

Reconstructing two-decade daily high-resolution seamless global land XCO₂ records using a hybrid Transformer–BiLSTM model

5 Yu Qu^{1,2}, Xian Shi³, Yulong Fan^{1,3}, Zhihui Wang^{4,5}, Jing Wei^{1*}

1. MEEKL-AERM, College of Environmental Sciences and Engineering, Institute of Tibetan Plateau, and Center for Environment and Health, Peking University, Beijing 100871, China.
2. School of Geographical Sciences, South China Normal University, Guangzhou, 510631, China.
- 10 3. College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China
4. University of Science and Technology of China, Hefei 230026, China
5. Anhui Province Key Laboratory of Optical Quantitative Remote Sensing, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China

15

Correspondence: Jing Wei (jingwei@pku.edu.cn)

Abstract: Accurate and temporally continuous global observations of atmospheric carbon dioxide (XCO₂) are essential for climate monitoring and emission assessment. However, satellite-based XCO₂ observations are often spatially incomplete and temporally discontinuous, while existing products typically suffer from coarse spatial resolutions, hindering the detection of fine-scale emission changes. Here, we developed a novel spatiotemporal Transformer-BiLSTM deep-learning network, which integrates the local temporal feature extraction capability of bidirectional long short-term memory with the global spatial dependency modeling strength of Transformer via self-attention mechanisms. The network assimilates multisource data, from satellite observations, meteorological reanalysis, and precursor gases, to reconstruct a global, daily, and seamless XCO₂ dataset over land at 0.1° resolution from 2003 to 2022. Independent validation of the data-fused XCO₂ 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 R² of 0.99

20
25
30

and an RMSE of 0.36 ppm in the sample-based ten-fold cross-validation. Our dataset enables accurate characterization of daily XCO₂ concentrations over global land surfaces, facilitating the detection of spatial heterogeneity associated with emission hotspots and point-source activities. The record reveals a persistent global increase in atmospheric XCO₂ over the past two decades, with a mean growth rate of 2.24 ppm/yr ($p < 0.001$). It reliably resolves global XCO₂ variability across a wide range of temporal scales, from day-to-day fluctuations to long-term trends. It consistently captures large-scale climate-driven signals, such as ENSO-related interannual variability, and short-lived XCO₂ enhancements associated with major wildfire events, demonstrating its capability to represent both persistent and episodic emission signals. This high-resolution, daily global XCO₂ product (GlobalHighXCO₂) provides a valuable benchmark for carbon cycle studies, atmospheric model evaluation, and emission monitoring, and is publicly available at <https://doi.org/10.5281/zenodo.18220961> (Qu and Wei, 2026).

1 Introduction

Carbon dioxide (CO₂) is a principal greenhouse gas that plays a significant role in driving global climate change (Romanov, 2017). Driven by anthropogenic activities and extreme wildfires, the global mean atmospheric CO₂ concentration reached 419.31 ± 0.15 parts per million (ppm) in 2023, approximately 50% higher than preindustrial levels, making it a major contributor to ongoing climate change (Budget, 2023). In response to rising CO₂ levels, the international community adopted the Paris Agreement in 2015, aiming to limit global warming to well below 2°C compared to preindustrial levels. Owing to its long atmospheric lifetime, CO₂ accumulates in the atmosphere and exerts sustained radiative forcing, resulting in long-term climate impacts (Lee et al., 2023; Kemp et al., 2022). Accurate long-term monitoring of atmospheric CO₂ is therefore essential for advancing understanding of the global carbon cycle, verifying national emission reduction commitments, and informing effective climate mitigation policies.

