly dampen extremely large residual corrections and could therefore reduce the magnitude of rare local extremes. However, because the truncation is imposed on the residual term rather than on the full concentration fields, the large-scale and physically consistent extreme signals already present in the GEOS-Chem simulations, TROPOMI observations, or DCT/SVD reconstruction are retained. Thus, the constraint mainly suppresses anomalous correction spikes while preserving the dominant extreme-event signals.

Comment 37: *lines 382–383. "Only those cases in which the satellite data exhibit a missing rate exceeding 0.5 at the site are retained for comparison", while lines 417–419 mention stations with "highest availability ratios" for a different selection. Please clarify.*

Response 37:

We thank you for pointing out this inconsistency. In the original manuscript, the description of the site-selection criterion was not sufficiently clear and may have caused misunderstanding. The missing-rate threshold was used to select site-day collocations under sparse TROPOMI coverage conditions. In contrast, the representative sites used for scatterplot visualization were selected based on the number of valid collocated samples after applying this threshold. Therefore, the original description referring to stations with the “highest availability ratios” has been revised to indicate that the selected sites were those with a relatively large number of valid collocated samples after applying the MR > 0.5 criterion, which ensures sufficient samples for visualization. We have revised the text accordingly. We have added the relevant description in Lines 583–588 of the revised manuscript.

Revised text:

This selection was based on the number of valid collocated samples after applying the missing-rate criterion, rather than on validation performance. Specifically, for each TCCON

station, we first applied the qa_value > 0.5 quality filter, the 13:30 ± 1 h temporal window, the 2° spatial collocation radius, and the site-day missing-rate criterion (MR > 0.5). Stations with relatively sufficient high-missing-rate collocations were then selected as representative examples for visualization.

Comment 38: *line 370. How are p-values used here (for linear regression slope significance, correlation significance?). What hypothesis tests did you run?*

Response 38:

Thank you for pointing out this ambiguity. We agree that the original statement was inaccurate. No formal hypothesis tests were performed for the validation metrics reported in this study. The metrics R^2, RMSE, mean bias (μ), and standard deviation of the bias (σ) were used as descriptive validation statistics, and no p-value threshold was applied to their calculation. We have therefore removed the statement referring to $p<0.01$ from Section 2.2.6 and deleted the corresponding citation. The revised text now explicitly states that no formal hypothesis tests or p-value thresholds were applied to these validation metrics. We have added the relevant description in Lines 479–481 of the revised manuscript.

Revised text:

These metrics are used as descriptive validation statistics, and no formal hypothesis tests or p-value thresholds are applied.

Comment 39: *line 408. The statement "The average bias of the GEOS-Chem simulation data was corrected using a manual correction method…" is vague. Please provide a detailed description.*

Response 39:

We thank you for pointing out this ambiguity in our description. We clarified the treatment of GEOS-Chem in the validation. The raw, uncorrected GEOS-Chem simulations were used as the physical prior in the fusion and reconstruction framework. The diagnostic mean-bias adjustment was applied only for the GEOS-Chem–TCCON comparison and was not used to generate the fused datasets. This adjustment was site-specific: for each TCCON station and each gas, the mean difference between GEOS-Chem simulations and TCCON observations was calculated over all valid collocated days, and this station-specific mean bias was then removed from the GEOS-Chem time series at that station. This procedure was performed after the TCCON collocation step and only for diagnostic comparison, so it did not affect the fused products. All figures and tables involving this bias-adjusted GEOS-Chem diagnostic comparison are explicitly labeled to indicate that the GEOS-Chem results were bias-corrected before validation. We have added the relevant description in Lines 570–580 of the revised manuscript.

Revised text:

GEOS-Chem simulations show a systematic low bias relative to TCCON observations. The raw, uncorrected GEOS-Chem simulations are used as the physical prior in the fusion and

reconstruction framework. The mean-bias adjustment is applied only for the diagnostic GEOS-Chem–TCCON comparison and is not used in generating the fused datasets. For each TCCON station and each gas, we calculate the mean difference between GEOS-Chem simulations and TCCON observations over all valid collocated days, and then subtract this mean bias from the GEOS-Chem time series at that station. This procedure removes the station-mean offset while preserving the model's day-to-day variability, allowing a clearer assessment of whether GEOS-Chem captures the temporal variations observed by TCCON. All figures and tables involving this bias-adjusted GEOS-Chem diagnostic comparison are explicitly labeled to indicate that the GEOS-Chem results were bias-corrected before validation.

Comment 40: *line 411. What fraction of the TCCON dataset falls into this category?*

Response 40:

We thank you for this valuable suggestion. Based on the collocated validation dataset used in this analysis, approximately 20.7% of the XCO samples and 77.6% of the XCH_4 samples fall into the high-missing-rate category (MR > 0.5).

Comment 41*: line 436. Please indicate explicitly which metrics this refers to (R^2, RMSE, μ, σ).*

Response 41:

Thank you for pointing out this ambiguity. We have revised this statement to clearly specify the metrics used and recalculated the site-level statistics. The statement refers to sites where the fused data improve at least one of the three reported site-level metrics in Tables S4 and S5, namely R^2, RMSE, and μ, relative to the original TROPOMI retrievals. Here, improvement is defined as higher R^2, lower RMSE, or smaller absolute mean bias |μ|. Based on this definition, the fused data improve at least one metric at 14 of 18 XCO sites (77.8%) and 10 of 17 XCH_4 sites (58.8%). We have added the relevant description in Lines 602-605 of the revised manuscript.

Revised text:

Overall, as shown in Tables S4 and S5, the fused data improve at least one of the three reported site-level metrics relative to the original TROPOMI retrievals at 14 of 18 XCO sites (77.8%) and 10 of 17 XCH_4 sites (58.8%). Here, improvement refers to higher R^2, lower RMSE, or smaller |μ|.

Table S4. (GEOS) Single-site validation results for GEOS-Chem, TROPOMI, and fused XCO data, with RMSE & μ in parts per billion (ppb).

SITE NAME	R^2			RMSE			μ		
	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION
burgos01	0.237	0.919	0.935	16.909	5.769	5.246	9.667	-2.799	-2.425

darwin01	-0.808	0.370	0.646	13.896	8.546	6.330	-9.241	-4.196	-2.848
easttroutlake01	-0.487	0.827	0.903	19.017	10.070	7.372	4.540	0.688	1.092
garmisch01	0.664	0.795	0.806	5.903	3.897	4.444	-1.453	0.737	1.113
hefei01	0.583	0.930	0.980	16.501	7.362	3.421	-15.897	-1.450	1.444
izana01	0.093	0.185	0.513	6.943	7.120	5.482	-3.692	5.790	3.909
karlsruhe01	0.521	0.411	0.596	7.055	6.892	6.441	4.002	-3.137	-4.337
lamont01	-0.134	0.534	0.580	9.297	6.447	6.013	5.209	-4.803	-3.650
lauder03	-0.524	0.849	0.663	11.990	3.893	5.666	-8.828	-2.228	-2.424
nicosia01	-0.251	0.868	0.847	8.114	2.791	3.008	5.978	0.557	-1.147
orleans01	0.581	0.791	0.537	6.704	6.656	8.947	3.012	-2.665	-2.786
parkfalls01	-0.468	0.899	0.811	19.445	6.597	8.677	7.525	-2.116	-2.106
reunion01	0.109	0.893	0.896	8.529	3.113	3.061	2.327	0.448	-0.859
rikubetsu01	0.635	0.961	0.899	7.674	2.893	4.545	1.122	0.064	0.438
sodankyla01	0.388	0.267	0.917	6.255	6.203	2.230	-0.775	2.807	1.381
tsukuba02	0.955	0.825	0.958	4.446	9.713	4.031	-2.817	-6.798	-1.684
wollongong01	-0.319	0.776	0.745	14.709	6.562	6.969	-5.625	-4.501	-4.418
xianghe01	0.279	0.269	0.553	11.965	10.749	9.469	-2.922	6.306	2.628

Table S5. (GEOS) Single-site validation results for GEOS-Chem, TROPOMI, and fused XCH_4 data, with RMSE & μ in parts per billion (ppb).

SITE NAME	R^2			RMSE			μ		
	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION
bremen01	-0.140	0.134	0.631	13.602	10.634	8.139	-0.024	-6.005	-2.375
burgos01	-0.280	0.741	0.877	24.716	13.217	8.892	-11.397	-7.712	-5.010
easttroutlake01	-0.560	0.371	0.322	27.858	18.152	20.471	-7.800	-10.457	-12.565
edwards01	-1.666	0.938	0.908	33.363	5.655	6.750	-20.863	2.314	4.679
garmisch01	-0.539	0.693	0.831	23.574	11.552	8.763	-4.177	-4.162	-3.266
hefei01	-1.130	0.882	0.770	38.129	8.789	12.575	-26.749	-3.298	-5.167
karlsruhe01	-1.466	0.546	0.561	31.590	13.937	14.156	-17.057	-8.526	-11.191
lamont01	-2.086	0.838	0.830	34.635	9.308	9.313	-22.403	-3.808	-4.785
lauder03	-1.178	0.725	0.742	33.162	12.366	12.116	-16.148	-7.392	-9.347
nicosia01	-2.566	0.670	0.893	32.223	10.568	5.954	-23.891	2.413	-1.740
orleans01	-1.780	0.551	0.705	30.640	13.366	11.070	-16.560	-6.472	-6.654
parkfalls01	-0.923	0.137	0.221	29.718	21.329	20.668	-12.079	-14.704	-16.099
pasadena01	-3.177	0.955	0.910	41.561	4.659	6.603	-32.970	-0.942	-3.476
rikubetsu01	0.230	0.740	0.715	15.539	9.189	9.892	6.588	-2.146	-4.784
tsukuba02	-0.582	0.777	0.712	15.420	6.238	7.147	-1.677	-2.308	-3.793
wollongong01	-1.396	0.195	0.347	29.981	19.964	17.320	-14.676	-12.975	-13.672
xianghe01	0.033	0.855	0.888	25.528	10.349	9.120	-9.444	-0.354	-1.247

Comment 42: *Fig.8. Which % of grids are pure ML reconstruction? A map of the coverage ratio would be useful.*

Response 42:

We thank you for the comment regarding the identification of reconstructed grid cells. In the revised manuscript and the accompanying dataset, we have addressed this issue by providing two diagnostic variables: obs_mask and observation_fraction. The obs_mask variable indicates whether the value at each grid cell is a TROPOMI observation or reconstructed by the fusion framework. Specifically, obs_mask = 1 indicates that the value at that grid cell and day is a TROPOMI observation, whereas obs_mask = 0 indicates that the value was generated by the proposed reconstruction method. Therefore, the percentage of reconstructed grid cells for each daily map can be calculated directly as the percentage of grid cells with obs_mask = 0.

In addition, the observation_fraction variable provides a quantitative measure of observational coverage at the monthly scale. It is calculated as the fraction of days in a month for which valid TROPOMI observations are available at each grid cell. Accordingly, 1 − observation_fraction represents the fraction of days for which the grid-cell values are gap-filled by the fusion framework rather than directly constrained by valid TROPOMI observations. This variable allows users to quantify the spatial and temporal extent of reconstructed values and to generate coverage or reconstruction-ratio maps if needed.

We did not add a separate coverage-ratio map to the main text because this information is now provided quantitatively through observation_fraction in the released dataset, and users can readily generate coverage or reconstruction-ratio maps for any month or region of interest. We have added the relevant description in Lines 865 – 877 of the revised manuscript.

Revised text:

The main data variables are `XCO` and `XCH4`, representing the fused column-averaged dry-air mole fractions of CO and CH_4, respectively. In addition, an obs_mask variable is provided to indicate whether the value at each grid cell is a TROPOMI observation or reconstructed by the fusion framework. The obs_mask variable is a dimensionless binary mask, where obs_mask = 1 indicates that the value at that grid cell and day is a TROPOMI observation, while obs_mask = 0 indicates that the value was generated by the proposed reconstruction method.

An `observation_fraction` variable is also included in each monthly file. This variable is calculated as the mean of `obs_mask` over the `time` dimension and represents the fraction of days in that month for which valid TROPOMI observations are available at each grid cell. For example, `observation_fraction = 1` indicates that valid satellite observations are available for all days in the month at that grid cell, whereas `observation_fraction = 0` indicates that all daily values at that grid cell are reconstructed by the fusion framework. This information is provided to prevent users from interpreting the entire gap-free product as purely observation-based.

Comment 43: *line 501-505. The methodology for trend computation (regression model, handling of autocorrelation, treatment of missing values, and significance level) is not detailed. Also - short period.*

Response 43:

Thank you for this important comment. We agree that the trend-computation procedure and the limitation of the short study period should be clarified. We have revised the manuscript to specify

that annual mean XCO and XCH_4 values were first calculated from the daily fused datasets. The annual rate of change was then estimated using ordinary least-squares linear regression fitted to the five annual mean values from 2019 to 2023. Because the fused datasets are gap-free, no additional missing-value interpolation was applied in the trend calculation.

