

Response to Anonymous Referee #2:

This study reconstructed a global full-coverage XCO₂ product with a 0.05° spatial resolution using multi-component satellite data and an advanced deep learning method. However, the manuscript lacks innovation and sufficient detail in several aspects. My comments are listed below:

Response: We are truly grateful for reviewing our manuscript and providing us with constructive feedback. Considering your feedback, we have elaborated on the innovation of our study and provided additional details in section **2. materials and methods**, **3. Results** and **4. Discussion**. We responded to your comments point by point as below.

Q1. (1) The manuscript notes that the spatial resolution of current global full-coverage XCO₂ products is relatively coarse, ranging from approximately 0.25° to 2° (Line 128). However, global XCO₂ products with a 0.1° spatial resolution already exist (<https://doi.org/10.1016/j.envint.2023.108057>), indicating a need for more comprehensive literature review.

Response: Thank you for pointing out this oversight and we apologize for omitting this important reference. We have provided a more tailed literature review in **1. introduction** as follows:

“The second category is regression-based methods, which aim to fill the gap by capturing the nonlinear relationship between multi-source XCO₂ measurements and related covariates (He et al., 2022; Siabi et al., 2019; Zhang and Liu, 2023). The specific methods include traditional statistical models, geostatistical models and machine learning models. Siabi et al. (2019) employed the Artificial Neural Network (ANN) to establish correlation between XCO₂ and eight environmental variables. Zhang and Liu (2023) utilized the convolution neural networks (CNN) coupled with attention mechanisms to produce full-coverage XCO₂ data across China. Recently, Zhang et al. (2023) developed high spatial resolution global CO₂ concentration data based on deep forest model and multi-source satellite products.

Although the development of CO₂ observation satellites and the application of machine learning methods have significantly improved the estimation accuracy of

XCO₂, current studies still face several limitations. Firstly, due to the sparse distribution of satellite XCO₂ data, previous studies always relied on assimilation and reanalysis XCO₂ data, such as CAMS XCO₂ with coarse spatial resolution (0.75°). This reliance often results in final products that closely mirror the assimilation and reanalysis results, leading to an oversmoothed distribution that undermines the high-resolution advantages of satellite data. Furthermore, most current studies estimated the spatial distribution of CO₂ primarily based on vegetation and meteorological information, with limited consideration of the impact of human activities and emissions, despite these have significant influence on atmospheric CO₂ variability. This limitation also led to estimation results that fail to adequately capture the impact of anthropogenic emissions on atmospheric CO₂. In addition, most studies that employ regression models to estimate full-coverage XCO₂ are limited to regional or national scales due to the weak transferability of these models. Only a few studies (Zheng et al., 2023) have explored global-scale CO₂ estimation using machine learning approaches, highlighting the need for further research to enhance model generalizability and scalability.”

(2) Although this study improves the XCO₂ spatial resolution to 0.05°, its innovation and advantages compared to other datasets remain unclear. It is recommended to clearly articulate the study’s novelty and specific strengths.

Response: Thanks for this constructive suggestion. We added a more detailed comparison with other datasets to highlight the innovation and advantages of our study in section 4.1 **Comparison with previous studies**, as follows:

“To validate the effectiveness of our model and resulting XCO₂ products, we compared our results with current studies which focuses on global XCO₂ reconstruction (Table 5). As for the in-situ validation, most existing studies report high accuracy with almost all R² over 0.9, RMSE less than 2 ppm. Regarding spatial resolution, the various products differ substantially, ranging from 1° down to 0.01°. It should be noted that increasing spatial resolution tends to compromise the accuracy of XCO₂ retrievals. However, our XCO₂ product achieves an optimal balance between spatial detail and measurement precision, exhibiting both high spatial resolution (0.05°) and robust accuracy (R²=0.91, RMSE =1.54 ppm) in comprehensive evaluations.

Table 5. Comparison between current studies focusing on global XCO₂ reconstruction

Model	Spatial resolution	In-situ validation (with TCCON)			Reference
		R ²	RMSE (ppm)	MAE (ppm)	
Attentional-based LSTM	0.05°	0.91	1.54	1.22	Our study
Deep forest	0.1°	0.96	1.01	-	Zhang et al. (2023)
S-STDCT	0.25°	0.95	1.18	-	Wang et al. (2023)
Spatiotemporal kriging	1°	0.97	1.13	0.88	Sheng et al. (2022)
MLE & OI	0.5°	0.92	2.62	1.53	Jin et al. (2022)
ERT	0.01°	0.83	1.79	-	Li et al. (2022)

*S-STDCT: Self-supervised spatiotemporal discrete cosine transform, MLE & OI: maximum likelihood estimation method and optimal interpolation; ERT: Extremely randomized trees