Atmospheric column-averaged dry air mole fractions of CO₂ (XCO₂) are commonly quantified using ground-based networks and satellite-based remote sensing techniques (Petzold et al., 2015; Huang et al., 2024a). Ground-based networks, such as the Total Carbon Column Observing Network (TCCON) provide accurate and stable point-based measurements; however, their sparse spatial distribution

limits their ability to support refined assessments of the global carbon budget and regional carbon fluxes (Li et al., 2024a). Satellite remote sensing offers a complementary approach by providing broad spatial coverage for regional and global-scale XCO₂ monitoring, but has a lower temporal sampling frequency (Buchwitz et al., 2015). Since the early 2000s, a series of satellite missions has progressively advanced XCO₂ observations in accuracy, spatial resolution, and temporal coverage. The SCanning Imaging Absorption spectroMeter for Atmospheric CHartography (SCIAMACHY) aboard the Environmental Satellite, launched in 2002, was the first satellite instrument to provide global XCO₂ observations until 2012 (Bovensmann et al., 1999). Subsequently, a new generation of spaceborne sensors, including Japan's Greenhouse Gas Observing Satellite (GOSAT) (Kuze et al., 2009; Butz et al., 2011) NASA's Orbiting Carbon Observatory-2/3 (OCO-2/3) (Crisp et al., 2017), and China's TanSat (Yang et al., 2018), have substantially advanced top-down constraints on carbon sources and sinks. These satellite data have become indispensable for our understanding of large-scale carbon cycle processes. Despite major advances in satellite XCO₂ observations, existing products remain limited by cloud and aerosol contamination, surface effects, and orbital sampling constraints (He et al., 2022). These gaps severely restrict the accurate analysis of seasonal and interannual changes in regional carbon fluxes and impede the detection of carbon-cycle anomalies associated with extreme climate events (Ma et al., 2021; Li et al., 2022).

To overcome spatial gaps in satellite XCO₂ observations, researchers have explored a variety of data fusion and reconstruction approaches. Early studies primarily relied on geospatial interpolation techniques, such as kriging and its spatiotemporal extensions (Hammerling et al., 2012; He et al., 2020; Wang et al., 2025; Chen et al., 2024). While effective in data-rich regions, these methods often assume stationary spatial correlations, leading to large uncertainties in sparsely observed areas. Global reanalysis products, including CAMS and CarbonTracker, provide spatially complete fields but typically have coarse spatial resolution, limiting their ability to resolve regional emission sources (Hua et al., 2024). Traditional machine learning (ML) models, such as Random Forest and XGBoost, have been widely adopted for XCO₂ reconstruction due to their strong nonlinear fitting capabilities (Ma et al., 2021; Zhang et al., 2023; Liang et al., 2023). However, these approaches generally treat reconstruction as point-wise regression problems, neglecting the intrinsic spatiotemporal autocorrelation and physical coherence of atmospheric fields, which often results in spatially and

temporally discontinuous outputs (Siabi et al., 2019; He et al., 2023b). Although some recent studies have attempted to incorporate spatiotemporal information (Wang et al., 2020; Liu et al., 2024), their ability to capture long-range temporal dependencies and cross-regional teleconnections remains limited (He et al., 2023a). The rapid development of deep learning (DL) has created new opportunities
95 for reconstructing atmospheric data, overcoming the limitations of conventional linear models (Wei et al., 2024). Spatiotemporally aware DL architectures, particularly those integrating convolutional neural networks (CNNs) with recurrent neural networks such as long short-term memory (LSTM), have demonstrated strong capabilities in capturing fine-scale spatial patterns and temporal dependencies (He et al., 2024; Zhang and Liu, 2023; Wu et al., 2024). Recent studies have leveraged
100 these models to fuse multisource satellite data (Huang et al., 2024a), incorporate temporal sequence features for daily XCO₂ estimation (Tian et al., 2024), and generate high-resolution regional products using deep autoencoders (Antezana Lopez et al., 2025; Li et al., 2025; Wang, 2026).