We also clarified that no formal significance testing, p-value threshold, or autocorrelation correction was applied, because only five annual values were available, which is insufficient for a robust autocorrelation-adjusted trend test. Accordingly, we have revised the wording throughout this section to describe the results as "short-term linear changes" or "tendencies" over 2019-2023 rather than statistically robust long-term trends. We have added the relevant description in Lines 719-724 of the revised manuscript.

Revised text:

Figures S8 and S9 illustrate the spatial distribution of the annual rates of change in global XCO and XCH_4 from 2019 to 2023. For each grid cell, the rate of change is estimated by applying ordinary least-squares linear regression to the five annual mean values. Based on this analysis, XCO shows an increasing tendency over North America, whereas Central Africa and Eastern China show decreasing tendencies during 2019–2023 (Fig. S8). For XCH_4, the global annual mean increases over the same period, with a larger positive rate of change observed over Central Africa (Fig. S9).

Comment 44: *line 518. The reported mean increases (17.1 ppb XCO, 24.5 ppb XCH_4) depend critically on the definition of "fire-affected area", the temporal window considered as fire period vs. reference, and whether the baseline is climatological or just adjacent days. Please specify.*

Response 44:

Thank you for this helpful comment. We agree that the definitions of the fire-affected area, fire period, and reference baseline should be specified. We have revised Section 3.3 to clarify the calculation of the reported mean concentration enhancements during the Chongqing hill fire. We have added the relevant description in Lines 731-744 of the revised manuscript.

Revised text:

To quantify the concentration enhancements, the fire-affected area is defined as a localized domain of 29°–30.5°N and 105.5°–107.5°E, based on the spatial extent of thermal anomalies and smoke plumes observed during the peak of the event. The active fire period is defined as 21–31 August 2022, while 10–20 August 2022 is used as the reference period. The reference period is selected as an adjacent pre-fire baseline, rather than a climatological baseline, to reduce the influence of broader seasonal variability and to better isolate short-term concentration changes associated with the fire episode. Mean XCO and XCH_4 values are calculated by averaging all grid cells within the defined fire-affected domain for each day. The reported enhancements are then calculated as the difference between the mean concentrations during the active fire period and those during the reference period. Based on this definition, the fused data show average increases of 17.1 ppb for XCO and 24.5 ppb for XCH_4 during the fire period. The comparison also indicates that the fused datasets preserve local spatial features observed by TROPOMI while providing

more complete spatial coverage over the fire-affected region.

Comment 45: *line 524. How is the annual growth rate computed (linear fit to daily values?), and over what exact years? Similarly, for XCO decline rates (line 536), clarify whether the % differences are relative or absolute, and provide confidence intervals or at least an idea of uncertainty.*

Response 45:

Thank you for your constructive comment. Following your suggestion in Comment 8, we improved the quantitative analysis of short-term temporal variations in XCO and XCH_4 over the major rice-growing regions in northeastern and southeastern China in the revised manuscript. Specifically, to reduce the influence of seasonal cycles, harmonic regression was applied to the daily regional mean XCO and XCH_4 time series for 2019-2023. The model includes a linear term, an annual periodic term, and a semiannual periodic term. The coefficient of the linear term represents the short-term rate of change after removing the seasonal cycle, and the 95% confidence interval was calculated from the regression standard error.

For the XCO decline rates mentioned in the original manuscript, we have clarified the expression by reporting absolute rates of change in ppb/year rather than percentage differences. The corresponding rates and 95% confidence intervals are now provided in Tables S9 and S10. We have added the relevant description in Lines 749-753 of the revised manuscript.

Revised text:

Figures 8 and 9 show the temporal variations in XCO and XCH_4 derived from TROPOMI and the fused data over these regions, respectively. To reduce the influence of the seasonal cycle, the short-term rates of change were derived using harmonic regression with annual and semi-annual terms, and the corresponding 95% confidence intervals are provided in Tables S9 and S10.

Table S9. Seasonally adjusted short-term rates of change and 95% confidence intervals for XCO over rice-growing regions during 2019–2023. Unit: ppb/year.

Region	Rate of change		95% confidence interval	
	TROPOMI	Fusion	TROPOMI	Fusion
Northeast China	−2.36	−2.28	−3.30 to −1.43	−3.30 to −1.27
Southeast China	−1.29	−1.76	−2.15 to −0.43	−2.79 to −0.72

Table S10. Seasonally adjusted short-term rates of change and 95% confidence intervals for XCH_4 over rice-growing regions during 2019–2023. Unit: ppb/year.

Region	Rate of change		95% confidence interval	
	TROPOMI	Fusion	TROPOMI	Fusion
Northeast China	12.08	12.16	11.24 to 12.92	11.27 to 13.04
Southeast China	12.16	11.78	10.83 to 13.49	9.72 to 13.83

Comment 46: *line 559. Add evidence.*

Response 46:

Thank you for your valuable comment. We agree that the original conclusion was too general and needed to provide clear supporting evidence. In the revised manuscript, we have expanded the conclusion section by adding quantitative validation results from the representative TCCON time-series comparison analysis and the comprehensive validation using all qualified TCCON station-day collocation data. We have also added validation of the performance of our fused data under different TROPOMI coverage conditions. We have added the relevant description in Lines 901-913 of the revised manuscript.

Revised text:

Validation against independent TCCON observations shows that the fused datasets improve spatiotemporal continuity while still maintaining good consistency with TCCON data. In the time-series comparison analysis at representative sites, the fused XCO and XCH_4 products generally conform to the temporal variation patterns observed by TCCON and TROPOMI, while reducing the systematic underestimation phenomenon associated with GEOS-Chem simulations. To further evaluate the performance under different TROPOMI data coverage rates, we conducted experiments based on missing-rate thresholds, with three MR intervals: $MR < 0.3$, $0.3 \leq MR \leq 0.5$, and $MR > 0.5$. The results show that when $MR < 0.3$ and $0.3 \leq MR \leq 0.5$, the fused products perform comparably to the original TROPOMI retrievals; whereas under sparse TROPOMI data coverage conditions ($MR > 0.5$), the fused XCO product obtained a higher R^2 value and a lower RMSE value than the TROPOMI product, with $R^2 = 0.91$ and RMSE = 5.65 ppb. The fused XCH_4 product obtained a higher R^2 value and a lower RMSE value, with $R^2 = 0.83$ and RMSE = 15.40 ppb. Compared with GEOS-Chem simulations, the fused products significantly reduced systematic bias.

Comment 47: *line 569. "excellent instrument for monitoring severe weather and pollution occurrences". More neutral language would be better.*

Response 47:

We thank you for the suggestion to use more neutral language. We have revised the sentence at the original Line 569 to replace "excellent instrument" with a more objective description. We have added the relevant description in Lines 935-937 of the revised manuscript.

Revised text:

Overall, this study provides continuous, high-resolution, and validated datasets of XCO and XCH_4 that can support future research on atmospheric composition, air quality, emission monitoring, and climate-related applications.

Comment 48: *Lack of the limitations section, please highlight user-relevant aspects:*

-Potential biases in regions without TCCON or other ground validation.

-Dependence on GEOS-Chem priors in persistent-cloud regions.

-Known artefacts or discontinuities (e.g., around June 2019 TROPOMI pixel size change).

-Discuss ML "hallucination" risk. Residual U-Net loss masked, model learns: what usually happens here, not what happened that day.

Response 48:

Thank you for this important suggestion. We agree that the limitations of the fused datasets should be explicitly discussed from the perspective of potential users. In the revised manuscript, we have added a dedicated section entitled “Discussion of uncertainties and limitations” to clarify the main user-relevant limitations of the dataset.

Specifically, we now discuss the possible bias and representativeness uncertainties in regions lacking independent ground-based validation, especially where TCCON coverage is sparse or absent. We also clarify that, in regions or periods with persistently sparse valid TROPOMI retrievals, such as cloudy, high-latitude, or aerosol-affected regions, the reconstruction relies more strongly on the GEOS-Chem prior and the DCT/SVD spatiotemporal structure. Therefore, systematic biases in GEOS-Chem may propagate into the fused fields.

In addition, we added a discussion of possible artefacts or discontinuities related to changes in the input satellite product, particularly the TROPOMI pixel-size change around June 2019 from 7.0 km × 7.0 km to 7.0 km × 5.5 km. Although gridding, quality filtering, and residual correction can mitigate this effect, subtle local-scale or short-term discontinuities may still remain.

We further addressed the uncertainty associated with the masked residual-learning strategy. Because the residual U-Net is trained only at valid TROPOMI pixels, it learns the statistical relationship between the preliminary reconstruction error and the input predictors under observed conditions. In persistently missing regions, the model therefore infers the most likely residual pattern under similar spatiotemporal conditions rather than directly observing the exact atmospheric state on a given day.

Finally, to help users assess these limitations, we emphasize the use of the provided observation-support variables, including obs_mask, MR, and observation_fraction, when interpreting the fused datasets, especially for local anomaly detection, persistently missing regions, or areas without independent validation data.

The revised text has been added to Lines 778–851 of the revised manuscript.

Revised text:

Although the proposed fusion framework substantially improves the spatial continuity of TROPOMI XCO and XCH_4 products and provides daily gap-free fields, several limitations should be considered when using the fused datasets.

Figures 10–11 show the temporal and latitudinal distributions of daily biases between GEOS, Fused, TROPOMI, and TCCON. Overall, GEOS exhibits a clear systematic underestimation of both XCO and XCH_4 relative to TCCON, especially after 2020, when strong negative biases appear at many sites. After fused, the biases are substantially reduced, with colors generally shifting from dark blue to light blue or near white. This indicates that the fusion method effectively corrects the systematic underestimation in GEOS and improves its consistency with TCCON. Compared with GEOS, TROPOMI shows smaller overall biases, but it has more missing

data and still exhibits localized positive or negative biases at some sites and periods. Overall, the fusion product performs better than GEOS in reducing biases and improving reliability, although residual uncertainties remain at individual sites.

AKs are an important source of uncertainty in comparisons between satellite retrievals and model simulations because they describe the vertical sensitivity of TROPOMI retrievals. To evaluate this effect, we conducted an AK sensitivity test using the 2019 CO data by comparing GEOS-Chem XCO with TROPOMI XCO before and after applying the TROPOMI CO column averaging kernels. As shown in Table S11, applying the TROPOMI CO column averaging kernels slightly increased the R^2 from 0.543 to 0.565, while the RMSE increased from 28.17 ppb to 28.66 ppb and the mean bias changed from −17.09 ppb to −18.34 ppb. These results indicate that AK treatment has a limited influence on the column-concentration-based fusion results in this experiment.

Another concern relates to the dependence of the fused datasets on the accuracy of the GEOS-Chem prior and the representativeness of valid TROPOMI observations. GEOS-Chem provides physically consistent large-scale background information, but it carries uncertainties from emission inventories, transport, chemistry, boundary conditions, vertical mixing, and other model parameterizations. In regions or periods where valid TROPOMI retrievals are persistently sparse—such as the Intertropical Convergence Zone, high-latitude zones, or heavily polluted areas with dense cloud or aerosol cover—the reconstruction relies more strongly on the GEOS-Chem prior and the DCT/SVD spatiotemporal structure. Consequently, systematic biases in the prior model may propagate into the fused fields. Users are therefore advised to consult the obs_mask variable provided in the dataset to assess observational support and to interpret the reconstructed values cautiously in persistently missing regions.

Figure S12 shows that the fused XCO data generated directly at the native 2° × 2.5° resolution (Fig. S12a) and the 0.25° fused XCO data aggregated to the same 2° × 2.5° grid (Fig. S12b) present relatively similar global spatial distributions, with relatively high XCO over East Asia, central Africa, and parts of South America. The quantitative comparison results in Figure S13 also show high consistency, with R^2 = 0.926, RMSE = 6.93 ppb, and mean bias = 0.15 ppb. These results indicate that the 0.25° fused XCO data preserve the coarse-scale spatial structure, and that the resolution mismatch does not have a substantial influence on the coarse-scale reconstruction fidelity of XCO.