To evaluate the advancement of our XCO₂ product, we compared it with original OCO-2 observations and publicly available global XCO₂ datasets (Wang et al., 2023; Sheng et al., 2022; Zhang et al., 2023) across four regions: North America, Europe with northern Africa, Asia, and Oceania (Fig. 13) in January 2015. Despite monthly aggregation, OCO-2 data exhibit persistent spatial discontinuities, limiting the capacity to analyze monthly XCO₂ variability at regional and national scales. Existing XCO₂ products (spatial resolution of 0.25°, 1°, and 0.1°, respectively) broadly reproduce large-scale XCO₂ patterns but fail to resolve fine-scale heterogeneity. In comparison, our reconstructed XCO₂, with the highest spatial resolution, provides a more detailed and accurate representation of the regional XCO₂ patterns. For example, lower XCO₂ concentrations are clearly identified in eastern Canada (The first row of Fig.13) and Papua New Guinea (The fourth row of Fig. 13), regions characterized by dense forest cover. This correspondence highlights the substantial carbon sink potential of these forested areas. Our high-resolution product better identifies the CO₂ heterogeneity associated with different land cover types, whereas the coarse-resolution products smooth these signals. This limitation primarily stems from the neglect of high-resolution land cover dynamics and dependence on coarse-resolution assimilated/reanalysis datasets (e.g., CAMS XCO₂, CarbonTracker), resulting in oversmoothed spatial patterns that obscure satellite-derived high-resolution signals. Unlike assimilation-dependent approaches, our method avoids XCO₂ reanalysis inputs, preserving satellite-scale fidelity through high-resolution environmental variables modeling while maintaining precision.”

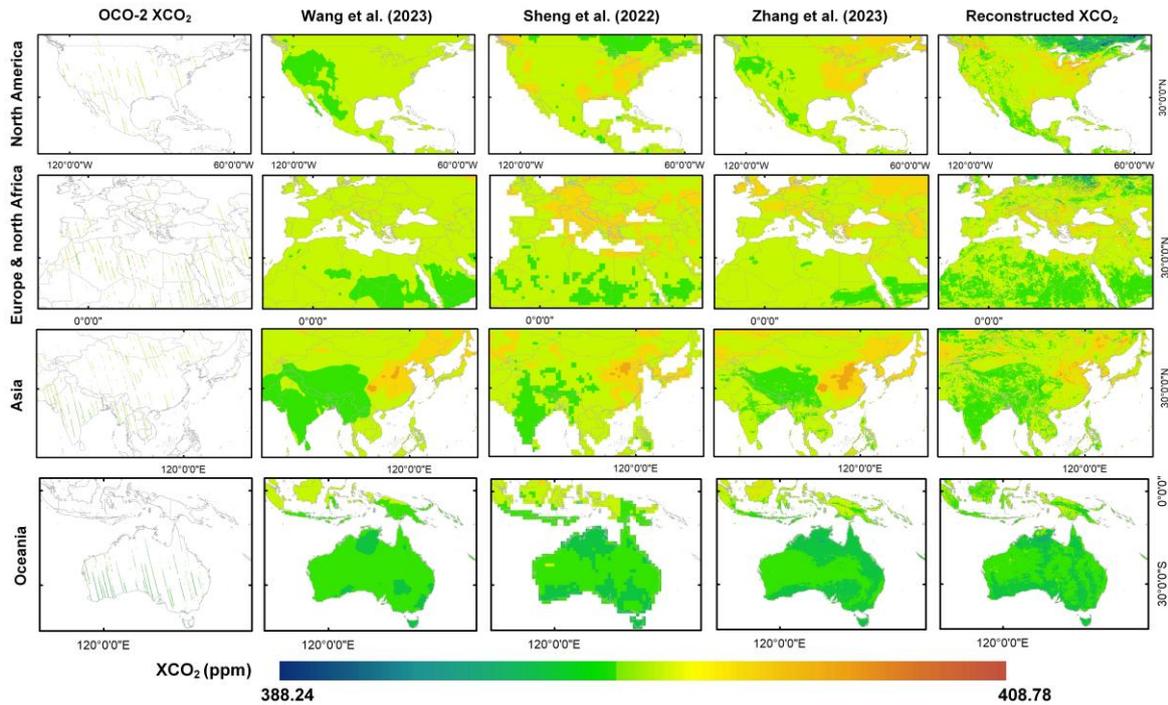


Figure 13. Comparison between the OCO-2 XCO₂ data, accessible XCO₂ products from Wang et al. (2023), Sheng et al. (2022), Zhang et al. (2023), and our reconstructed XCO₂ data in four regions, using the products of January of 2015 as an example.

And we also summarized the advantages and contributions of our study in **6. Conclusion** as follows:

“As a major driver of global warming, the monitoring of CO₂ changes, especially anthropogenic CO₂ emissions, is of critical importance. The launch of carbon satellites offers a significant advancement for CO₂ monitoring. However, the limited spatial coverage of satellite observations constrains the utility of XCO₂ data. **While current XCO₂ products exhibit relatively high validation accuracy, their coarse spatial resolution remains inadequate for applications such as regional- or county-level emission monitoring, as well as for the detection and inversion of large emission sources.** To address these issues, we reconstructed a global full-coverage XCO₂ product at a fine spatial resolution of 0.05° and temporal resolution of 1 month from 2015 to 2021. The advanced deep learning method was adopted to model time-series XCO₂ and incorporate terrestrial flux, anthropogenic flux and climatic impacts into the parameterization process. Through comprehensive evaluations, including cross-validation, in-situ validation, spatial distribution assessment and comparison with other XCO₂ products, our reconstructed XCO₂ products demonstrates significant improvements in both accuracy and spatial resolution. **The main conclusions and**

contributions are as following:

(1) The advanced At-BiLSTM model could successfully established the nonlinear relationship between satellite-derived XCO₂ and a set of key environmental variables. And the reconstructed XCO₂ based on our model shows relatively good agreement with TCCON XCO₂, with R², RMSE, and MAE values of 0.91, 1.58 ppm, and 1.22 ppm, respectively.

(2) The reconstructed XCO₂ product overcomes the extensive data gaps typically caused by narrow satellite swaths and retrieval interference from clouds and aerosols, achieving complete global coverage. Moreover, relative to existing publicly available full-coverage XCO₂ datasets, our product offers the finest spatial resolution (0.05°) while maintaining comparable accuracy.