Despite these methodological advances, generating a truly seamless global daily XCO₂ dataset
105 remains challenging. Harmonizing long-term records from multiple satellite missions, including SCIAMACHY, GOSAT, and OCO-2, introduces inter-sensor inconsistencies due to differences in instrument design and retrieval algorithms (Li et al., 2022; Chen et al., 2024). Although TCCON observations have been used for post hoc harmonization in some studies (Huang et al., 2024b; Li et al., 2024b), bias correction is not yet fully integrated into the reconstruction framework. Moreover,
110 existing global products often face a trade-off between spatial resolution and temporal frequency, providing either coarse daily fields or high-resolution monthly averages, thereby limiting their ability to capture short-term variability associated with extreme events such as wildfires (Liu et al., 2022; Rodrigues et al., 2025) and climate anomalies during El Niño episodes (Chatterjee et al., 2017; Betts et al., 2016). Consequently, a unified methodology that simultaneously addresses inter-satellite biases
115 and captures long-range spatiotemporal dependencies at the global daily scale is still lacking.

To address these limitations, we develop a hybrid spatiotemporal deep learning framework based on a Transformer–BiLSTM architecture for global daily XCO₂ reconstruction. Through jointly learning temporal evolution and spatial coherence from multisource satellite observations, reanalysis data, and auxiliary predictors, the proposed approach effectively captures complex spatiotemporal variability
120 in atmospheric XCO₂ fields. Leveraging this framework, we generate a seamless global daily XCO₂

dataset over land at 0.1° resolution spanning 2003–2022. A TCCON-based bias correction strategy is incorporated to harmonize multiple satellite mission records and ensure long-term temporal consistency. In addition, explainable artificial intelligence techniques are employed to enhance model interpretability and to quantify the relative contributions of key drivers of XCO_2 variability. Together, these advances provide a consistent long-term XCO_2 product suitable for carbon-cycle analysis, emission monitoring, and climate-related applications.

2 Materials and methods

2.1 Multisource data

2.1.1 Satellite, reanalysis, and ground-based XCO_2 data

Satellite-based XCO_2 retrievals from SCIAMACHY, GOSAT, and OCO-2 were used as the primary observational constraints in this study. SCIAMACHY, onboard ENVISAT and operating from 2002 to 2012, provided global XCO_2 observations at a resolution of $30 \times 60 \text{ km}^2$ approximately every 30 days, retrieved using the Bremen Optimal Estimation DOAS algorithm (BOED) (Bovensmann et al., 1999; Buchwitz et al., 2005; Reuter et al., 2011). GOSAT, launched in 2009, provides near-global coverage at 10.5 km^2 every three days through its Fourier Transform Spectrometer. OCO-2, operating since 2014, provides XCO_2 measurements at $1.29 \times 2.25 \text{ km}^2$ resolution every ~ 16 days using three grating spectrometers (Crisp, 2015). Here, we used the SCIAMACHY BOED products from 2003 to 2009, GOSAT Level 2 XCO_2 Version 9r products from 2010 to 2014, OCO-2 Level 2 XCO_2 Lite Version 11r product from 2015 to 2022. All retrievals underwent standard quality control, and only high-quality observations were retained. In addition, the Copernicus Atmosphere Monitoring Service (CAMS) provides a global atmospheric composition reanalysis, offering continuous greenhouse gas fields from 2003 to the present at $0.75^\circ \times 0.75^\circ$ horizontal and three-hour temporal resolution (Agustí-Panareda et al., 2023). CAMS provides spatially complete and consistent estimates of large-scale XCO_2 patterns and was also employed in this study.

Ground-based observations were obtained from the TCCON, which provides high-precision column carbon dioxide (XCO_2) measurements (average uncertainty $\sim 0.2\%$), with sub-daily temporal sampling, typically consisting of multiple observations per clear-sky day, and are widely used to evaluate satellite retrievals (Wunch et al., 2017). We used all available 31 stations in the GGG2020

release, distributed across global land (Fig. 1), serving as the reference for assessing reconstruction performance and guiding bias correction. To independently evaluate the bias-corrected daily XCO₂, we used surface flask measurements from the Observations Package (ObsPack) dataset, including 41 land-based sites (Fig. 1). Only stations with a quality inspection flag of one or two were retained to ensure consistency with the World Meteorological Organization X2019 calibration scale.