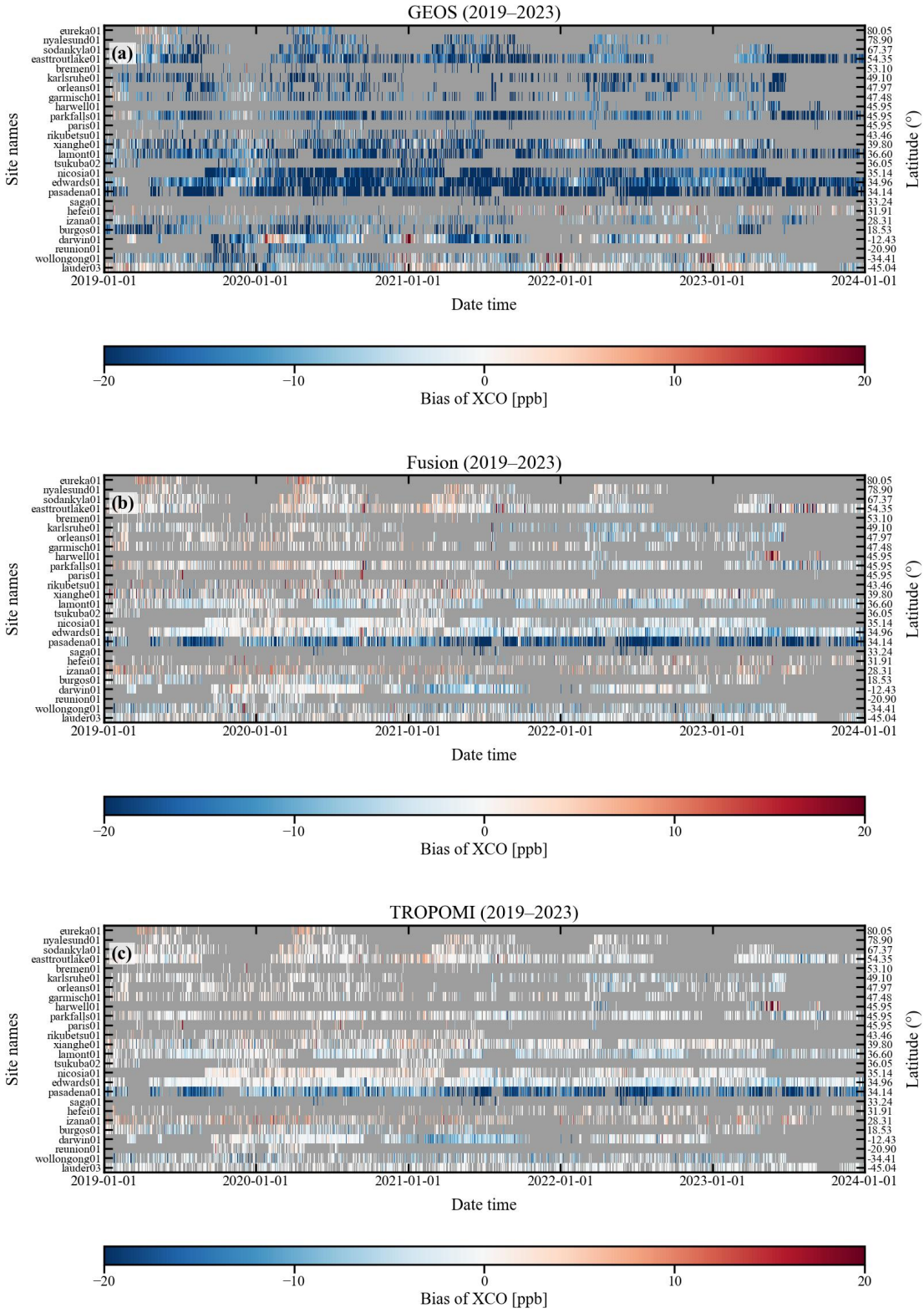


Figure 10. Heat maps of the biases between daily (a) GEOS, (b) fused, and (c) TROPOMI XCO and TCCON XCO over time and latitude. Color ramps stand for the biases of XCO. Background colors (grey) indicate the missing data.

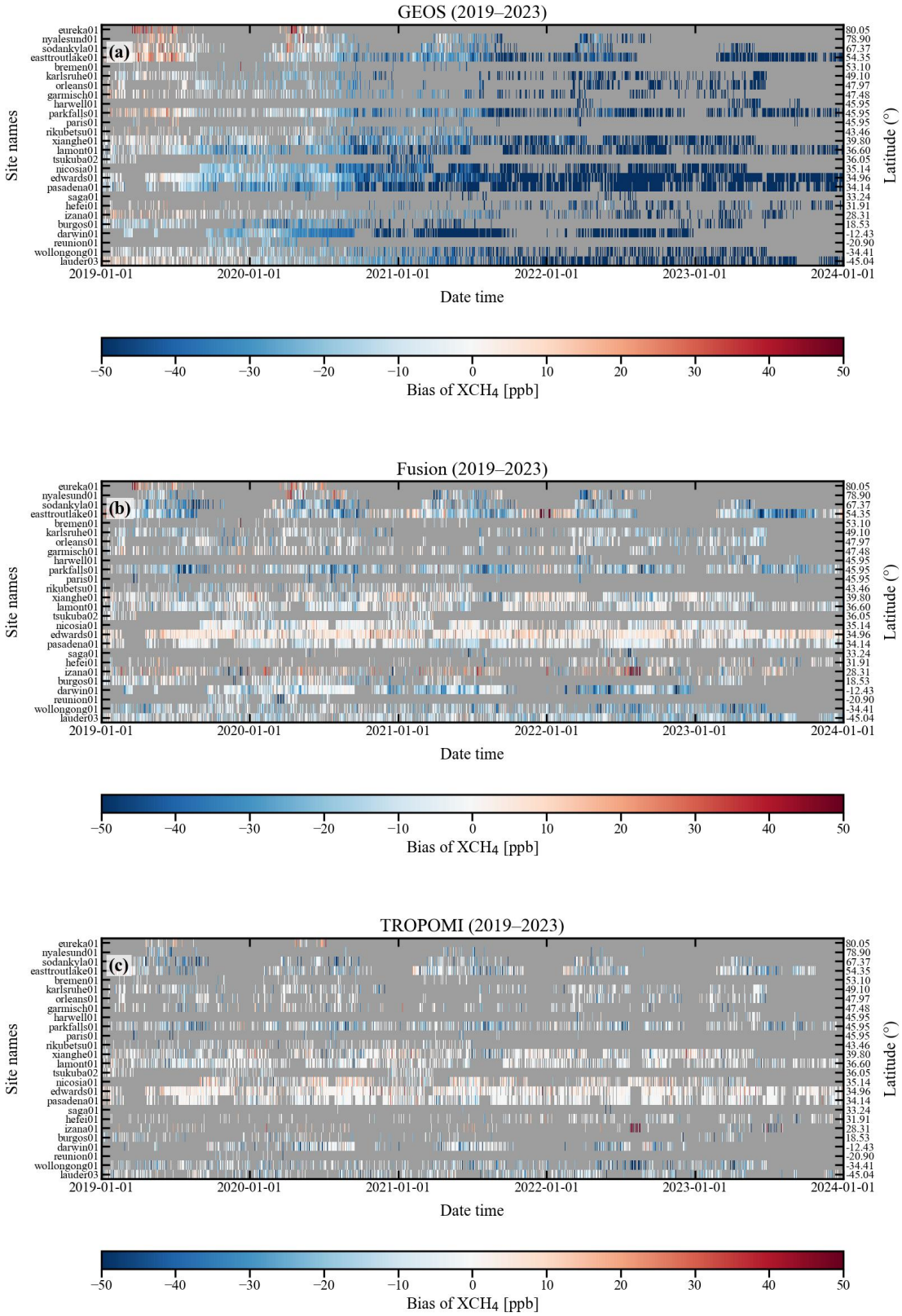


Figure 11. Heat maps of the biases between daily (a) GEOS, (b) fused, and (c) TROPOMI XCH_4 and TCCON XCH_4 over time and latitude. Color ramps stand for the biases of XCH_4. Background colors (grey) indicate the missing data.

The masked residual-learning strategy itself introduces uncertainty in persistently missing regions. Because the residual U-Net is trained only at valid TROPOMI pixels, the model learns the statistical relationship between the preliminary reconstruction error and the input predictors under observed conditions. When applied to unobserved regions—especially those with persistent cloud cover or aerosol contamination—the model infers the most likely residual pattern given similar meteorological and spatiotemporal conditions, rather than the exact unobserved atmospheric state on a specific day. This does not invalidate the gap-filling framework, but it implies that reconstructed values in long-term data gaps should be understood as observation-constrained estimates rather than direct measurements. Users should be aware of this extrapolation uncertainty when analyzing local short-lived anomalies in persistently cloudy regions.

Finally, possible discontinuities associated with changes in the input satellite product should be considered. The TROPOMI pixel size changed around June 2019 from 7.0 km × 7.0 km to 7.0 km × 5.5 km, which may affect sampling density and local spatial representativeness. Although the fusion framework uses gridding, quality filtering, and residual correction to mitigate such effects, subtle artifacts or temporal discontinuities may remain. Caution is advised when interpreting local-scale or short-term variations that span this instrumental transition.

In summary, the fused datasets are most suitable for studying daily to seasonal spatiotemporal variability, improving data continuity and spatial coverage, examining regional-scale events (e.g., biomass burning emissions), and supporting comparative analyses during 2019–2023. For fine-scale local anomaly detection in persistently missing regions, for applications in areas lacking independent ground-based validation, or for studies requiring strict profile-sensitive validation or long-term climate trend attribution, users should consult the observation-support variables, including obs_mask, MR, and observation_fraction, and carefully consider all the limitations discussed above.

Comment 49: *line 577. "All created fusion data may be obtained upon request from the authors for academics and policymakers." Different info is given below (line 592).*

Response 49:

We thank the reviewer for pointing out this inconsistency regarding data access. We have revised the manuscript to harmonize the description of data availability. Specifically, the previous statement that the fusion data may be obtained upon request has been removed, and the Data Availability statement has been moved to the end of the manuscript and aligned with the repository information. The revised text has been added to Lines 939–941 of the revised manuscript.

Revised text:

The global (0.25°) and China-specific (0.05°) daily gap-free XCO and XCH_4 datasets (2019–2023) generated in this study are openly available in the Zenodo repository at https://doi.org/10.5281/zenodo.17936461. (An et al., 2026).

Comment 50: *line 592-595. Data availability statement. This should be given in a separate section (which is placed below).*

Response 50:

We appreciate the reviewer's guidance on the manuscript structure. We have moved the Data Availability statement to its own separate section at the end of the manuscript, as per the journal's standard formatting requirements. The revised Data Availability statement can be found in Lines 939–941 of the revised manuscript.

Revised text:

Data availability statement

The global (0.25°) and China-specific (0.05°) daily gap-free XCO and XCH_4 datasets (2019–2023) generated in this study are openly available in the Zenodo repository at https://doi.org/10.5281/zenodo.17936461. (An et al., 2026).

Comment 51: *A lot of references have incomplete information (e.g., lines 713, 847, 875).*

Response 51:

We sincerely thank you for carefully checking the references in our manuscript. We have thoroughly reviewed the reference list and corrected all incomplete entries, including those previously identified at Lines 713, 847, and 875. We have ensured that all cited references now include complete bibliographic information, such as volume numbers, page ranges, and available DOIs, in accordance with the journal's publication standards.

Specifically, the reference originally listed at Line 713 has been removed. The reference previously at Line 847 has been moved to Line 1221 of the revised manuscript, and the reference previously at Line 875 has been moved to Line 1271 of the revised manuscript.

Revised text:

Wang, Y., Yuan, Q., Li, T., Yang, Y., Zhou, S., Zhang, L., 2023. Seamless mapping of long-term (2010–2020) daily global XCO2 and XCH4 from the Greenhouse Gases Observing Satellite (GOSAT), Orbiting Carbon Observatory 2 (OCO-2), and CAMS global greenhouse gas reanalysis (CAMS-EGG4) with a spatiotemporally self-supervised fusion method. Earth Syst. Sci. Data 15, 3597-3622. https://doi.org/10.5194/essd-15-3597-2023

Yang, Y., Zhou, M., Langerock, B., Sha, M.K., Hermans, C., Wang, T., Ji, D., Vigouroux, C., Kumps, N., Wang, G., 2020. New ground-based Fourier-transform near-infrared solar absorption measurements of XCO_2, XCH_4 and XCO at Xianghe, China. Earth Syst. Sci. Data 12, 1679-1696. https://doi.org/10.5194/essd-12-1679-2020

Response to Reviewer 4:

This manuscript presents a fusion framework combining signal-domain reconstruction (DCT + SVD) with a residual deep learning model to generate gap-free global XCO and XCH_4 datasets

based on TROPOMI and GEOS-Chem simulations. The topic is clearly relevant and producing continuous datasets of atmospheric composition is important for many applications. I am not a specialist in all aspects of the methodology, but I found the general idea interesting. At the same time, I think the manuscript still requires major revision before it can be considered for publication. Several points related to method clarity, validation and dataset description need to be strengthened.

Response

Thank you for your positive evaluation of our study and for recognizing the potential value of the fused datasets generated using the two-stage framework of signal-domain reconstruction and residual learning. At the same time, we also thank you for your concern that this study needs to provide clearer methods, dataset validation, and dataset description.

In response to your comments on methodological and validation details, we have substantially revised the manuscript to improve transparency and reproducibility. Specifically, to enhance methodological clarity, we clarified the satellite product versions and data sources, the definition of the missing rate, the DCT/SVD and residual learning methods, and the masked loss strategy. To clarify the validation process, we specified the site selection criteria and the calculation methods for relevant statistics, including R^2, RMSE, and μ. In terms of dataset description, we added a summary of the fused dataset so that users can use the data more clearly.

