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 XCO_2 product with strong validation performance. However, our study focuses on constructing a temporally seamless, cross-mission-consistent, and physically coherent daily global XCO_2 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 XCO_2 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 XCO_2 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 XCO_2 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 XCO_2 observations, we incorporate multiple physically relevant predictors, including CAMS XCO_2, meteorological variables, surface variables, and emission-related precursor gases such as NO_2, CO, and XCH_4. These variables help characterize atmospheric circulation, fossil-fuel combustion, biomass burning, and biospheric activity, thereby improving the physical realism of the reconstructed XCO_2 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 XCO_2 dataset, but the development of a temporally seamless, physically consistent, and cross-satellite-harmonized daily global XCO_2 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 XCO_2 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 XCO_2 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 XCO_2 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 XCO_2 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 XCO_2

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 NO2 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.

Reviewer 2:

This manuscript presents a timely and potentially valuable high-resolution daily XCO_2 dataset by combining multi-mission satellite retrievals with a hybrid Transformer–BiLSTM framework, followed by a bias-correction step. The topic is well suited to ESSD, and the dataset has clear potential applications in carbon-cycle analysis, emission monitoring, and climate studies. Overall, the manuscript is well written, and I have several suggestions to further improve the study.

Specific comments:

(1) The spatial scope of the dataset should be described more consistently. The title and several sections of the manuscript give the impression of a fully global XCO_2 product, whereas the abstract states that the dataset characterizes XCO_2 "over global land surfaces." This scope should be made consistent across the title, abstract, main text, and dataset description.

Response: We have revised the manuscript to consistently describe the dataset as a global land XCO_2 product. The title and the corresponding descriptions in the Abstract, Introduction, Conclusions, and data description have been updated accordingly.

(2) The description of the Bi-LSTM architecture needs clarification. The statement that the model "consists of 64 hidden layers with a dimension of 128" is difficult to interpret and seems unlikely in its current form. It would be helpful to clarify whether this refers to hidden units, hidden states, or stacked layers.

Response: We have revised the description of the Bi-LSTM architecture to: "The temporal encoder was implemented as a one-layer bidirectional LSTM (Bi-LSTM) with 64 hidden units in each direction, resulting in a 128-dimensional output feature." .

(3) To enhance readability, Figure 5 should include a corresponding legend.

Response: We have added a corresponding legend to Figure 5 in the revised manuscript to improve its clarity and readability.

(4) Several language and wording issues should be refined. For example, "till 2012" should be revised to "until 2012," and "Electric power plant in Algerian" should be corrected to "in Algeria." A careful language edit would improve overall readability.

Response: We have corrected the noted language issues and have carefully proofread the entire manuscript to improve overall readability.

(5) The ENSO discussion contains a minor numerical inconsistency. A reported correlation of R = 0.55 corresponds to approximately 30% explained variance rather than ~34%. Please check the correlation coefficient or revise the corresponding statement accordingly.

Response: The explained variance has been corrected from 34% to 30% to accurately reflect the reported correlation (R = 0.55).

(6) The performance metrics reported in the abstract and main text would benefit from clearer interpretation. Different statistics are presented for independent TCCON validation, bias-corrected results, and various cross-validation strategies, but the abstract may give the impression that all values correspond to a single evaluation setting. It would be helpful to explicitly indicate which metric corresponds to which validation framework.

Response: To improve clarity, we have revised the abstract to explicitly distinguish between different validation frameworks as follows:

"Independent validation of the data-fused XCO_2 product against Total Carbon Column Observing Network (TCCON) measurements shows excellent agreement, with a correlation coefficient (R) of 0.99, a root mean square error (RMSE) of 1.10 ppm, and a mean bias of 0.01 ppm. After bias correction, the product further improves cross-satellite consistency, achieving an R^2 of 0.99 and an RMSE of 0.36 ppm in the sample-based ten-fold cross-validation."

(7) In the caption for Figure 7, the years are listed as "May 30, 2009, 2013, 2020." For grammatical correctness, this should be revised to "May 30, 2009, 2013, and 2020."

Response: We have corrected the figure caption as suggested.

(8)The seasonal averages should be presented in a consistent format, for example (MAM; average = 395.36 ± 2.20 ppm).

Response: We have standardized all seasonal-average expressions to a consistent format throughout the manuscript as suggested.

(9) The terminology used for the reconstructed product should be standardized. Terms such as "seamless," "gap-free," and "gapless" are used interchangeably; selecting one consistent term would improve clarity.

Response: We have revised the manuscript to use "seamless" consistently throughout when referring to the reconstructed product.

(10) Some minor grammatical issues should also be corrected, for example "Sample-based CV show" should be revised to "Sample-based CV shows."

Response: We have corrected this grammatical error and carefully revised similar minor grammatical issues throughout the manuscript.

(11) The model name should be used consistently throughout the manuscript. In most sections it is referred to as "Transformer–BiLSTM," whereas in the XAI section it appears as "4D-STransformer-BiLSTM," which may confuse readers.

Response: We have unified all model references to "Transformer-BiLSTM" throughout the manuscript.

(12) It would also be helpful to clarify whether coarser-resolution products were regridded prior to comparison in Figure 7.

Response: We have clarified in the manuscript that no regridding was applied to Figure 7, as the coarser-resolution products are plotted at their native resolutions for qualitative comparison.

(13) The interpretation of localized XCO_2 hotspots should be phrased more cautiously. Some descriptions appear overly definitive for a column-integrated quantity that can also be influenced by atmospheric transport and meteorological conditions. Expressions such as "consistent with localized enhancements associated with…" would be more appropriate unless supported by stronger independent evidence.

Response: We have revised this sentence as "These patterns are consistent with localized regions of relatively high XCO_2 concentrations associated with fossil-fuel combustion and extraction-related activities."

(14) The formatting of several references should be carefully checked, as there are minor issues such as missing spaces or compressed publisher/year formatting.

Response: We have carefully checked and corrected the formatting of all references throughout the manuscript, including missing spaces and compressed publisher/year formatting issues.

(15) In Table 2, it would be beneficial to include comparisons with the most recent relevant studies to better position the dataset within the current literature.

Response: We have expanded this table to include several recent global XCO_2 studies, and the associated text in Section 3.3.2 has been revised accordingly. These additions better position our dataset within the context of the current literature.

(16) Finally, the names of datasets and networks should be presented consistently. For example, "ObsPack" and "Obspack" appear in different places and should be unified.

Response: We have standardized all dataset and network names throughout the manuscript, correcting "Obspack" to "ObsPack" as suggested.