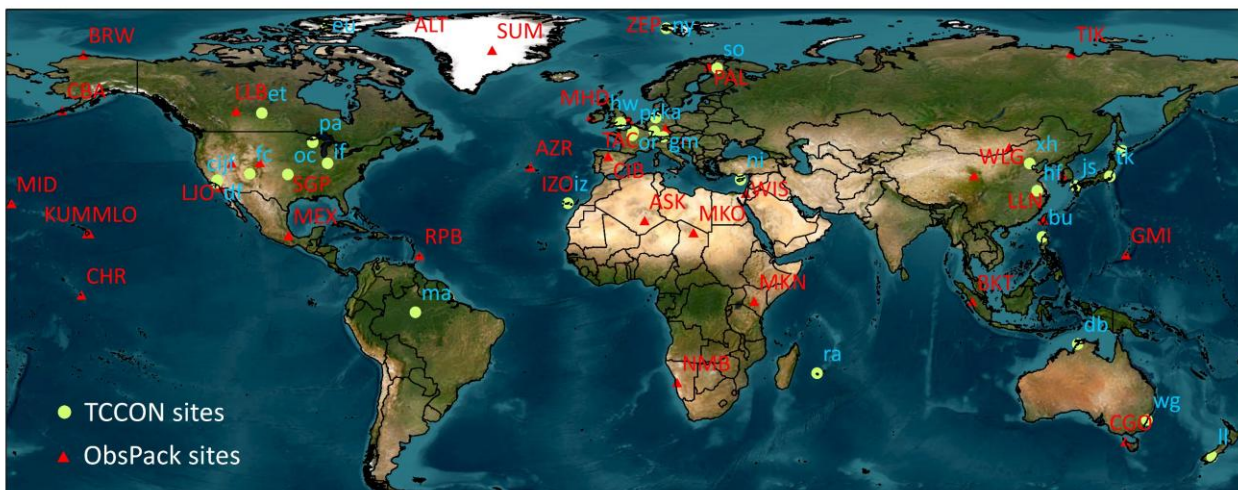


Figure 1. Geographic locations of TCCON (Total Carbon Column Observing Network, green dots) and ObsPack (Observations Package, red triangles) sites.

160

2.1.2 Meteorological, precursor gases, and auxiliary variables

Meteorological variables relevant to atmospheric CO₂ variability were obtained from the ERA5-Land hourly reanalysis at a spatial resolution of $0.1^\circ \times 0.1^\circ$, including surface pressure (SP), 2 m air temperature (T2M), surface net solar radiation (SNSR), evaporation (ET), and 10 m wind components (U10 and V10). Boundary layer height (BLH) and relative humidity (RH) were collected from the ERA5 reanalysis at $0.25^\circ \times 0.25^\circ$ resolution. To capture short-lived atmospheric precursors associated with anthropogenic and biomass-burning emissions, we included daily nitrogen dioxide (NO₂) and carbon monoxide (CO) from CAMS analysis, as well as column-averaged methane (XCH₄) derived from our previous study (Qu et al., 2025), which are used as proxies for fossil-fuel combustion and fire-related emissions (Reuter et al., 2019). Land-surface conditions were characterized using MODIS products, including monthly normalized difference vegetation index (NDVI) and land surface temperature (LST), which are closely related to terrestrial carbon uptake and surface-atmosphere carbon exchange (Huang et al., 2015; Li et al., 2025). Topographic effects were represented using the

170

Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) at a spatial resolution of 90 m. Table 1 provides an overview of all data sources used in this study.

Table 1. Summary of the data used for the XCO₂ reconstruction in our study.