Major comments

Comment 1:

Positioning and novelty

It would help to better clarify what is new compared to existing gap-filling or fusion approaches. The method seems to combine known components (signal processing + deep learning), but it is not fully clear what specific limitation of previous work is being addressed, or how much improvement is actually achieved.

Response 1:

We thank you for this comment. In the revised manuscript, We clarified the limitations of existing studies in this research direction and the innovations of the present work.We also conducted ablation experiments to demonstrate the specific contribution of each stage of our framework. Finally, we evaluated the fused datasets under different levels of TROPOMI coverage and demonstrated that, under low-coverage conditions, the fused datasets show slightly better performance than the TROPOMI datasets.

Revised text:

The generation of continuous TROPOMI XCO and XCH_4 products has been addressed through a range of gap-filling approaches. First, machine learning-based interpolation methods (X. Chen et al., 2022; Hu et al., 2022; Valerio et al., 2025; Wei et al., 2025) reconstruct missing observations by learning nonlinear spatiotemporal relationships. However, their performance strongly depends

on the quality and representativeness of the training data, and biased or insufficient samples may lead to unreliable predictions. Second, enhanced spectral fitting algorithms (Borsdorff et al., 2019; Guanter et al., 2015; Schneising et al., 2023; Wang et al., 2020) improve retrieval continuity by refining the spectral inversion process, for example through higher-order polynomial fitting to reduce reflectance-induced biases, albeit at increased computational cost. Third, physics–data–driven coupling approaches, including model fusion, data assimilation, and hybrid machine learning frameworks, integrate satellite retrievals with chemical transport models or reanalysis datasets (Fritz et al., 2022; Schneising et al., 2023; Sicard et al., 2021). By combining physical constraints with data-driven learning or observational corrections, these methods improve spatial consistency and temporal continuity while reducing retrieval bias. However, their performance remains dependent on uncertainties in prior model simulations and the representativeness of assimilated observations (Inness et al., 2022; Wang et al., 2023). (The limitations of the current method are discussed in Lines 95–110 of the revised manuscript.)

The proposed method specifically addresses the shortcomings of existing approaches in three key aspects: (1) Addressing Data Sparsity and Non-linearity: Unlike pure machine learning interpolation that may struggle with unbalanced datasets, we employ a two-stage strategy. We first use 3D Discrete Cosine Transform (DCT) and Singular Value Decomposition (SVD) to perform low-rank signal-domain reconstruction (Rao and Yip, 2014; Wall et al., 2003). This step exploits the shared spatiotemporal structure between GEOS-Chem simulations and TROPOMI observations to approximate missing values in the frequency domain, providing a robust prior; (2) Enhancing Physical Consistency: To correct residual biases without overfitting, we implement a lightweight residual U-Net for pixel-level refinement (Ronneberger et al., 2015; Tang, 2022). Crucially, instead of using the TROPOMI data mask as a direct input feature, we incorporate it as a weighting mask in the loss function (Wei et al., 2022). This masked learning strategy ensures that the model is supervised exclusively by valid observations, thereby preserving the physical consistency of the DCT/SVD-derived signal-domain reconstructed fields; (3) Improving Efficiency and Accuracy: By fusing model-driven priors and observational constraints in both frequency and spatial domains, our method significantly reduces computational load compared to enhanced spectral fitting while avoiding the "black-box" limitation of pure data-driven methods. (The novelty of the method proposed in this study is described in Lines 118–133 of the revised manuscript)

2.2.5. Ablation study design

To explicitly isolate the contribution of each component in the proposed two-stage framework, we conducted an ablation study using site-based validation across all TCCON stations for the year 2021. Five configurations were compared: (1) only DCT fusion(Garcia, 2010), (2) DCT and SVD mixed-signal reconstruction(M. Bengherabi et al., 2008; Majhi and Pal, 2021), (3) residual CNN(Y. Jiang et al., 2024), (4) residual U-Net (Ronneberger et al., 2015; Yan et al., 2022), and (5) residual XGBoost(Naseem et al., 2024). The performance was assessed using the coefficient of determination (R^2), root mean square error (RMSE), and mean bias (μ). This design allows the contribution of the DCT/SVD signal-domain reconstruction stage and the residual-correction stage to be evaluated separately. The site-based validation results are summarized in Tables S2 and S3. (The ablation experiments added to distinguish the contributions of different research stages have been included in Lines 463–472 of the revised manuscript.)

Table S2. Ablation study of site-based XCO validation using different reconstruction configurations. Metrics include coefficient of determination (R^2), root mean square error (RMSE, ppb), and mean bias (μ, ppb).

Configuration	R^2	RMSE	μ
DCT-only	0.8227	10.439	-3.815
DCT/SVD	0.8335	10.1166	-3.7682
DCT/SVD + residual CNN	0.8385	9.9619	-3.7677
DCT/SVD + residual U-Net	**0.845**	**9.7616**	**-3.686**
DCT/SVD + residual XGBoost	0.8392	9.9406	-3.7435

Table S3. Ablation study of site-based XCH_4 validation using different reconstruction configurations. Metrics include coefficient of determination (R^2), root mean square error (RMSE, ppb), and mean bias (μ, ppb).

Configuration	R^2	RMSE	μ
DCT-only	0.7056	17.5156	-6.9345
DCT/SVD	0.732	16.7111	-6.9635
DCT/SVD + residual CNN	0.7469	16.2404	-6.9162
DCT/SVD + residual U-Net	**0.7598**	**15.8212**	**-6.7858**
DCT/SVD + residual XGBoost	0.7476	16.2188	-6.9563

To demonstrate the performance of our fused datasets under different TROPOMI coverage, we conducted experiments using different MR thresholds. Tables S6 and S7 further quantify how the performance of the fused datasets varies with TROPOMI coverage. For XCO, when the missing rate is low or moderate ($MR < 0.3$ and $0.3 \leq MR \leq 0.5$), TROPOMI shows slightly higher R^2 values and lower RMSE values than the fused dataset, indicating the high reliability of TROPOMI retrievals under relatively good coverage conditions. However, under sparse TROPOMI coverage conditions ($MR > 0.5$), the performance of TROPOMI decreases, with its R^2 decreasing to 0.89 and RMSE increasing to 6.16 ppb. In contrast, the fused XCO data maintain a higher R^2 value of 0.91 and a lower RMSE value of 5.65 ppb, indicating greater robustness under high-missing-rate conditions. For XCH_4, a similar pattern is observed. TROPOMI performs better than the fused product under low and moderate missing-rate conditions, with higher R^2 and lower RMSE. However, when $MR > 0.5$, TROPOMI performance decreases markedly, with R^2 declining to 0.81 and RMSE increasing to 15.84 ppb. The fused XCH_4 data show a slightly higher R^2 of 0.83 and a lower RMSE of 15.40 ppb under the same high-missing-rate condition. Although the fused XCH_4 bias is larger than that of TROPOMI, the higher R^2 and lower RMSE suggest that the fused dataset reduces overall errors under sparse-observation conditions. Overall, under conditions of good TROPOMI coverage, the fused XCO and XCH_4 datasets perform comparably to the TROPOMI datasets. When TROPOMI retrievals are sparse, the fused datasets show relatively high reliability. These results indicate that our fusion framework is effective for gap-filling applications. (The performance of the fused datasets under different levels of TROPOMI coverage is evaluated in Lines 663–682 of the revised manuscript.)

Table S6. Validation statistics of XCO under different TROPOMI missing-rate intervals. GEOS-Chem results were bias-corrected before validation, and RMSE and μ are reported in parts per billion (ppb).

MR	R^2			RMSE			μ		
	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION
MR < 0.3	0.65	0.97	0.94	16.60	2.04	5.69	-0.22	-0.43	-1.10
0.3 ≤ MR ≤ 0.5	0.63	0.96	0.93	13.44	2.96	5.16	0.02	-0.88	-1.21
MR > 0.5	0.48	0.89	0.91	11.99	6.16	5.65	0.79	-0.87	-1.40

Table S7. Validation statistics of XCH_4 under different TROPOMI missing-rate intervals. GEOS-Chem results were bias-corrected before validation, and RMSE and μ are reported in parts per billion (ppb).

MR	R^2			RMSE			μ		
	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION
MR < 0.3	-0.15	0.95	0.87	29.10	3.05	9.66	-2.87	-0.57	-4.37
0.3 ≤ MR ≤ 0.5	-0.17	0.94	0.85	28.39	3.84	10.78	-3.19	-0.65	-2.17
MR > 0.5	0.34	0.81	0.83	29.84	15.84	15.40	-0.80	-5.48	-7.30

Comment 2:

Method description and justification

Some methodological choices are not fully explained:

1. *The use of DCT + SVD and the choice of keeping 80% of the energy are not justified or tested*
2. *It is difficult to understand the relative contribution of each step (reconstruction vs. neural network refinement)*
3. *Important implementation details (training setup, parameters, etc.) are missing*

Overall, more transparency would help readers assess robustness and reproducibility.

Response 2:

We thank you for this constructive comment. In the revised manuscript, we have addressed the methodological issues you raised, and we provide point-by-point explanations for each issue in this response. The detailed explanations are as follows:

1. we clarified the rationale for using DCT + SVD. The DCT-based reconstruction is used to exploit the smooth low-frequency spatiotemporal structure shared by GEOS-Chem and TROPOMI, while SVD is applied to retain the dominant low-rank components and suppress weak noisy components in the reconstructed parameter matrix. To justify the empirical 80% cumulative singular-value energy threshold, we added a threshold-sensitivity experiment using 2019 daily global XCO. Thresholds of 70%, 75%, 80%, 85%, 90%, and 95% were tested. The reconstructed fields were highly stable across this range, with spatial correlations > 0.9998 relative to the 80% reference case, RMSE differences < 0.62 ppb, spatial standard deviation ratios close to 1, and hotspot overlap ratios > 0.97. These results support the use of 80% as a stable and computationally efficient threshold.

Table S8. Sensitivity of the 2019 daily XCO reconstruction to the SVD cumulative-energy threshold.

SVD energy threshold (%)	Spatial Rvs. 80%	RMSE vs. 80% (ppb)	μ vs. 80% (ppb)	SD ratio vs. 80%	Hotspot overlap
70	0.99998	0.206	-0.088	0.99989	0.983
75	0.99999	0.160	-0.086	0.99990	0.984
80	1.00000	0.000	0.000	1.00000	1.000
85	0.99996	0.271	0.000	1.00000	1.000
90	0.99992	0.382	0.003	0.99965	0.999
95	0.99984	0.616	0.285	1.00019	0.975

Note: The 80% threshold was used as the reference case. Spatial R denotes the Pearson correlation coefficient calculated across valid grid cells between each threshold experiment and the 80% case. RMSE, μ, and SD ratio were also calculated relative to the 80% case, where SD denotes the spatial standard deviation. Hotspot overlap was calculated using the upper 10% of reconstructed XCO grid cells. All statistics were derived from the 2019 global daily 0.25° × 0.25° CO reconstruction experiments.

2. we added a stage-wise ablation analysis to clarify the relative contribution of each component. The comparison between "Only DCT" and "DCT/SVD mixed-signal reconstruction" isolates the contribution of SVD, while the comparison between DCT/SVD reconstruction and residual-learning models isolates the contribution of neural-network refinement. For XCO, adding SVD increased R^2 from 0.8227 to 0.8335 and reduced RMSE from 10.439 to 10.1166 ppb; the residual U-Net further improved R^2 to 0.8450 and reduced RMSE to 9.7616 ppb. For XCH_4, adding SVD increased R^2 from 0.7056 to 0.7320 and reduced RMSE from 17.5156 to 16.7111 ppb; the residual U-Net further improved R^2 to 0.7598 and reduced RMSE to 15.8212 ppb. These results show that both the signal-domain reconstruction and residual refinement stages make measurable contributions.

Table S2. Ablation study of site-based XCO validation using different reconstruction configurations. Metrics include coefficient of determination (R^2), root mean square error (RMSE, ppb), and mean bias (μ, ppb).

Configuration	R^2	RMSE	μ
DCT-only	0.8227	10.439	-3.815
DCT/SVD	0.8335	10.1166	-3.7682
DCT/SVD + residual CNN	0.8385	9.9619	-3.7677
DCT/SVD + residual U-Net	**0.845**	**9.7616**	**-3.686**
DCT/SVD + residual XGBoost	0.8392	9.9406	-3.7435

Table S3. Ablation study of site-based XCH_4 validation using different reconstruction configurations. Metrics include coefficient of determination (R^2), root mean square error (RMSE, ppb), and mean bias (μ, ppb).