(3) Our method avoids coarse XCO₂ reanalysis inputs, preserving satellite-scale fidelity through high-resolution environmental variables modeling. Consequently, the products enable enhanced ability in identifying regional- and county-level XCO₂ hotspots, carbon emissions and fragmented carbon sinks, providing a robust basis for targeted global carbon governance policies.”

Q2. The model methodology section lacks essential explanations. The study employs the Attention-based Bidirectional Long Short-Term Memory (At-BiLSTM) model for global XCO₂ reconstruction, but it does not justify the choice of this model or clarify its advantages over traditional LSTM models. Additionally, the model’s interpretation remains unclear. It is recommended to provide a rationale for selecting At-BiLSTM and elucidate its specific benefits and interpretive framework.

Response: Thanks for this constructive suggestion. Given the complex temporal dependencies and nonlinear relationships between atmospheric XCO₂ and a wide range of environmental variables, we selected the Attention-based Bidirectional Long Short-Term Memory (At-BiLSTM) model for this study. This choice is motivated by several key considerations:

Firstly, LSTM networks are well-suited for modeling temporal sequences and capturing long-range dependencies, which is essential for understanding the seasonal variations of XCO₂ and dynamic feedbacks between XCO₂ and environmental drivers such as temperature, vegetation activity, and surface pressure. The bidirectional structure enhances this capability by allowing the model to consider both past and future

context in the time series, thereby providing a more comprehensive representation of the underlying temporal dynamics.

Secondly, the incorporation of the attention mechanism enables the model to dynamically focus on the most critical time steps when making predictions. This is particularly important when dealing with high-dimensional input data comprising multi-timestep variables, as it allows the model to assign different weights to different input features, thereby improving interpretability and predictive performance.

Finally, the At-BiLSTM model's ability to capture nonlinear relationships is crucial in the context of atmospheric CO₂ modeling, where interactions between variables are complex and nonlinear. By leveraging the strengths of deep learning, the model can learn intricate patterns from the multi-source data that are difficult to capture with traditional statistical or linear models.

Therefore, we chose At-BiLSTM model as a robust and flexible framework to reconstructing XCO₂ at fine spatial resolution with improved accuracy and spatiotemporal consistency.

We have included the necessary clarifications of its advancement in **2.2 Deep learning-based XCO₂ reconstruction**:

“The LSTM model is a variant of RNN that excels in **modeling temporal sequences and capture long-range dependencies** (Hochreiter and Schmidhuber, 1997; Graves et al., 2005), which is essential for understanding the seasonal variations of XCO₂ and dynamic feedbacks between XCO₂ and environmental drivers we selected. Each LSTM cell includes an input gate, a forget gate and an output gate. The forget gate f_t determines which information from the previous time step to forget (Eq. 1):

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

where σ , W_f , $[h_{t-1}, x_t]$, and b_f denotes the sigmoid activation function, vectors of weights, concatenation of the hidden state at timestep $t-1$ and the current input, and the bias vector, respectively.

The input gate i_t governs the selective storage of the data in current time step, and the output from forget gate f_t and input gate i_t are combined in the cell state C_t (Eq. 2-3):

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

where W_i and W_C denote the weight matrix for the input gate and the current cell

state, respectively; b_i and b_c are the bias vector of the input gate and the current cell state, respectively; C_{t-1} and \tanh represent the cell state at timestep $t-1$ and the activation function.

Lastly, the output gate o_t controls the flow of information from the cell state to the next time step (Eq. 4).

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (4)$$

where W_o and b_o denotes the weight matrix and the bias vector of the output gate, respectively.

These gate structures effectively manage the flow of information within the LSTM, enabling it to capture the temporal dependencies present in the data (Yuan et al., 2020; Su et al., 2021). Bidirectional LSTM consists of two directional LSTM, in which the data flows forward and backward (Graves et al., 2013). The bidirectional structure was chosen to enhance the capability of LSTM by allowing the model to consider both past and future context in the time series, thereby providing a more comprehensive understanding of the underlying temporal dynamics.

We also defined a multi-dimensional attention layer behind the BiLSTM to focus on more critical timesteps and give them higher weights (Bahdanau et al., 2016). This is particularly important when dealing with high-dimensional input data comprising multi-timestep variables, as it allows the model to assign different weights to different timesteps, thereby improving interpretability and predictive performance (Liu and Guo, 2019). Based on this framework, the At-BiLSTM model offers a robust and flexible framework for linking XCO₂ with multiple environmental variables and reconstructing XCO₂ at fine spatial resolution with improved accuracy and spatiotemporal consistency.”

And we have also added the detailed architecture of At-BiLSTM model as follows:

“The At-BiLSTM consists of one input layer, three Bidirectional LSTM (Bi-LSTM) layers, one attention layer, one dropout layer to prevent overfitting, and one fully connected layer (i.e., dense layer) for the final output. Each Bi-LSTM includes 512 hidden units and tanh activation in both forward and backward directions. The attention mechanism learns a weight distribution over the time dimension using a Dense layer with softmax activation, then multiplies these weights with the BiLSTM output to emphasize important time steps. The detailed deployment and output are provided in Table 3. The model was implemented using the TensorFlow and Keras deep learning

APIs in Python. A time step of 3 was used, and the model was trained for 200 epochs with the mean squared error (MSE) as the loss function. A step-wise decay strategy was applied to the learning rate, where the rate was reduced by a factor of 10 every 50 epochs to enhance training stability and convergence. Prior to training, all input data were normalized using the mean and standard deviation of the dataset.”