Type	Source	Product	Spatial resolution	Temporal resolution	Period
Satellite	SCIAMACHY	XCO ₂	30×60 km ²	~30 days	2003-2009
	GOSAT	XCO ₂	10.5 km ²	~3 days	2010-2014
	OCO-2	XCO ₂	1.29×2.25 km ²	~16 days	2015-2022
Meteorological variables	ERA5	Surface pressure (SP)	0.1°	Hourly	2003-2022
		2-meter temperature (T2M)			
		10m u wind (U10)			
		10m v wind (V10)			
		Surface net solar radiation (SNSR)			
		Total evaporation (ET)			
		Relative humidity (RH)			
Precursor	CAMS (Qu et al., 2025)	CAMS-XCO ₂	0.75°	3-hour	2003-2022
		Nitrogen dioxide (NO ₂)	0.75°	3-hour	2003-2022
		CO	0.75°	3-hour	2003-2022
		GlobalHighXCH ₄	0.1°	Daily	2003-2022
		Boundary layer height (BLH)	0.25°		
Surface-related variables	MODIS	NDVI	0.05°	Monthly	2003-2022
		LST			
	SRTM	Elevation	90 m	-	-

2.2 Model development and construction

2.2.1 Transformer-BiLSTM framework

In this study, we proposed a hybrid spatiotemporal Transformer-BiLSTM deep-learning framework to capture both temporal and spatial dependencies. Temporal sequences of all input predictors are first processed by a Bidirectional Long Short-Term Memory (Bi-LSTM) network, which encodes forward and backward dependencies to generate comprehensive temporal feature representations for each grid cell, integrating both past and future context (Zhang et al., 2020). This sequence-to-sequence formulation captures temporal dependencies arising from diurnal cycles, synoptic weather variations, and seasonal biospheric processes affecting XCO₂ concentrations.

These temporally encoded features were then fed into the Transformer encoder (Vaswani et al., 2017),

190 which models long-range spatial dependencies at regional and global scales via multi-head self-attention and position-wise feed-forward layers, without relying on local receptive fields. Each location’s feature vector is projected into a Query, a Key, and a Value, and the resulting attention scores quantify the influence of all other locations on that point. The multi-head mechanism integrates complementary spatial relationships across multiple subspaces, enabling the model to explicitly learn
 195 planetary-scale influences, such as atmospheric teleconnections and large-scale weather systems, on regional XCO₂ concentrations. The resulting spatiotemporal features are subsequently passed through a multilayer perceptron (MLP) block, followed by a fully connected linear layer that projects the fused features to scalar daily XCO₂ values for each grid cell. To effectively capture the complex spatiotemporal dynamics of atmospheric XCO₂, we designed a weighted spatiotemporal loss function
 200 that accounts for intrinsic temporal continuity and spatial correlation, in contrast to standard Mean Squared Error (MSE) loss functions:

$$L_{XCO_2} = \frac{1}{N} \sum_{i=1}^N (|y_i - \hat{y}_i| + \chi_1 * (|\Delta y_i - \Delta \hat{y}_i|) + \chi_2 * L_{spatial}) \quad (1)$$

205 where the first term ensures point-wise accuracy, the second term (weighted by χ_1) penalizes errors in temporal derivatives, and the third term (weighted by χ_2) enforces spatial smoothness via a Laplacian constraint. Optimal weights were determined via grid search ($\chi_1 = 0.5, \chi_2 = 0.01$), ensuring robust reconstruction of seasonal cycles without over-smoothing high-emission features.

210 During model construction, the dataset was randomly partitioned into training, validation, and testing sets at a ratio of 8:1:1. 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. A dropout rate of 0.5 was applied to mitigate overfitting. The Transformer module comprises 4 encoder layers, each with 64-dimensional hidden layers and 4 attention heads, with ReLU activation functions
 215 applied between layers to enhance learning and facilitate sparse feature extraction. The Transformer-BiLSTM model was trained using the Adam optimizer with an initial learning rate of 0.001 and a batch size of 256. To optimize convergence, a learning rate scheduler reduced the rate by a factor of 0.1 every 10 epochs if the validation loss did not improve. Training continued for up to 200 epochs,

with early stopping applied if the validation loss failed to decrease for 15 consecutive epochs (Yeom
220 et al., 2021).