Column	R^2	RMSE	μ
DCT-only	0.7056	17.5156	-6.9345
DCT/SVD	0.732	16.7111	-6.9635
DCT/SVD + residual CNN	0.7469	16.2404	-6.9162

DCT/SVD + residual U-Net	**0.7598**	**15.8212**	**-6.7858**
DCT/SVD + residual XGBoost	0.7476	16.2188	-6.9563

3. we added implementation details for reproducibility. The revised manuscript now describes the residual U-Net architecture, inputs, masked loss, and training setup. The network uses an encoder-decoder structure with skip connections and lightweight residual blocks, containing 10 convolutional layers, 32 filters in hidden layers, two input channels, and one output channel for the residual field. The model was trained using daily global maps from 2019-2023 for 5 epochs, with a batch size of 4, the Adam optimizer, and a fixed learning rate of 1×10^{-4}. Input variables were normalized before training, and the TROPOMI mask was used only in the loss function rather than as a model input.

Comment 3:

Validation strategy

The validation is useful but still raises some questions:

1. *The comparison with TCCON does not fully address the mismatch between point measurements and gridded data*
2. *The selection of validation sites is a bit unclear and seems inconsistent in places*
3. *Since the goal is gap filling, it would be important to show performance in regions with high missing data (e.g. cloudy areas)*

Uncertainty estimates or more systematic validation would also strengthen the results.

Response 3:

We thank you for your valuable comments on our validation strategy. We agree that the original manuscript did not sufficiently clarify the mismatch between TCCON point-scale observations and gridded satellite/model products, the site-selection criteria, or the validation performance under sparse TROPOMI coverage conditions. We have revised the TCCON validation section to address these issues.

1. we added a discussion of the spatial representativeness mismatch. TCCON provides near-point-scale measurements, whereas TROPOMI, GEOS-Chem, and the fused datasets provide gridded values. Therefore, exact point-to-grid equivalence cannot be expected. We clarified that the 2° radius was used as a compromise to include sufficient satellite and gridded samples around each station while limiting the influence of distant air masses.

2. We clarified the site-selection criteria. For the time-series examples, after applying the 13:30 ± 1 h local-time window, the 2° spatial collocation radius, and the qa_value > 0.5 quality filter, we selected stations with relatively continuous multi-year records and limited long data gaps. To evaluate the performance of the fused datasets in regions with sparse TROPOMI coverage, we applied the 13:30 ± 1 h local-time window, the 2° spatial collocation radius, the qa_value > 0.5 quality filter, and the MR > 0.5 criterion to obtain collocated samples. Representative stations with a relatively large number of valid collocated samples were then selected for individual site-level visualization. Finally, for the overall validation of the fused dataset performance, all site-day

collocations with MR > 0.5 from all qualified stations were included in the statistical calculations and used to draw the final conclusions.

3. because gap filling is a central objective of this study, we explicitly evaluated the fused datasets under sparse TROPOMI coverage conditions. We defined the TROPOMI missing rate (MR) as the fraction of 0.25° grid cells within the 2° collocation radius that had no valid TROPOMI retrieval after quality filtering. Site-day collocations with MR > 0.5 were used to assess performance under high-missing-rate conditions. The representative scatterplots in Figs. 4 and 5 and the overall validation in Fig. 6 now explicitly indicate that they are based on sparse TROPOMI coverage conditions.

4. To provide a more systematic validation, we added Tables S6 and S7 to quantify the validation performance under different MR intervals (MR < 0.3, 0.3 ≤ MR ≤ 0.5, and MR > 0.5). The results show that, under low and moderate missing-rate conditions, the fused XCO and XCH_4 datasets perform comparably to the TROPOMI datasets, whereas under high-missing-rate conditions, the fused datasets slightly outperform the TROPOMI datasets.

Revised text:

1. We acknowledge that TCCON provides near-point-scale column measurements, whereas the TROPOMI and fused datasets represent spatially averaged grid-cell values at 0.25° resolution. Therefore, a spatial representativeness mismatch is unavoidable, especially in regions with strong horizontal gradients, complex terrain, coastal influences, or localized emission sources. The 2° radius was used as a compromise to include a sufficient number of satellite and gridded data samples around each station while limiting the influence of distant air masses. (Lines 496–502)

2. Figures 2 and 3 illustrate representative time series of daily GEOS-Chem simulations, TROPOMI observations, fused data, and TCCON XCO and XCH_4 measurements at selected ground stations. The ground stations were selected according to the following criteria. For each gas, we first screened TCCON stations that retained a relatively large number of valid spatially collocated data pairs after applying the 13:30 ± 1 h temporal window, the 2° spatial matching radius, and the TROPOMI quality filter (qa_value > 0.5). Among these qualified stations, we further selected those with relatively continuous multi-year records and limited long data gaps to illustrate the temporal consistency among GEOS-Chem, TROPOMI, the fused data, and TCCON observations.

Following the representative time-series comparison, we further evaluated site-level statistical agreement using scatterplots at selected representative stations under sparse TROPOMI coverage conditions in Figs. 4 and 5. This selection was based on the number of valid collocated samples after applying the missing-rate criterion, rather than on validation performance. Specifically, for each TCCON station, we first applied the qa_value > 0.5 quality filter, the 13:30 ± 1 h temporal window, the 2° spatial collocation radius, and the site-day missing-rate criterion (MR > 0.5). Subsequently, stations with a relatively large number of valid collocated samples after applying the missing-rate screening criterion were selected for visualization, whereas the overall validation statistics were calculated using all qualified high-missing-rate site-day collocations rather than only these representative stations. (Lines 581–590)

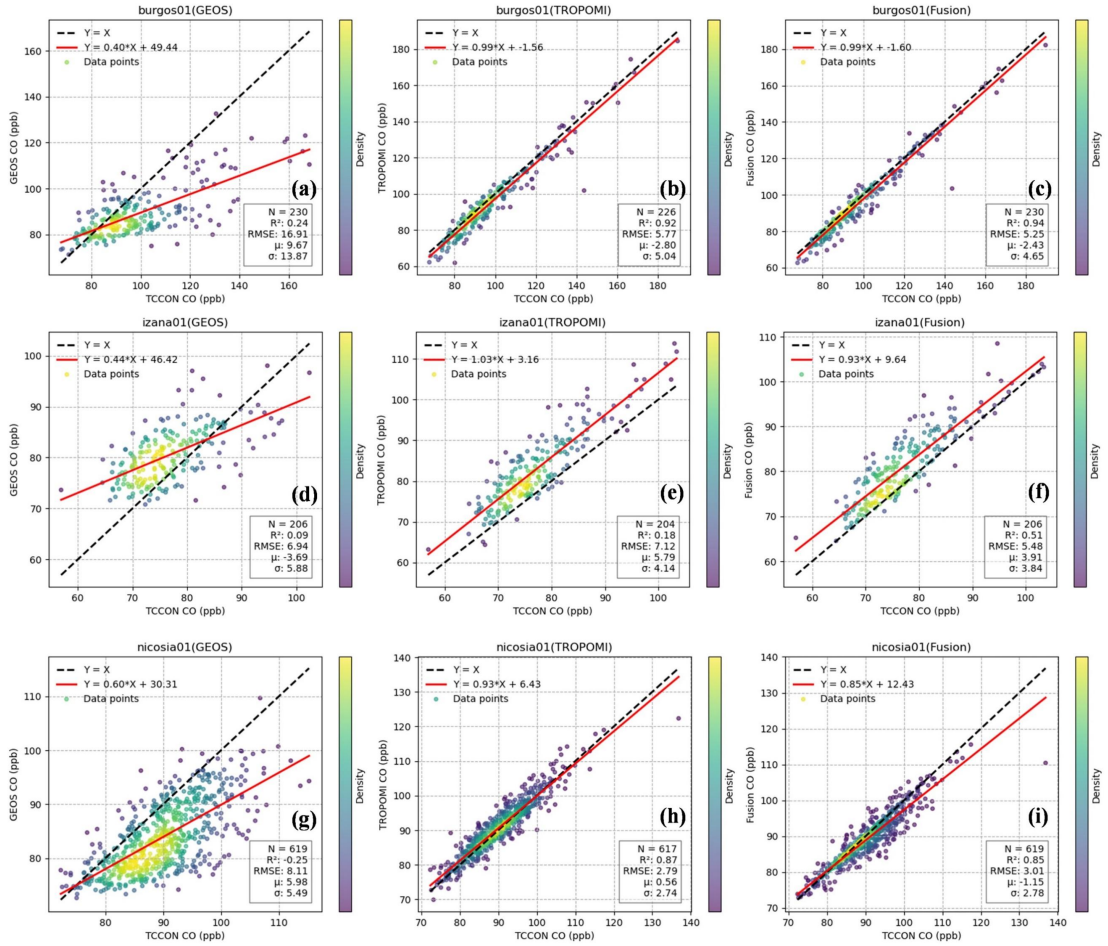


Figure 4. Density scatterplots of the independent validation results for GEOS-Chem, TROPOMI, and fused XCO data at burgos01, izana01, and nicosia01 under sparse TROPOMI coverage conditions (MR > 0.5). Panels (a, d, g), (b, e, h), and (c, f, i) correspond to GEOS-Chem, TROPOMI, and fused XCO data, respectively. Black dotted and red solid lines indicate the 1:1 line and fitted regression line, respectively. X: TCCON data; Y: GEOS-Chem, TROPOMI, and fused data. Units: ppb for RMSE, μ, and σ.

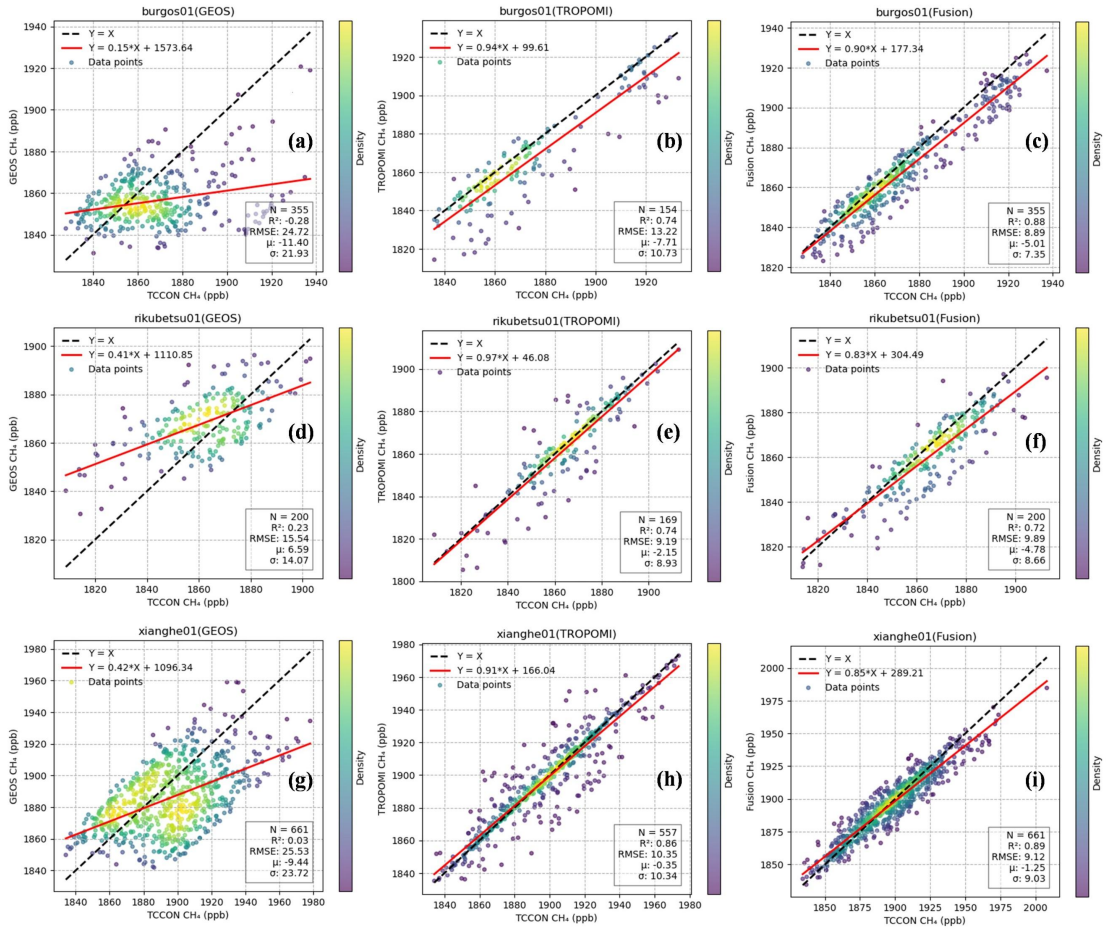


Figure 5. Density scatterplots of the independent validation results for GEOS-Chem, TROPOMI, and fused XCH_4 data at burgos01, rikubetsu01, and xianghe01 under sparse TROPOMI coverage conditions (MR > 0.5). Panels (a, d, g), (b, e, h), and (c, f, i) correspond to GEOS-Chem, TROPOMI, and fused XCH_4 data, respectively. Black dotted and red solid lines indicate the 1:1 line and fitted regression line, respectively. X: TCCON data; Y: GEOS-Chem, TROPOMI, and fused data. Units: ppb for RMSE, μ, and σ.