Table 3. Architecture of the At-BiLSTM model

Layer Name	Layer	Parameters	Output size
Bi-LSTM	Input layer	-	3×16
	Bi-LSTM1	units = 512, activation = ‘tanh’	3×1024
	Bi-LSTM2	units = 512, activation = ‘tanh’	3 × 1024
	Bi-LSTM3	units = 512, activation = ‘tanh’	3 × 1024
Attention	Permute	-	1024×3
	Dense	units = 3, activation = ‘softmax’	1024×3
	Permute	-	3 × 1024
	Multiply	-	3 × 1024
Dropout		rate = 0.5	
Full-connect	Dense	units = 1	1

Q3. (1) The discussion section requires further elaboration. It should comprehensively address the advantages of the model used and the resulting full-coverage XCO₂ product compared to other models and datasets.

Response: Thank you for this valuable comment. We have revised the discussion section, and added two sub-section: **4.1 Comparison with previous studies** and **4.2 Limitations and future improvements**. In added section **4.1 Comparison with previous studies**, we elaborated the comparison with previous studies, and clarified the advantages of the full-coverage XCO₂ product we generated:

“To validate the effectiveness of our model and resulting XCO₂ products, we compared our results with current studies which focuses on global XCO₂ reconstruction (Table 5). As for the in-situ validation, most existing studies report high accuracy with almost all R² over 0.9, RMSE less than 2 ppm. Regarding spatial resolution, the various products differ substantially, ranging from 1° down to 0.01°. It should be noted that increasing spatial resolution tends to compromise the accuracy of XCO₂ retrievals. However, our XCO₂ product achieves an optimal balance between spatial detail and

measurement precision, exhibiting both high spatial resolution (0.05°) and robust accuracy ($R^2=0.91$, $RMSE = 1.54$ ppm) in comprehensive evaluations.

Table 5. Comparison between current studies focusing on global XCO₂ reconstruction

Model	Spatial resolution	In-situ validation (with TCCON)			Reference
		R ²	RMSE (ppm)	MAE (ppm)	
Attentional-based LSTM	0.05°	0.91	1.54	1.22	Our study
Deep forest	0.1°	0.96	1.01	-	Zhang et al. (2023)
S-STDCT	0.25°	0.95	1.18	-	Wang et al. (2023)
Spatiotemporal kriging	1°	0.97	1.13	0.88	Sheng et al. (2022)
MLE & OI	0.5°	0.92	2.62	1.53	Jin et al. (2022)
ERT	0.01°	0.83	1.79	-	Li et al. (2022)

*S-STDCT: Self-supervised spatiotemporal discrete cosine transform, MLE & OI: maximum likelihood estimation method and optimal interpolation; ERT: Extremely randomized trees

To evaluate the advancement of our XCO₂ product, we compared it with original OCO-2 observations and publicly available global XCO₂ datasets (Wang et al., 2023; Sheng et al., 2022; Zhang et al., 2023) across four regions: North America, Europe with northern Africa, Asia, and Oceania (Fig. 13) in January 2015. Despite monthly aggregation, OCO-2 data exhibit persistent spatial discontinuities, limiting the capacity to analyze monthly XCO₂ variability at regional and national scales. Existing XCO₂ products (spatial resolution of 0.25° , 1° , and 0.1° , respectively) broadly reproduce large-scale XCO₂ patterns but fail to resolve fine-scale heterogeneity. In comparison, our reconstructed XCO₂, with the highest spatial resolution, provides a more detailed and accurate representation of the regional XCO₂ patterns. For example, lower XCO₂ concentrations are clearly identified in eastern Canada (The first row of Fig. 13) and Papua New Guinea (The fourth row of Fig. 13), regions characterized by dense forest cover. This correspondence highlights the substantial carbon sink potential of these forested areas. Our high-resolution product better identifies the CO₂ heterogeneity associated with different land cover types, whereas the coarse-resolution products smooth these signals. This limitation primarily stems from the neglect of high-resolution land cover dynamics and dependence on coarse-resolution assimilated/reanalysis datasets (e.g., CAMS XCO₂, CarbonTracker), resulting in oversmoothed spatial patterns that obscure satellite-derived high-resolution signals. Unlike assimilation-dependent approaches, our method avoids XCO₂ reanalysis inputs,

preserving satellite-scale fidelity through high-resolution environmental variables modeling while maintaining precision.