2.2.2 Two-phase reconstruction workflow

To address the spatial discontinuity of raw satellite observations and the limited coverage of TCCON
ground-based stations, we designed a sequential two-phase “fusion-then-correction” reconstruction
225 workflow (Fig. 2), decoupling spatial pattern learning from magnitude calibration.

Phase 1: Data Fusion. The goal of this stage is to generate a spatially continuous global XCO₂ field
from sparse and discontinuous satellite swaths. The Transformer-BiLSTM model first takes XCO₂
retrievals from SCIAMACHY, GOSAT, and OCO-2 as the target and ingests multisource inputs,
230 including atmospheric fields, i.e., XCH₄ from our previous study (Qu et al., 2025), CAMS XCO₂ and
NO₂ concentrations; ERA5 meteorology variables (i.e., SP, T2M, SNSR, ET, U10, V10, BLH, RH);
and surface or terrestrial variables (i.e., NDVI, LST, DEM). In addition, spatiotemporal encoding
vectors are incorporated: spatial coordinates are represented by three Euclidean spherical functions
(P_S), and temporal dynamics by three helix-shaped trigonometric vectors (P_T), enabling the model to
235 capture spatial heterogeneity and temporal variability (Wei et al., 2023; Qu et al., 2025). The output
of this phase is a preliminary fused daily XCO₂ estimates at 0.1°×0.1° resolution, providing a spatially
continuous field to support the subsequent bias correction stage.

$$Fusion_{XCO_2} \sim f_{TF-BiLSTM}(CAMS_{XCO_2}, CAMS_{NO_2}, Qu_{XCH_4}, Pre_{gas}, Meteo, NDVI, LST, DEM, P_S, P_T)$$

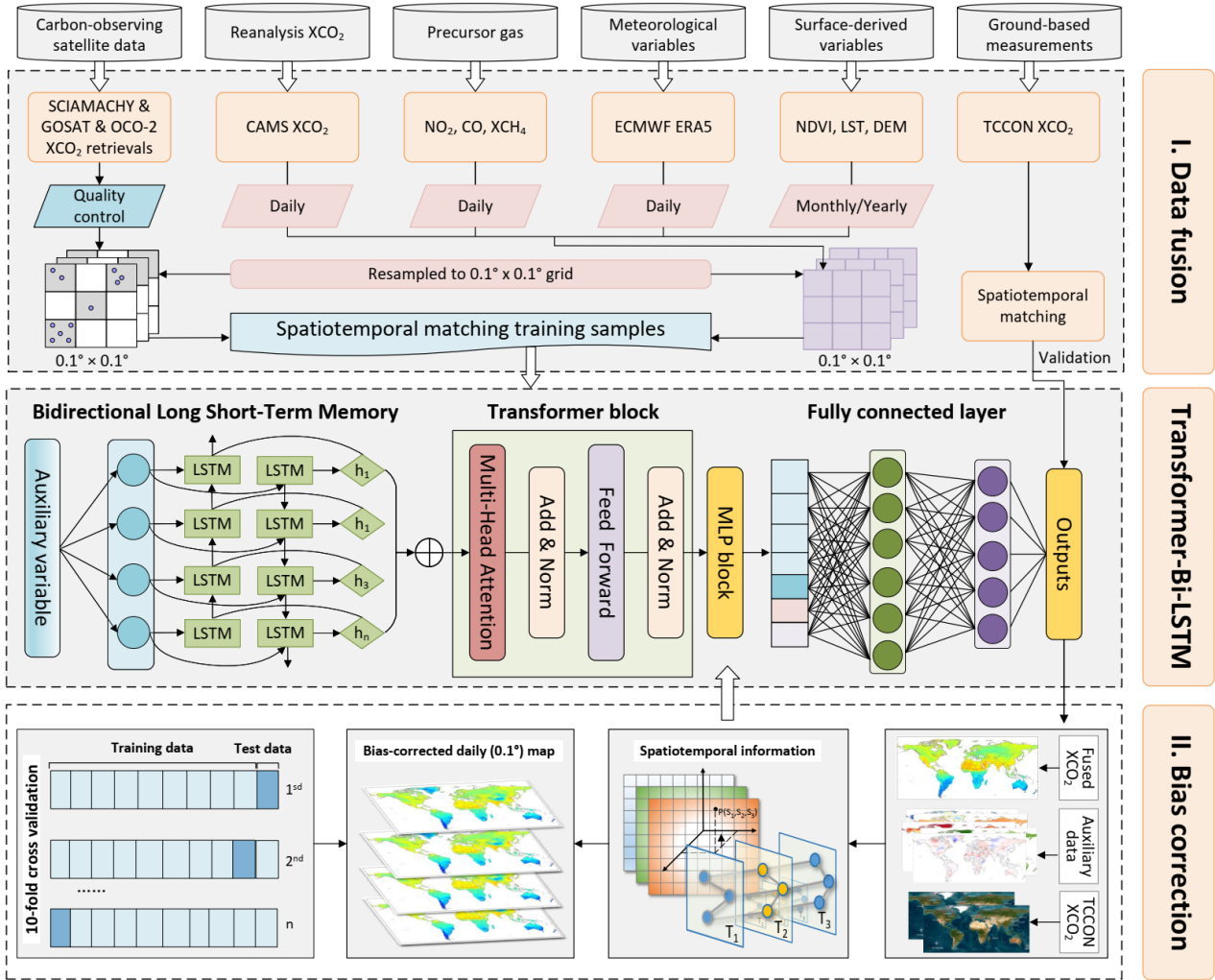
240 (2)