3. To demonstrate the performance of our fused datasets under different TROPOMI coverage, we conducted experiments using different MR thresholds. Tables S6 and S7 further quantify how the performance of the fused datasets varies with TROPOMI coverage. For XCO, when the missing rate is low or moderate (MR < 0.3 and 0.3 ≤ MR ≤ 0.5), TROPOMI shows slightly higher R^2 values and lower RMSE values than the fused dataset, indicating the high reliability of TROPOMI retrievals under relatively good coverage conditions. However, under sparse TROPOMI coverage conditions (MR > 0.5), the performance of TROPOMI decreases, with its R^2 decreasing to 0.89 and RMSE increasing to 6.16 ppb. In contrast, the fused XCO data maintain a higher R^2 value of 0.91 and a lower RMSE value of 5.65 ppb, indicating greater robustness under high-missing-rate conditions. For XCH_4, a similar pattern is observed. TROPOMI performs better than the fused product under low and moderate missing-rate conditions, with higher R^2 and lower RMSE. However, when MR > 0.5, TROPOMI performance decreases markedly, with R^2 declining to 0.81 and RMSE increasing to 15.84 ppb. The fused XCH_4 data show a slightly higher R^2 of 0.83 and a lower RMSE of 15.40 ppb under the same high-missing-rate condition. Although the fused XCH_4

bias is larger than that of TROPOMI, the higher R^2 and lower RMSE suggest that the fused dataset reduces overall errors under sparse-observation conditions. Overall, under conditions of good TROPOMI coverage, the fused XCO and XCH_4 datasets perform comparably to the TROPOMI datasets. When TROPOMI retrievals are sparse, the fused datasets show relatively high reliability. These results indicate that our fusion framework is effective for gap-filling applications. (Lines 644–662)

Table S6. Validation statistics of XCO under different TROPOMI missing-rate intervals. GEOS-Chem results were bias-corrected before validation, and RMSE and μ are reported in parts per billion (ppb).

MR	R^2			RMSE			μ		
	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION
MR < 0.3	0.65	0.97	0.94	16.60	2.04	5.69	-0.22	-0.43	-1.10
0.3 ≤ MR ≤ 0.5	0.63	0.96	0.93	13.44	2.96	5.16	0.02	-0.88	-1.21
MR > 0.5	0.48	0.89	0.91	11.99	6.16	5.65	0.79	-0.87	-1.40

Table S7. Validation statistics of XCH_4 under different TROPOMI missing-rate intervals. GEOS-Chem results were bias-corrected before validation, and RMSE and μ are reported in parts per billion (ppb).

MR	R^2			RMSE			μ		
	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION	GEOS	TROPOMI	FUSION
MR < 0.3	-0.15	0.95	0.87	29.10	3.05	9.66	-2.87	-0.57	-4.37
0.3 ≤ MR ≤ 0.5	-0.17	0.94	0.85	28.39	3.84	10.78	-3.19	-0.65	-2.17
MR > 0.5	0.34	0.81	0.83	29.84	15.84	15.40	-0.80	-5.48	-7.30

Comment 4:

Interpretation of results

Some of the conclusions feel a bit strong given the analysis:

(1) Statements such as the dataset being "superior to TROPOMI" should probably be phrased more carefully;

(2) The time period (2019–2023) is relatively short for trend analysis, so conclusions about long-term changes should be treated with caution;

(3) The discussion of spatial patterns is mostly descriptive, and could benefit from more quantitative support.

Response 4:

Thank you for this constructive comment. We agree that several statements in the original manuscript were too strong and that the interpretation of the results should be more cautious. We have revised the manuscript accordingly.

1. We softened statements such as "superior to TROPOMI" and replaced them with more objective descriptions, such as "comparable to or slightly improved relative to TROPOMI" or "improved for several validation metrics." This better reflects the validation results, because the

fused datasets improve spatial and temporal continuity while maintaining validation performance comparable to TROPOMI.

2. We agree that the 2019-2023 period is relatively short for robust long-term trend analysis. To avoid overinterpretation, we replaced terms such as “trend” and “long-term trend” with “short-term change” or “short-term rate of change.” We also added harmonic-regression-based estimates and 95% confidence intervals for the short-term changes in XCO and XCH_4 over the rice-growing regions of Northeast and Southeast China. This content can be found in Lines 745–777 of the revised manuscript.

Revised text:

Therefore, we further analyzed the short-term temporal changes in XCO and XCH_4 over the major rice-growing regions of Northeast and Southeast China to evaluate the applicability of the high-resolution fused datasets in agricultural emission-related regions. Figures 8 and 9 show the temporal variations in XCO and XCH_4 derived from TROPOMI and the fused data over these regions, respectively. To reduce the influence of the seasonal cycle, the short-term rates of change were derived using harmonic regression with annual and semiannual terms, and the corresponding 95% confidence intervals are provided in Tables S9 and S10.

The results show that XCO exhibits negative short-term rates of change in both the satellite and fused datasets across the rice cultivation areas of Northeast and Southeast China. In Northeast China, the XCO rate is −2.36 ppb/year for TROPOMI and −2.28 ppb/year for the fused data. In Southeast China, the XCO rate is −1.29 ppb/year for TROPOMI and −1.76 ppb/year for the fused data. The consistent negative rates indicate that the fused dataset preserves the main decreasing XCO signal observed by TROPOMI. Differences between the two datasets may be partly related to the uneven temporal sampling of TROPOMI caused by clouds, aerosols, and retrieval filtering, whereas the fused dataset provides spatially and temporally continuous fields.

In contrast, XCH_4 exhibits positive short-term rates of change in both the satellite and fused datasets. In Northeast China, the XCH_4 rate is 12.08 ppb/year for TROPOMI and 12.16 ppb/year for the fused data. In Southeast China, the XCH_4 rate is 12.16 ppb/year for TROPOMI and 11.78 ppb/year for the fused data. The close agreement between the two datasets indicates that the fused dataset preserves the main increasing XCH_4 signal observed by TROPOMI while providing improved temporal continuity.

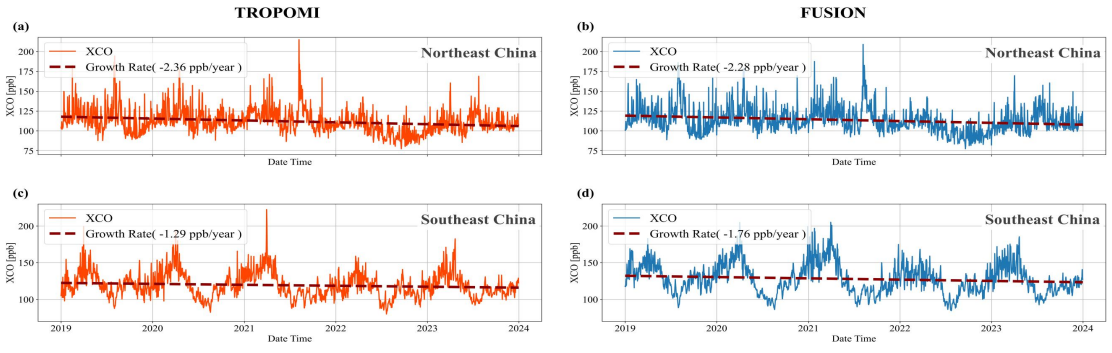


Figure 8. Daily time series of TROPOMI and fused XCO data over rice-growing regions in (a, b) Northeast China and (c, d) Southeast China. Red dashed lines indicate short-term rates of change

derived from harmonic regression after accounting for the seasonal cycle, with values shown in parentheses. Unit: parts per billion per year.

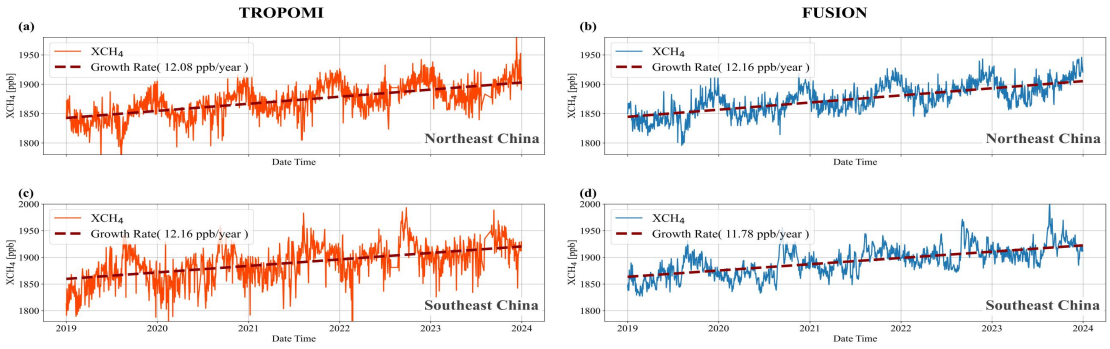


Figure 9. Daily time series of TROPOMI and fused XCH_4 data over rice-growing regions in (a, b) Northeast China and (c, d) Southeast China. Red dashed lines indicate short-term rates of change derived from harmonic regression after accounting for the seasonal cycle, with values shown in parentheses. Unit: parts per billion per year.

Table S9. Seasonally adjusted short-term rates of change and 95% confidence intervals for XCO over rice-growing regions during 2019–2023. Unit: ppb/year.

Region	Rate of change		95% confidence interval	
	TROPOMI	Fusion	TROPOMI	Fusion
Northeast China	−2.36	−2.28	−3.30 to −1.43	−3.30 to −1.27
Southeast China	−1.29	−1.76	−2.15 to −0.43	−2.79 to −0.72

Table S10. Seasonally adjusted short-term rates of change and 95% confidence intervals for XCH_4 over rice-growing regions during 2019–2023. Unit: ppb/year.

Region	Rate of change		95% confidence interval	
	TROPOMI	Fusion	TROPOMI	Fusion
Northeast China	12.08	12.16	11.24 to 12.92	11.27 to 13.04
Southeast China	12.16	11.78	10.83 to 13.49	9.72 to 13.83

3. We added more quantitative support to the discussion of spatial patterns and application cases. These additions include TCCON-based validation statistics, site-level and all-site metrics, bias heat maps, short-term rates of change with confidence intervals, and quantified concentration enhancements during the Chongqing fire event. These revisions make the interpretation more balanced and better supported by quantitative evidence.

Comment 5:

Dataset

From a data perspective, more information is needed:

(1) The dataset files appear to lack some basic metadata and clear variable descriptions;

(2) It is not fully clear what variables are included and how they should be used

(3) Some form of uncertainty or quality information would be very useful for users

A dedicated section summarizing the dataset (variables, format, limitations) would improve usability.

Response 5:

Thank you for this helpful suggestion. We agree that clearer dataset documentation is important for users. In the revised manuscript, we added a dedicated section entitled “Fused dataset summary” to describe the dataset structure, variables, format, metadata, and limitations.

1. We clarified that the final products are daily gap-free fused XCO and XCH_4 datasets for 2019–2023, provided at two spatial scales: a global 0.25° × 0.25° product and a China-specific 0.05° × 0.05° product. The datasets are distributed in NetCDF4 format, with monthly files containing daily gridded fields along the dimensions time,lat, and lon.

2. We also added explicit descriptions of all included variables. The main variables are XCO and XCH4, representing fused column-averaged dry-air mole fractions in ppb. To provide quality and uncertainty-related information, we included two additional variables: obs_mask and observation_fraction. The obs_mask variable indicates whether each grid-cell value on each day is directly constrained by a valid TROPOMI retrieval (obs_mask = 1) or reconstructed by the fusion framework (obs_mask = 0). The observation_fraction variable gives the monthly fraction of days with valid TROPOMI observations at each grid cell, helping users assess the observational support behind the gap-free product.

In addition, we revised the NetCDF files to include CF-style metadata, including coordinate attributes, variable long names, units, missing-value definitions, temporal coverage, spatial resolution, product region, data source, author, institution, and contact information. We also clarified that the monthly file organization is only for file-size and repository-management purposes and does not indicate monthly averaging; each file still contains daily fields.

3. We added guidance on dataset limitations, emphasizing that users should consult obs_mask andobservation_fraction when interpreting reconstructed values, especially in regions with persistent missing TROPOMI observations or limited independent validation.

We have added the relevant description in Lines 853–892 of the revised manuscript.

Revised text:

The final data products generated in this study consist of daily gap-free fused XCO and XCH_4 datasets for the period 2019–2023. Two spatial scales are provided. The global product contains daily XCO and XCH_4 fields at 0.25° × 0.25° resolution, covering the global domain. The regional product provides higher-resolution daily XCO and XCH_4 fields over China at 0.05° × 0.05° resolution. Both products are reported as column-averaged dry-air mole fractions in parts per billion (ppb).

