

Reviewer 1:

The authors reconstructed two-decade global daily high-resolution XCO₂ data based on a hybrid Transformer–BiLSTM model. However, the topic is usual, and the method should be more innovative in this study, which requires large improvements. Specific comments are as follows:

Major comments:

1. There are many studies providing global daily high-resolution XCO₂ for a long term, such as DOI: 10.1016/j.eiar.2025.108146 with better spatial resolution (1 km) than yours (2003–2023, $R^2 = 0.988$, RMSE = 1.10 ppm against TCCON) [1]. What are your strengths compared to them? What is it that your dataset can present but others cannot? Please state in detail.

Response: Thank you for raising this important question. We acknowledge the previous related studies, including Wang (2026), which provides a 1 km global XCO₂ product with strong validation performance. However, our study focuses on constructing a temporally seamless, cross-mission-consistent, and physically coherent daily global XCO₂ dataset spanning two decades (2003–2022), which is particularly important for long-term carbon-cycle and climate analyses, rather than simply pursuing the highest spatial resolution.

The primary innovation of our study lies in the explicit treatment of inter-satellite inconsistencies among SCIAMACHY, GOSAT, and OCO-2. Previous long-term reconstruction studies generally fuse multiple satellite products directly, but often do not sufficiently address systematic biases caused by differences in sensor characteristics, orbital sampling, retrieval algorithms, and mission transitions. These inconsistencies can introduce artificial discontinuities and temporal drifts that compromise long-term trend analyses. To address this issue, we developed a TCCON-guided bias-correction framework that harmonizes observations across different satellite missions and minimizes artificial step changes during the SCIAMACHY–GOSAT and GOSAT–OCO-2 transition periods (Fig. 5). Compared with the uncorrected data-fused product, the bias-corrected XCO₂ dataset shows improved accuracy (sample-based CV-RMSE reduced from 1.10 ppm to 1.03 ppm; spatial CV-RMSE further reduced to 0.97 ppm) and substantially enhanced temporal consistency. This demonstrates that cross-mission harmonization is a critical component for generating reliable long-term XCO₂ records.

Another key strength of our study is the emphasis on temporal continuity and daily dynamics. While previous studies mainly focus on improving spatial resolution, our hybrid Transformer–BiLSTM framework is specifically designed to capture both long-range spatial dependencies and temporal evolution. The BiLSTM module extracts bidirectional temporal features to preserve daily continuity, while the Transformer module leverages self-attention to characterize non-local spatial relationships. In addition, we introduced a weighted spatiotemporal loss function to jointly constrain point-wise accuracy, temporal smoothness, and spatial coherence. These designs enable the reconstruction of gap-free daily XCO₂ fields and improve the representation of

temporal variability associated with atmospheric transport, biospheric exchange, and anthropogenic emissions.

Importantly, our study also places stronger emphasis on physical interpretability. In addition to satellite XCO₂ observations, we incorporate multiple physically relevant predictors, including CAMS XCO₂, meteorological variables, surface variables, and emission-related precursor gases such as NO₂, CO, and XCH₄. These variables help characterize atmospheric circulation, fossil-fuel combustion, biomass burning, and biospheric activity, thereby improving the physical realism of the reconstructed XCO₂ fields. SHAP analysis further confirms the meaningful contribution of these predictors to the reconstruction process. We believe this is an important advantage because high spatial resolution alone cannot compensate for missing physical constraints or temporal inconsistencies.

Furthermore, our validation framework is more comprehensive for assessing long-term robustness. In addition to evaluation against TCCON observations, we also validate the dataset using 41 independent ObsPack stations distributed globally. These independent evaluations demonstrate that the reconstructed dataset maintains strong spatial transferability and temporal stability across diverse regions and atmospheric conditions.

Therefore, the novelty of our study is not simply the generation of another high-resolution XCO₂ dataset, but the development of a temporally seamless, physically consistent, and cross-satellite-harmonized daily global XCO₂ record suitable for investigating long-term carbon-cycle dynamics, interannual variability, and climate-related changes. In this sense, our dataset is complementary to previous ultra-high-resolution products (Wang, 2026) are optimized for fine-scale spatial mapping, whereas our product is specifically designed for long-term temporal analysis and climate applications requiring stable cross-mission continuity.

We have added a new section titled “Strengths” to the revised manuscript to better highlight these innovations and clarify the distinct scientific contribution of our study.

2. The BiLSTM and attention model have been applied to estimate global XCO₂, such as DOI: 10.5194/essd-17-5355-2025. It seems that your model structure presenting similarity to this paper [2]. Please justify your innovation.

Response: Thank you for this important comment. We acknowledge that BiLSTM and attention-based models have previously been applied to global XCO₂ reconstruction, such as Wang et al. (2025). However, our framework differs substantially in both model architecture and reconstruction strategy.

The key innovation of our study lies in the synergistic integration of Transformer and BiLSTM modules to jointly model long-range spatial dependencies and temporal continuity. Specifically, the Transformer encoder utilizes self-attention mechanisms to capture non-local spatial relationships and global contextual information, which are

critical for representing large-scale atmospheric transport and spatially heterogeneous carbon dynamics. In contrast, the BiLSTM module focuses on bidirectional temporal evolution, enabling the model to better characterize daily continuity, temporal autocorrelation, and persistent XCO₂ variability. Compared with conventional BiLSTM-attention frameworks (Wang et al., 2025), which mainly enhance sequential learning using local attention operations, our Transformer–BiLSTM architecture provides a substantially larger receptive field and stronger capability for jointly learning global spatiotemporal dependencies. In addition, we further introduce a weighted spatiotemporal loss function that jointly constrains reconstruction accuracy, temporal smoothness, and spatial coherence, thereby improving the temporal consistency and physical continuity of reconstructed daily XCO₂ fields. This differs from previous studies that mainly optimize point-wise reconstruction errors using standard loss functions such as MSE.