Phase 2: Bias Correction. Satellite XCO₂ retrievals often exhibit systematic biases due to differences
in sensor characteristics, retrieval algorithms, and temporal sampling. In this phase, the fused XCO₂
field from Phase 1 is further calibrated (bias-corrected) using TCCON ground-based measurements
as reference standards. The Transformer-BiLSTM model ingests the fused XCO₂ data together with
245 the same auxiliary predictors used in the fusion stage, including precursor gases (XCO₂, XCH₄, and
NO₂), meteorological variables, surface/terrestrial features, and spatiotemporal encoding vectors. By
learning the complex nonlinear relationships between satellite-dependent biases and environmental
conditions, this framework achieves robust global generalization. The output is a bias-corrected, gap-

free, daily global XCO₂ dataset at 0.1° resolution, consistent across all satellite platforms:

250

$$Bias_{XCO_2} \sim f_{TF-BiLSTM}(Fusion_{XCO_2}, CAMS_{NO_2}, Qu_{XCH_4}, Pre_{gas}, Meteo, NDVI, LST, DEM, P_S, P_T) \quad (3)$$



255

Figure 2. Workflow for generating global daily gapless XCO₂ concentrations with a 0.1°x0.1° resolution using the developed Transformer-BiLSTM framework.

2.2.3 Innovations of the framework

260

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.

We further 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. Ablation analysis comparing different backbone models for XCO₂ reconstruction

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

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.