The datasets are distributed in NetCDF4 (.nc/.nc4) format. Each monthly file contains gridded daily fields with the dimensions `time`, `lat`, and `lon`. The `time` coordinate is defined using CF-style physical time units, for example, `days since 2019-01-01 00:00:00`, with `time = 0` corresponding to the first day of the month. Users can therefore retrieve a specific daily field by selecting the corresponding `time` index in the monthly NetCDF file. The coordinate variables

`lat` and `lon` define the spatial grid of each product.

The main data variables are `XCO` and `XCH4`, representing the fused column-averaged dry-air mole fractions of CO and CH_4, respectively. In addition, an `obs_mask` variable is provided to indicate whether the value at each grid cell and time step is directly constrained by a valid TROPOMI retrieval or reconstructed by the fusion framework. The `obs_mask` variable is a dimensionless binary mask, where `obs_mask = 1` indicates that the value at that grid cell and day is constrained by a valid satellite observation, and `obs_mask = 0` indicates that the value was generated by the proposed reconstruction method.

An `observation_fraction` variable is also included in each monthly file. This variable is calculated as the mean of `obs_mask` over the `time` dimension and represents the fraction of days in that month for which valid TROPOMI observations are available at each grid cell. For example, `observation_fraction = 1` indicates that valid satellite observations are available for all days in the month at that grid cell, whereas `observation_fraction = 0` indicates that all daily values at that grid cell are reconstructed by the fusion framework. This information is provided to prevent users from interpreting the entire gap-free product as purely observation-based.

Because of file-size and repository upload limitations, the daily data are organized into monthly NetCDF files rather than being provided as a single file for the entire study period. This monthly organization does not indicate temporal averaging. Each monthly file contains all daily fields for that calendar month.

The revised NetCDF files include CF-style metadata, including coordinate attributes, variable long names, units, missing-value definitions, temporal coverage, spatial resolution, product region, data source, author, institution, and contact information. The main variables included in the released NetCDF files are summarized as follows: `XCO` (ppb), fused column-averaged dry-air mole fraction of carbon monoxide; `XCH4` (ppb), fused column-averaged dry-air mole fraction of methane; `obs_mask` (unitless binary mask), indicating whether the grid-cell value is derived from a valid satellite observation (`1`) or reconstructed by the proposed fusion method (`0`); `observation_fraction` (unitless fraction from 0 to 1), indicating the monthly fraction of days with valid satellite observations at each grid cell; `time`, the daily time coordinate within each monthly file, defined using CF-style units; and `lat` and `lon`, the latitude and longitude coordinates of the gridded product.

Comment 6:

Physical consistency

A few aspects of the satellite–model comparison are not discussed:

(1) The role of averaging kernels (AKs) is not mentioned, which may affect comparisons;

(2) The difference in resolution between GEOS-Chem and the final product is not evaluated.

These points could have implications for the interpretation of the results.

Response 6:

Thank you for raising this important concern. We agree that both averaging kernels (AKs) and the spatial-resolution mismatch between GEOS-Chem simulations and the final fused products may affect the interpretation of satellite–model comparisons and reconstruction fidelity. To address these two issues, we added two sensitivity analyses in the revised manuscript.

1. We added a sensitivity analysis of the TROPOMI CO AKs. Because satellite retrievals and model profiles differ in their vertical sensitivity, representativeness errors may be introduced if AKs are not considered when comparing satellite data with model results. In this test, both the original GEOS-Chem XCO and the AK-smoothed GEOS-Chem XCO were compared with TROPOMI XCO in ppb. The results show that applying the TROPOMI CO column averaging kernels slightly increased the R^2 from 0.543 to 0.565, while the RMSE increased from 28.17 ppb to 28.66 ppb and the mean bias changed from −17.09 ppb to −18.34 ppb. These results indicate that AK treatment has a limited influence on the column-concentration-based fusion results, although AK-related effects remain an uncertainty source for applications requiring strict profile-sensitive validation. The detailed statistics are provided in Table S11.

2. To evaluate the influence of the spatial-resolution mismatch, we added a one-year resolution-sensitivity experiment using the 2019 CO data. In this experiment, we used GEOS-Chem simulation data at the native 2° × 2.5° resolution and aggregated TROPOMI XCO to the 2° × 2.5° resolution using the area-weighted averaging method. The fusion calculation framework was directly implemented on the 2° × 2.5° grid. The obtained fused XCO data at the 2° × 2.5° resolution were then compared with the fused product obtained by aggregating the fused XCO data at the 0.25° resolution to the same 2° × 2.5° grid using the area-weighted averaging method.

Figure S12 shows that the fused XCO data generated directly at the native 2° × 2.5° resolution and the 0.25° fused XCO data aggregated to the same 2° × 2.5° grid present relatively similar global spatial distributions. The quantitative comparison in Figure S13 also shows high consistency, with R^2 = 0.926, RMSE = 6.93 ppb, and mean bias = 0.15 ppb. These results indicate that the 0.25° fused XCO data preserve the coarse-scale spatial structure and that the resolution mismatch does not have a substantial influence on the coarse-scale reconstruction fidelity of XCO.

We also clarified that fine-scale structures below the native GEOS-Chem resolution should still be interpreted with caution, because they are constrained by TROPOMI spatial patterns, the DCT/SVD-reconstructed field, and residual learning rather than being directly resolved by GEOS-Chem. The corresponding descriptions have been added to Sections 2.2.1 and 2.2.2 and to the “Discussion of uncertainties and limitations” section. The detailed results are provided in Table S11, Figure S12, and Figure S13.

Revised text related to AKs:

2.2.2 Averaging-kernel sensitivity test

We conducted an additional sensitivity test to examine the influence of TROPOMI averaging kernels (AKs) on the comparison between satellite retrievals and model simulations. For 2019, quality-screened TROPOMI CO Level-2 retrievals (qa_value > 0.5) were collocated with GEOS-Chem CO vertical profiles. Two GEOS-Chem XCO datasets were then generated for comparison. For the first dataset, GEOS-Chem XCO was diagnosed from the simulated CO

profiles by vertically averaging the dry-air mole fractions using pressure-layer thickness as weighting, consistent with the standard post-processing procedure described above. For the second dataset, the GEOS-Chem CO profiles were first interpolated to the TROPOMI pressure layers and then smoothed using the TROPOMI CO column averaging kernels. Finally, both the original GEOS-Chem XCO and the AK-smoothed GEOS-Chem XCO were compared with TROPOMI XCO in ppb. (Lines 270–280)

AKs are an important source of uncertainty in comparisons between satellite retrievals and model simulations because they describe the vertical sensitivity of TROPOMI retrievals. To evaluate this effect, we conducted an AK sensitivity test using the 2019 CO data by comparing GEOS-Chem XCO with TROPOMI XCO before and after applying the TROPOMI CO column averaging kernels. As shown in Table S11, applying the TROPOMI CO column averaging kernels slightly increased the R^2 from 0.543 to 0.565, while the RMSE increased from 28.17 ppb to 28.66 ppb and the mean bias changed from −17.09 ppb to −18.34 ppb. These results indicate that AK treatment has a limited influence on the column-concentration-based fusion results in this experiment. (Lines 800–808)

Table S11. Overall statistics of the TROPOMI CO averaging-kernel sensitivity test in 2019, with RMSE and μ in parts per billion (ppb).

Comparison	R^2	RMSE	μ
Original GEOS-Chem vs. TROPOMI	0.543	28.17	-17.09
AK-smoothed GEOS-Chem vs. TROPOMI	0.565	28.66	-18.34

Revised text related to the evaluation of resolution differences:

To evaluate the influence of the resolution mismatch between GEOS-Chem simulation results and the final fused product, we used the 2019 CO data to conduct a one-year resolution sensitivity experiment. In this experiment, we used GEOS-Chem simulation data at the native 2° × 2.5° resolution and aggregated TROPOMI XCO to the 2° × 2.5° resolution using the area-weighted averaging method. The related fusion calculation framework was directly implemented on the 2° × 2.5° grid; subsequently, the obtained fused XCO data at the 2° × 2.5° resolution were compared with the fused product obtained by aggregating the fused XCO data at the 0.25° resolution to the same 2° × 2.5° grid using the area-weighted averaging method. (Lines 262–269)

Figure S12 shows that the fused XCO data generated directly at the native 2° × 2.5° resolution (Fig. S12a) and the 0.25° fused XCO data aggregated to the same 2° × 2.5° grid (Fig. S12b) present relatively similar global spatial distributions, with relatively high XCO over East Asia, central Africa, and parts of South America. The quantitative comparison results in Figure S13 also show high consistency, with R^2 = 0.926, RMSE = 6.93 ppb, and mean bias = 0.15 ppb. These results indicate that the 0.25° fused XCO data preserve the coarse-scale spatial structure, and that the resolution mismatch does not have a substantial influence on the coarse-scale reconstruction fidelity of XCO. (Lines 820–827)

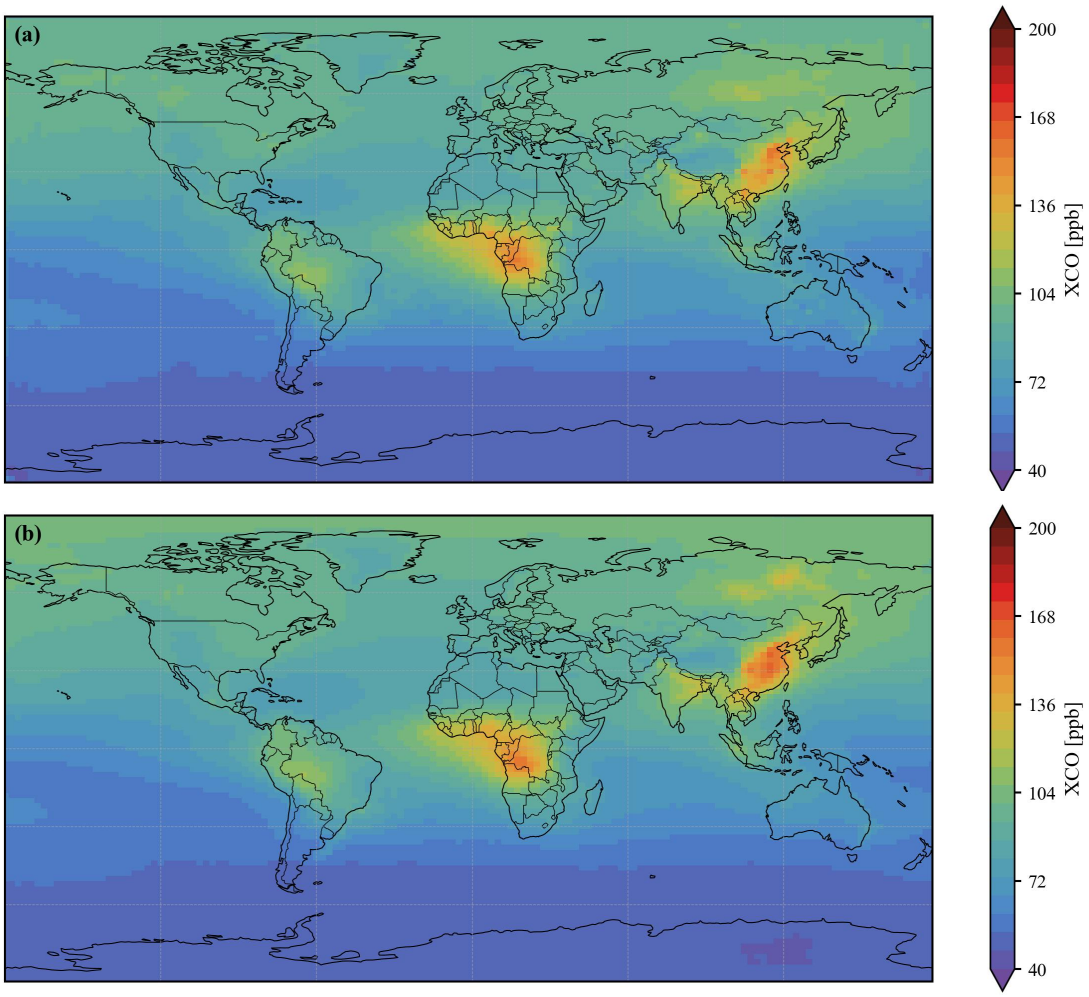


Figure S12. Spatial comparison of the 2019 XCO resolution-sensitivity experiment. (a) Fused XCO data generated directly at the native 2° × 2.5° GEOS-Chem resolution. (b) Fused XCO data obtained by aggregating the 0.25° resolution fused XCO data to the same 2° × 2.5° grid using the area-weighted averaging method.