Another important difference is that our framework is specifically designed for long-term multi-mission reconstruction. We propose a “data fusion + bias correction” workflow with explicit TCCON-guided bias correction to harmonize systematic discrepancies among SCIAMACHY, GOSAT, and OCO-2. Unlike previous studies that mainly focus on single-mission reconstruction performance, our framework explicitly addresses cross-mission inconsistencies and minimizes artificial temporal discontinuities caused by sensor replacement and orbital differences.

Therefore, while previous studies have explored BiLSTM or attention-based approaches (Wang et al., 2025), the novelty of our study lies in the development of a Transformer–BiLSTM framework specifically designed for long-term, multi-mission, temporally seamless, and physically consistent daily XCO₂ reconstruction.

We have added a new section titled “Innovations of the framework” to the revised manuscript to better clarify these methodological innovations and their differences from previous work.

Minor comments:

Line 180: Please add some ablation experiments of your model.

Response: In the revised manuscript, we have conducted an architecture-level ablation analysis comparing three backbone variants: Bi-LSTM, Transformer, and Transformer-BiLSTM (Table 2). The hybrid Transformer-BiLSTM model achieved the best overall performance, demonstrating the advantage of jointly modeling long-range spatial dependencies and temporal continuity.

Table 2. Comparison of performance across various other models for retrieving XCO₂

Model	CV-R ²	RMSE	MAE
Transformer	0.78	1.78	1.55
Bi-LSTM	0.82	1.42	1.36
Transformer-BiLSTM	0.85	1.11	0.82

Figure 6: A station presents very low R^2 value (< 0.1). Please add some discussions.

Response: The low R^2 at the AZR station is mainly associated with the limited number of matched observations and relatively small temporal variability at this site, which can lead to unstable correlation statistics. Importantly, the RMSE and bias at this station remain relatively small, suggesting that the low R^2 does not indicate a systematic deficiency of the reconstructed dataset. The corresponding discussion has been added to the revised manuscript.

Figure 8: The selected scenarios are only located on one or two pixels. I think the high values also may be some noises.

Response: These selected examples are intended to illustrate localized elevated XCO_2 signals rather than strict point-source detections; however, some isolated high-value pixels may partly reflect retrieval uncertainty or noise, especially in regions with limited observations. We have added this discussion to the revised manuscript.

Figure 14: I notice the mean SHAP value of NO_2 is the smallest. How to interpret this result?

Response: The relatively low mean SHAP value of NO_2 suggests that its overall contribution to the global XCO_2 reconstruction is smaller compared with variables such as CAMS XCO_2 and meteorological factors. This is expected because NO_2 mainly reflects localized anthropogenic combustion emissions, whereas XCO_2 variability at the global scale is more strongly controlled by large-scale atmospheric transport, background concentration fields, and biospheric exchange processes. Nevertheless, NO_2 still provides useful complementary information for identifying regional anthropogenic emission patterns, particularly over urban and industrial areas. We have added this clarification to the revised manuscript.

Table 2: Some important works are missing, such as [1] and [3]-[5].

Response: We have revised Table 3 by adding the missing studies suggested by the reviewer, including Wang (2026) and other recent global XCO_2 reconstruction studies (Li et al., 2026; Hwang et al., 2026; Yu et al., 2026). We also updated the discussion section to better compare our dataset with these previous products and clarify the advantages of our framework in temporal seamlessness, cross-mission harmonization, and long-term daily XCO_2 reconstruction.

References:

[1] Wang, J. (2026). Global daily 1 km gapless XCO_2 (2003– 2023) derived from multi-satellite observations and a spatiotemporal deep learning framework. *Environmental Impact Assessment Review*, 117, 108146.

[2] Wang, Z., Zhang, C., Shi, K., Shangguan, Y., Hu, B., Chen, X., ... & Zhang, Q. (2025). A full-coverage satellite-based global atmospheric CO_2 dataset at 0.05°

resolution from 2015 to 2021 for exploring global carbon dynamics. *Earth System Science Data*, 17(10), 5355-5375.

[3] Li, J., Zhang, Z., Li, T., Yuan, Q., & Zhang, L. (2026). Global daily seamless XCO₂ Mapping (2016–2020): Spatio-temporal trends and variations during wildfire events. *International Journal of Applied Earth Observation and Geoinformation*, 146, 105092.

[4] Hwang, S., Choi, H., Kang, Y., & Im, J. (2026). Reconstructing long-term (2003–2019) global high-resolution XCO₂: bridging observational gaps with machine learning. *GIScience & Remote Sensing*, 63(1), 2627042.

[5] Yu, Y., Tian, W., Zhang, L., Yu, T., Wu, Y., & Cheng, T. (2026). MCF-XCO₂: A cross-mission consistency and fusion framework for integrating multi-satellite XCO₂ observations. *Atmospheric Research*, 108747.