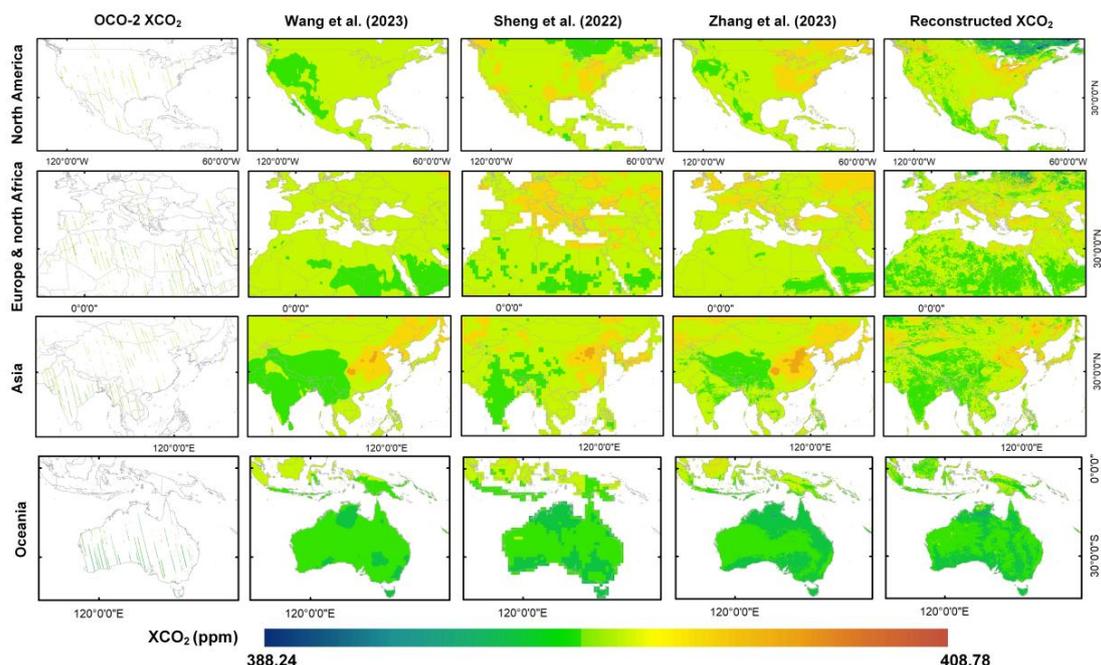


Figure 13. Comparison between the OCO-2 XCO₂ data, accessible XCO₂ products from Wang et al. (2023), Sheng et al. (2022), Zhang et al. (2023), and our reconstructed XCO₂ data in four regions, using the products of January of 2015 as an example.”

(2) Additionally, the global spatial distribution characteristics of XCO₂ need more detailed discussion.

Response: Many thanks for this comment. We have added two figures and more description of the global spatial distribution characteristics of XCO₂ in **3.2 Spatiotemporal pattern of global XCO₂** as follows:

The global distribution of annual mean XCO₂ concentration from 2015 to 2021 is illustrated in Fig. 8. The results reveal pronounced spatial heterogeneity in XCO₂ concentrations, characterized by a marked hemispheric asymmetry. Specifically, the Northern Hemisphere exhibited systematically elevated XCO₂ levels compared to the Southern Hemisphere, consistent with latitudinal gradients driven by anthropogenic emission patterns and atmospheric transport dynamics. Regionally, North America, East Asia, Central Africa, and northwest of Southern America were identified as persistent hotspots of enhanced XCO₂. The high concentrations of XCO₂ in North America and East Asia stem primarily from the fossil fuel emission from energy production and transportation sectors. Whereas the tropical regions (i.e., Central Africa

and South America) are influenced by coupled biomass burning and land-use changes.

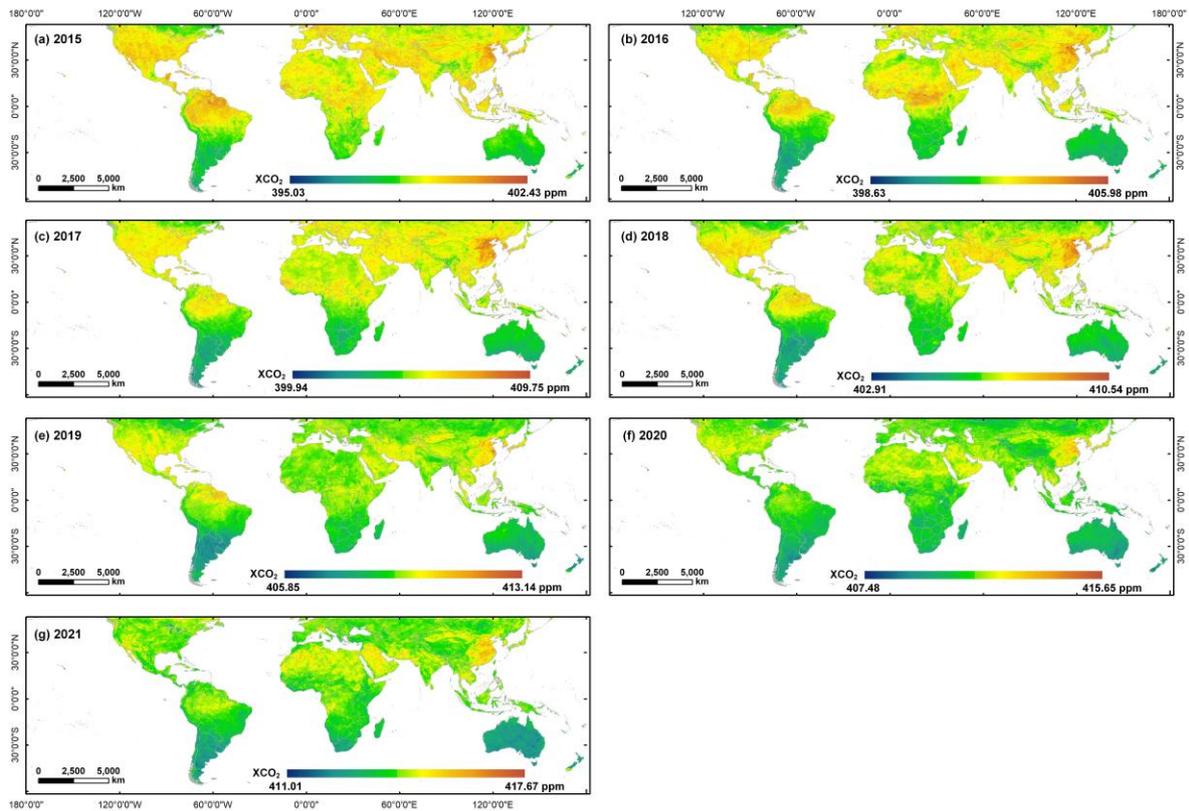


Figure 8. The global spatial distribution of reconstructed annual mean XCO₂ concentration from 2015 to 2021.