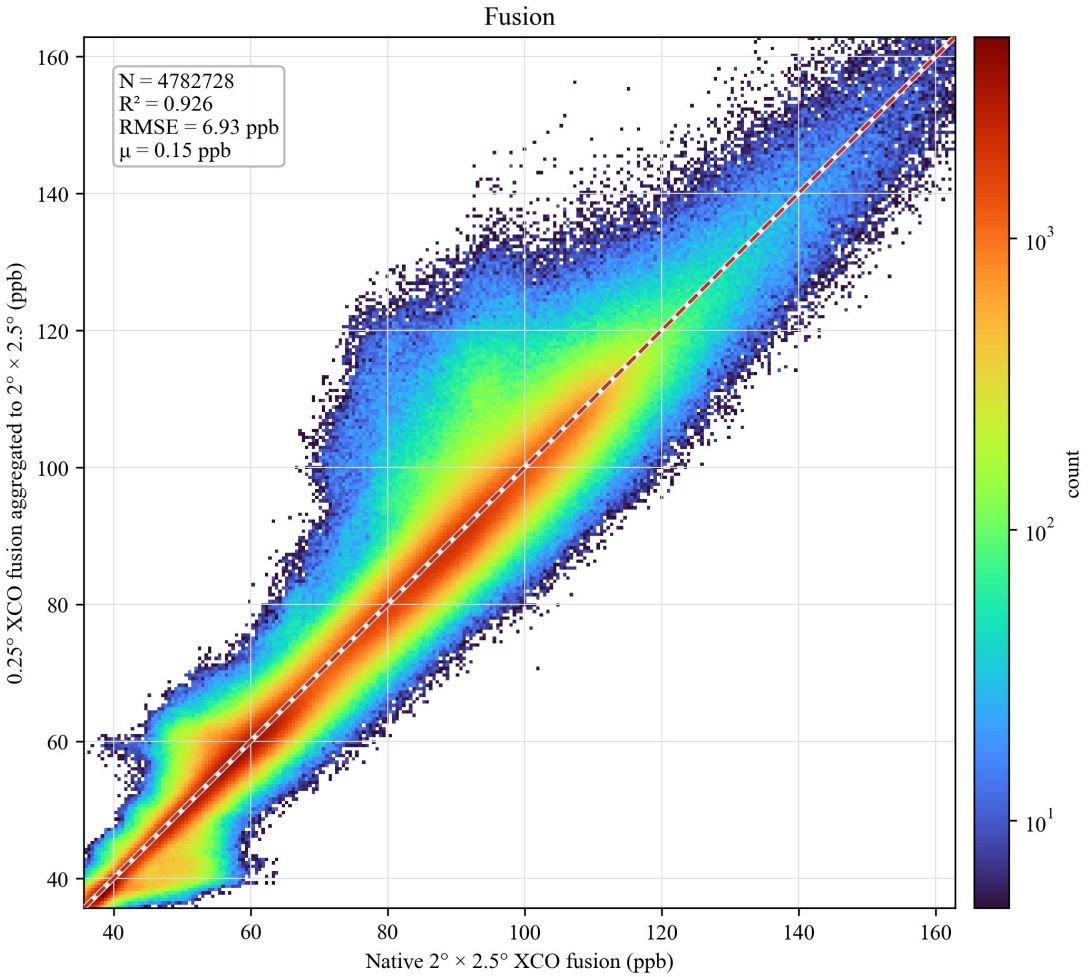


Figure S13. Quantitative comparison of the 2019 XCO resolution-sensitivity experiment.

Minor comments

Comment 7: *I think, the introduction could better explain why CO and CH_4 are studied together, rather than treating them separately.*

Response 7:

Thank you for this valuable suggestion. We have substantially revised the Introduction to better clarify the scientific rationale for the simultaneous reconstruction of XCO and XCH_4. Specifically, we added a new paragraph emphasizing the chemical and dynamical coupling between CO and CH_4 through shared oxidation pathways, common emission sources, and atmospheric transport processes. The relevant text has been added to Lines 57–72 of the revised manuscript.

Revised text:

Although CO and CH_4 are often discussed separately in terms of their sources and impacts, they are chemically and dynamically coupled in the atmosphere. Both gases are predominantly removed by the hydroxyl radical (OH), meaning that changes in one can affect the atmospheric lifetime and concentration of the other through OH competition (Lelieveld et al., 2016). Moreover,

they share major emission sources, particularly biomass burning and fossil-fuel-related activities such as incomplete combustion and flaring, where they can be emitted simultaneously (Andreae and Merlet, 2001; Akagi et al., 2011; Andreae, 2019). Their co-emission and co-transport imply that the spatial patterns of CO and CH_4 may be correlated near fire hotspots, industrial regions, and other combustion-influenced areas. Therefore, joint reconstruction of XCO and XCH_4 offers two key advantages: (1) shared spatial structure and common meteorological drivers can provide complementary constraints for deep learning models, improving generalization in data-sparse regions; (2) simultaneous monitoring can support more consistent estimation of emission ratios and source attribution, as demonstrated by studies using CH_4/CO or related trace-gas ratios to constrain biomass-burning and fossil-fuel methane emissions (Worden et al., 2013; Worden et al., 2017; Ramsden et al., 2022). A comprehensive understanding of these coupled gases is fundamental to improving air quality predictions and formulating climate mitigation policies.

Comment 8: *There are several language issues and repeated sentences; overall readability could be improved.*

Response 8:

We sincerely apologize for the language issues and redundancies in the original manuscript, and we thank you for pointing this out. We fully agree that the readability needed significant improvement. To address this issue, we carefully reviewed the entire manuscript and removed repeated sentences and overlapping explanations. For instance, we streamlined the Introduction, particularly the description of satellite instruments, and the Methodology section, particularly the GEOS-Chem setup and site-selection description, to ensure a more concise and logical flow. In addition, the entire manuscript has undergone a thorough language revision, including improvements to grammar, sentence structure, terminology, and overall clarity.

Comment 9: *Some terminology is inconsistent (e.g. quality flag vs. quality value).*

Response 9:

We thank you for pointing out the inconsistency in our terminology. Following the official TROPOMI product documentation and to ensure technical precision, we have standardized the terminology throughout the manuscript to “quality value (qa_value)”, which represents the quality assurance value. We have replaced all previous instances of "Quality Value" and "Quality Flag" with this standard term. The relevant revisions have been made in the revised manuscript.

Revised text:

(1) we applied product-specific quality screening and retained only TROPOMI retrievals with a quality value (qa_value) greater than 0.5 for subsequent analysis. (Lines 164–166)

(2) TROPOMI XCO and XCH_4 retrievals with qa_value below 0.5 were excluded. (Lines 255–256)

(3) where N_{total} is the total number of 0.25° grid cells within the 2° radius and N_{valid} is the number of grid cells with valid TROPOMI retrievals after applying the qa_value > 0.5 criterion.

Grid cells without satellite retrievals or with qa_value ≤ 0.5 were treated as missing. We selected site-day observation cases with MR > 0.5, corresponding to less than 50% valid TROPOMI coverage within the collocation window, to evaluate the performance of the fused XCO and XCH_4 datasets under sparse TROPOMI coverage conditions. (Lines 508–513)

(4) For each gas, we first screened TCCON stations that retained a relatively large number of valid spatially collocated data pairs after applying the 13:30 ± 1 h temporal window, the 2° spatial matching radius, and the TROPOMI quality filter (qa_value > 0.5). (Lines 543–546)

Comment 10: *A few figures could potentially be simplified or moved to supplementary material.*

Response 10:

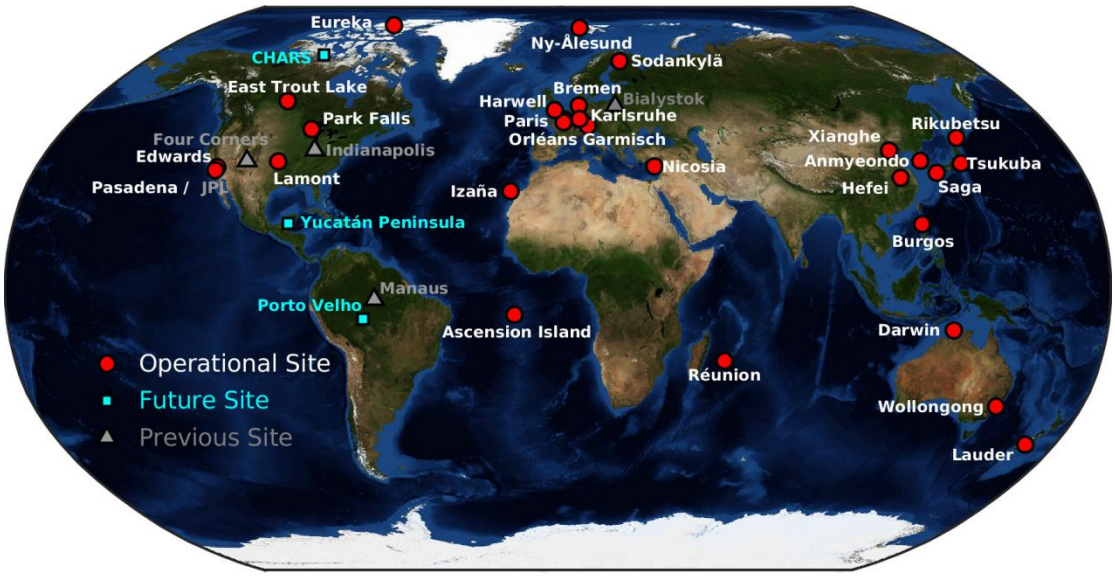
We thank you for the suggestion that some figures could be simplified or moved to the Supplementary Information. We have carefully reviewed all figures to enhance the focus and readability of the main manuscript. Specifically, we have already moved Figure 1 (global distribution of TCCON sites) and Table 1 (site details) to the Supplementary Information (now Figure S1 and Table S1). As a result, the figures in the main text now focus primarily on the methodological framework, core validation results, and key spatiotemporal patterns. We believe these revisions substantially improve the clarity and overall readability of the manuscript.

Revised Table and Figure:

Table S1. Details of the TCCON sites used in this study, No. is the serial number.

No.	Site name	Latitude	Longitude	Location	Start date	End date
1	bremen01	53.104	8.85	Bremen, Germany	2009/1/6	2021/6/24
2	burgos01	18.5325	120.6496	Burgos, Philippines	2017/3/3	2023/6/23
3	darwin01	-12.425	130.891	Darwin, Australia	2013/1/1	2022/12/27
4	easttroutlake01	54.354	-104.987	East Trout Lake, Canada	2016/10/3	2024/2/15
5	edwards01	34.9599	-117.881	AFRC, Edwards, CA, USA	2013/7/20	2024/2/22
6	garmisch01	47.476	11.063	Garmisch, Germany	2007/7/18	2023/5/4
7	hefei01	31.91	117.17	Hefei, China	2015/11/2	2023/12/25
8	izana01	28.31	-16.5	Izana, Tenerife, Spain	2014/1/2	2023/8/30
9	karlsruhe01	49.103	8.44	Karlsruhe, Germany	2014/1/15	2023/6/26
10	lamont01	36.604	-97.486	Lamont, Oklahoma, USA	2011/4/16	2024/2/25
11	lauder03	-45.038	169.684	Lauder, New Zealand	2018/10/2	2023/12/28
12	nicosia01	35.141	33.381	Nicosia, Cyprus	2019/9/1	2023/5/10
13	orleans01	47.97	2.113	Orleans, France	2009/9/6	2023/6/23
14	parkfalls01	45.945	-90.273	Park Falls, Wisconsin, USA	2004/6/2	2024/2/25
15	pasadena01	34.136	-118.127	Pasadena, California, USA	2012/9/20	2024/2/25
16	reunion01	-20.901	55.485	Reunion Island, France	2015/3/1	2020/7/18
17	rikubetsu01	43.4567	143.7661	Rikubetsu, Hokkaido, Japan	2014/6/24	2021/6/30
18	sodankyla01	67.367	26.631	Sodankylä, Finland	2009/5/16	2023/5/30
19	tsukuba02	36.0513	140.1215	Tsukuba, Ibaraki, Japan, 125HR	2014/3/28	2021/3/31
20	wollongong01	-34.406	150.891	Wollongong, Australia	2013/1/4	2023/6/27
21	xianghe01	39.8	116.96	Xianghe, China	2018/6/14	2023/5/29

Figure S1. Global distribution of TCCON sites, including operational sites, future sites, and previous sites, respectively.



Comment 11: *Acronyms should be defined once and used consistently.*

Response 11:

We apologize for the inconsistent use of acronyms in the original manuscript. We have reviewed the manuscript to improve acronym usage and reduce redundant expansions. Acronyms such as TROPOMI, DCT, SVD, and TCCON are defined at their first mention in the Abstract and again at their first mention in the main text, where appropriate. After definition, the abbreviated forms are used throughout the remainder of the manuscript. Some examples of the revisions are provided below, appearing in Line 346 and Line 351 of the revised manuscript.

Examples of revisions:

"Three-dimensional discrete cosine signal transform (DCT_3) rule," was revised to "The 3D DCT, denoted as DCT_3, is defined as:", and "Three-dimensional discrete cosine inverse transform ($IDCT_3$) rule," was revised to "The inverse 3D DCT, denoted as $IDCT_3$, is defined as:".