We also provided the annual OCO-2 XCO₂ data from 2015 to 2019 and OCO-3 XCO₂ data from 2020 to 2021 in Fig. 9. Spatially, our reconstructed XCO₂ dataset demonstrates robust consistency with satellite observations, particularly in mid-latitude industrialized regions where both datasets capture emission hotspots. Notably, OCO-3 exhibits denser observational sampling due to its improved spatial coverage and swath width compared to OCO-2's narrow tracks. However, persistent data gaps remain prevalent in both two satellite products after annual aggregating. These spatial coverage limitations hinder fine-scale global analysis, particularly in assessing localized emission sources and regional scale carbon flux.

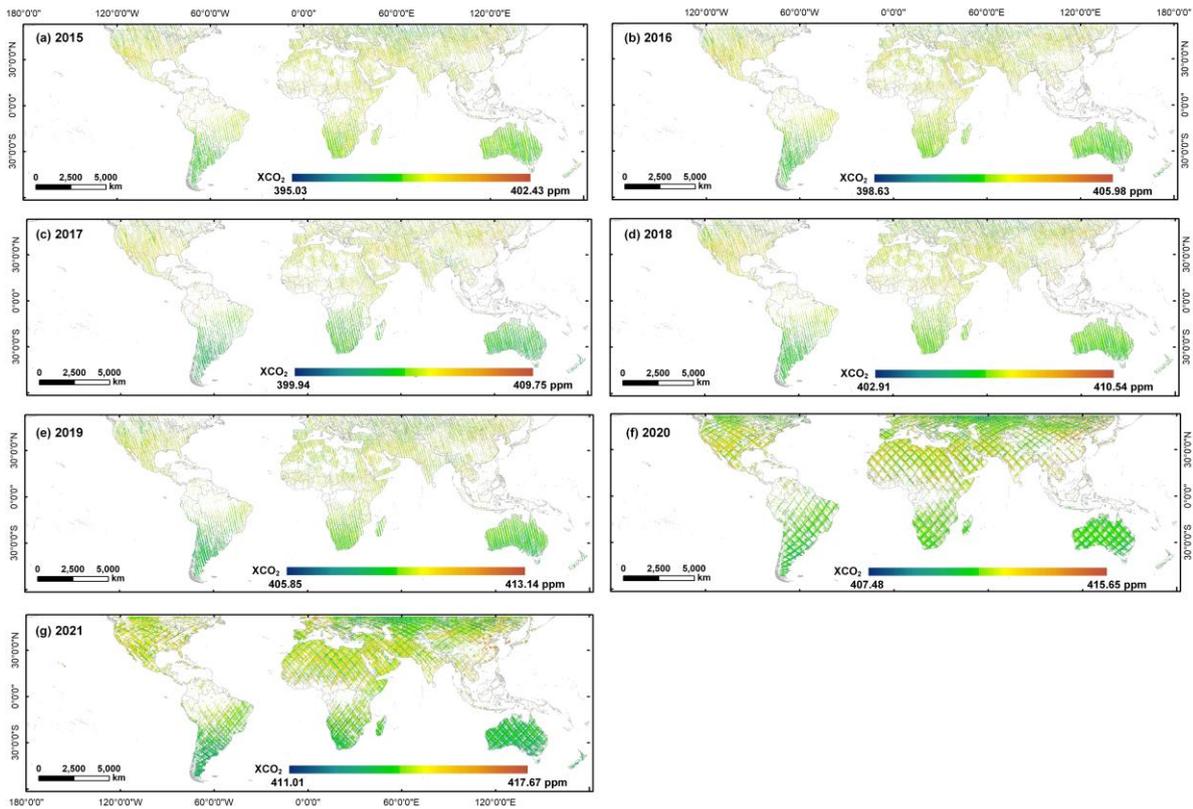


Figure 9. The global spatial distribution of annual mean OCO-2/OCO-3 XCO₂ concentration from 2015 to 2021.

Q4. (1) In the conclusion or discussion section, please clearly specify the concrete data or scientific significance of the high-resolution XCO₂.

Response: Thanks for this constructive suggestion. We have provided further examples to introduce the significance of high-resolution XCO₂ products, in **4.1 Comparison with previous studies** we supplemented the comparison with other coarse-resolution data products as follows:

“To evaluate the advancement of our XCO₂ product, we compared it with original OCO-2 observations and publicly available global XCO₂ datasets (Wang et al., 2023; Sheng et al., 2022; Zhang et al., 2023) across four regions: North America, Europe with northern Africa, Asia, and Oceania (Fig. 13) in January 2015. Despite monthly aggregation, OCO-2 data exhibit persistent spatial discontinuities, limiting the capacity to analyze monthly XCO₂ variability at regional and national scales. Existing XCO₂ products (spatial resolution of 0.25°, 1°, and 0.1°, respectively) broadly reproduce large-scale XCO₂ patterns but fail to resolve fine-scale heterogeneity. In comparison, our reconstructed XCO₂, with the highest spatial resolution, provides a more detailed

and accurate representation of the regional XCO₂ patterns. For example, lower XCO₂ concentrations are clearly identified in eastern Canada (The first row of Fig. 13) and Papua New Guinea (The fourth row of Fig. 13), regions characterized by dense forest cover. This correspondence highlights the substantial carbon sink potential of these forested areas. Our high-resolution product better identifies the CO₂ heterogeneity associated with different land cover types, whereas the coarse-resolution products smooth these signals. This limitation primarily stems from the neglect of high-resolution land cover dynamics and dependence on coarse-resolution assimilated/reanalysis datasets (e.g., CAMS XCO₂, CarbonTracker), resulting in oversmoothed spatial patterns that obscure satellite-derived high-resolution signals. Unlike assimilation-dependent approaches, our method avoids XCO₂ reanalysis inputs, preserving satellite-scale fidelity through high-resolution environmental variables modeling while maintaining precision.”

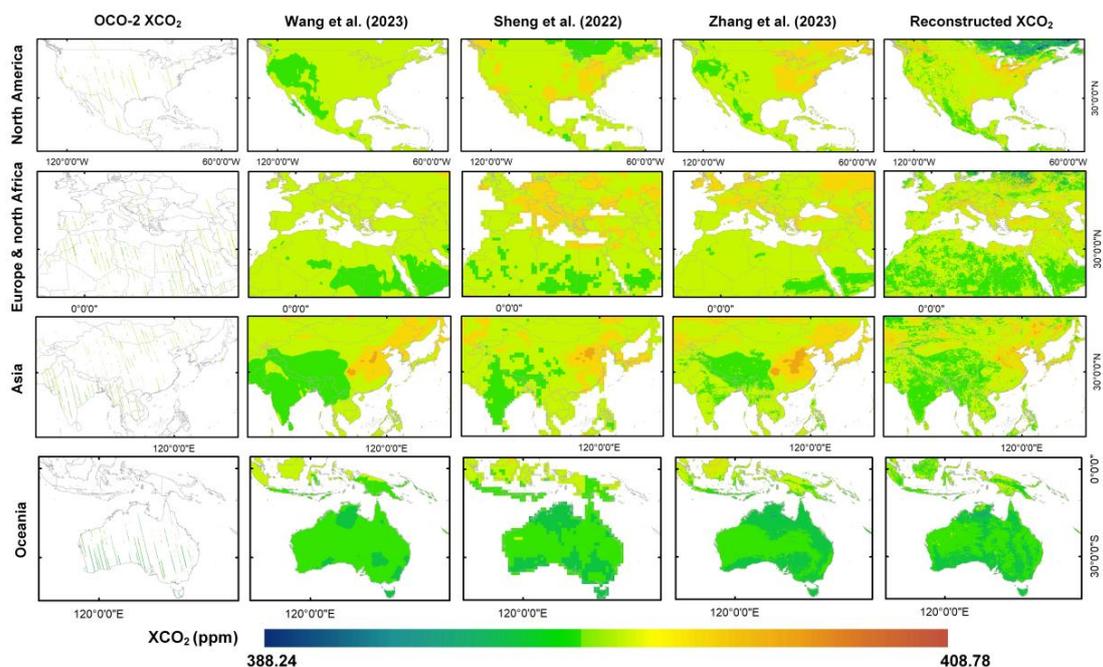


Figure 13. Comparison between the OCO-2 XCO₂ data, accessible XCO₂ products from Wang et al. (2023), Sheng et al. (2022), Zhang et al. (2023), and our reconstructed XCO₂ data in four regions, using the products of January of 2015 as an example.

And we also elaborated further in 6. **Conclusion** as follows:

“As a major driver of global warming, the monitoring of CO₂ changes, especially anthropogenic CO₂ emissions, is of critical importance. The launch of carbon satellites offers a significant advancement for CO₂ monitoring. However, the limited spatial coverage of satellite observations constrains the utility of XCO₂ data. **While current XCO₂ products exhibit relatively high validation accuracy, their coarse spatial**

resolution remains inadequate for applications such as regional- or county-level emission monitoring, as well as for the detection and inversion of large emission sources. To address these issues, we reconstructed a global full-coverage XCO₂ product at a fine spatial resolution of 0.05° and temporal resolution of 1 month from 2015 to 2021. The advanced deep learning method was adopted to model time-series XCO₂ and incorporate terrestrial flux, anthropogenic flux and climatic impacts into the parameterization process. Through comprehensive evaluations, including cross-validation, in-situ validation, spatial distribution assessment and comparison with other XCO₂ products, our reconstructed XCO₂ products demonstrates significant improvements in both accuracy and spatial resolution. The main conclusions and contributions are as following:

(1) The advanced At-BiLSTM model could successfully established the nonlinear relationship between satellite-derived XCO₂ and a set of key environmental variables. And the reconstructed XCO₂ based on our model shows relatively good agreement with TCCON XCO₂, with R², RMSE, and MAE values of 0.91, 1.58 ppm, and 1.22 ppm, respectively.

(2) The reconstructed XCO₂ product overcomes the extensive data gaps typically caused by narrow satellite swaths and retrieval interference from clouds and aerosols, achieving complete global coverage. Moreover, relative to existing publicly available full-coverage XCO₂ datasets, our product offers the finest spatial resolution (0.05°) while maintaining comparable accuracy.

(3) Our method avoids coarse XCO₂ reanalysis inputs, preserving satellite-scale fidelity through high-resolution environmental variables modeling. Consequently, the products enable enhanced ability in identifying regional- and county-level XCO₂ hotspots, carbon emissions and fragmented carbon sinks, providing a robust basis for targeted global carbon governance policies.”

(2) Additionally, provide an outlook for future research, outlining key issues to address in global XCO₂ or CO₂ concentration reconstruction studies, such as critical challenges or priorities that should be focused on.

Response: Thank you for this comment. We have added section **4.2 Limitations and future improvements** to provide a more detailed discussion of key challenges and the future outlook. Two key points are highlighted: firstly, the incorporation of auxiliary

variables that capture vertical CO₂ transport. Secondly, the enhancement of satellite observation coverage and accuracy to minimize data gaps and retrieval errors. The revised context is as follows:

“Additionally, though our model integrates multiple environmental variables associated with surface carbon flux variations, it does not account for vertical atmospheric transport. As XCO₂ represents the column-averaged CO₂ concentration, vertical redistribution of CO₂ through atmospheric transport (e.g., mixing, convection) can alter the relationship between surface carbon fluxes and column concentrations. The absence of such vertical transport indicators may reduce the model’s accuracy in regions or periods with strong vertical mixing. Future efforts will incorporate vertical transport-related variables, such as planetary boundary layer height, vertical wind components, and other reanalysis-derived indicators, to better represent the atmospheric processes that influence the column-averaged CO₂ signal.

Moreover, while OCO missions currently provide some of the most accurate carbon satellite-based XCO₂ retrievals, they still encounter some retrieval errors and data gaps driven by algorithmic limitations and variable meteorological conditions. A critical research frontier is the refinement of XCO₂ retrieval algorithms to mitigate systematic biases in high-aerosol-load regions (e.g., industrial regions and biomass-burning plumes). Additionally, next-generation hyperspectral satellites, such as the upcoming CO2M (Copernicus Anthropogenic CO₂ Monitoring Mission) with 2×2 km² resolution and GeoCarb (Geostationary Carbon Observatory) offering hourly monitoring, will enhance spatial-temporal coverage and reduce cloud-induced data gaps.”

Q5. The construction of the OCO dataset is unclear. For instance, it is not specified how grids containing both OCO-2 and OCO-3 data within the same time period were processed.

Response: Thanks for this comment. In this study, given that OCO-3 has more intensive observations, we utilized the OCO-3 XCO₂ data for all available year (i.e., 2020-2021) and used the data of OCO-2 for the other years (i.e., 2015-2019). Although OCO-3 began providing data in August 2019, we used OCO-2 data for entire 2019 to maintain consistency in our monthly estimates.

In addition, analysing OCO-2 and OCO-3 data simultaneously may introduce several uncertainties due to their different spatiotemporal coverages. However, OCO-3 has a similar sensor with OCO-2 and inherits the retrieval algorithms of OCO-2. According to Taylor et al. (2023), the mean differences between OCO-3 and OCO-2 are around 0.2 ppm over land. Therefore, we suppose that the discrepancies between their datasets are minimal, and the combined analysis of data from these two satellites will have a negligible impact on our results. And we discussed this in **4.2 Limitations and future improvements** as follows:

“In terms of the satellite data, OCO-2 and OCO-3 provide different spatiotemporal coverages. Analyzing OCO-2 and OCO-3 data simultaneously may introduce several uncertainties due to these differences. However, OCO-3 has a similar sensor and inherits the retrieval algorithms of OCO-2. According to Taylor et al. (2023), the mean differences between OCO-3 and OCO-2 are around 0.2 ppm over land. Therefore, we suppose that the discrepancies between their datasets are minimal, and the combined analysis of data from these two satellites will have a negligible impact on our results.”

Q6. The study utilized various satellite-derived variables, including land flux, anthropogenic flux, and climatic impacts, for global XCO₂. However, it is unclear whether these satellite data have gaps, particularly in high-latitude regions. If gaps exist, the study should specify how they were addressed.

Response: Many thanks for this constructive comment. Among all ancillary variables, those related to climatic impacts from ERA5-Land, as well as land use and cover change (LUCC), vegetation continuous fields (VCF), and nighttime lights (NTL), provide full spatial coverage. Given that the XCO₂ reconstruction was performed on a monthly scale, all satellite-derived variables were aggregated to monthly averages using Google Earth Engine (GEE). During this averaging process, most data gaps were effectively filled. For variables that still contained missing values after monthly aggregation, we applied bilinear interpolation methods to fill the remaining gaps.

We have added the description of data processing in **2.1.3 Environmental variables** as follows:

“All data were converted to monthly time-series. The bilinear interpolation approach was employed both to fill gaps in the ancillary data and to convert the data at

different spatial resolutions to 0.05° resolution. The data preprocessing was conducted on GEE, R and ArcGIS 10.3.”

Q7. Line 243-244. “...spatial resolutions to 1 km resolution”. The ‘1 km resolution’ is inconsistent with the study’s focus on a 0.05°

Response: Thank you for your correction. We apologize for the typographical error and all data have been processed at the resolution of 0.05° to match the XCO₂ products. We have revised these accordingly.

Q8. The manuscript contains several typo errors. For instance, in Line 303, “Figure 5. (a) Density scatterplots of sample-based...” includes an unnecessary ‘(a)’. In Line 314, the value ‘1.21’ should be corrected to ‘1.22 ppm’. These errors should be revised for accuracy and clarity.

Response: Thank you for the correction. We have revised these accordingly and checked the full manuscript to avoid such typo errors.