2.3 Validation method

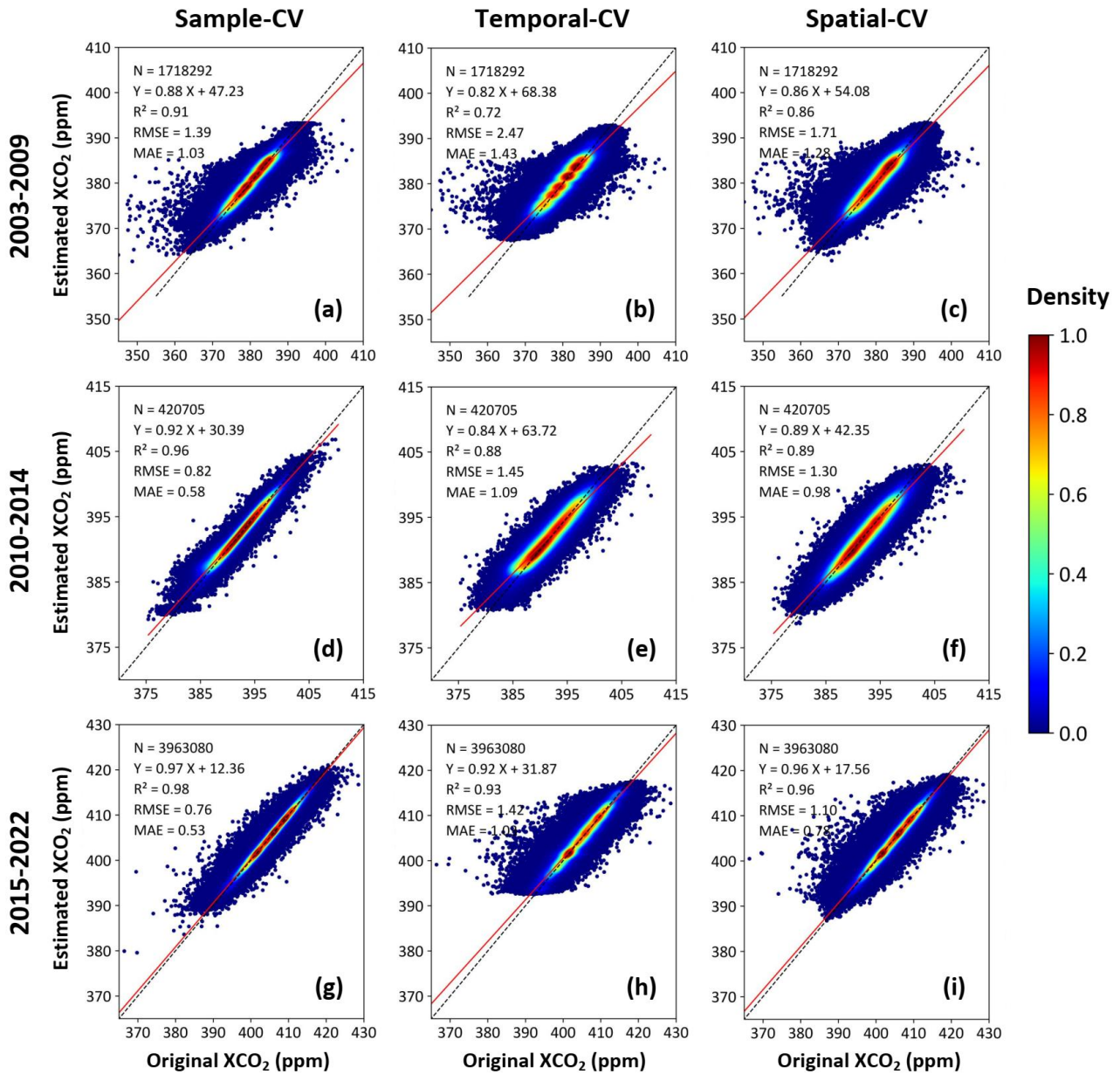
The data-fused XCO₂ estimates were independently evaluated against TCCON XCO₂ measurements. For both the data-fused and bias-corrected XCO₂ products, a multi-strategy ten-fold cross-validation (10-CV) framework was applied using satellite and TCCON observations. Specifically, the sample-based CV randomly withheld 10% of the daily grid-level samples to assess overall performance and internal consistency. Temporal-based CV withheld continuous blocks of daily data to evaluate the model's temporal generalization over periods without ground-based measurements. Spatial-based CV withheld spatially contiguous clusters of grid cells (1° × 1°) to assess the model's capability to generalize in regions lacking ground-based observations.

3 Results and discussion

3.1 Model performance

3.1.1 Validation of data-fused XCO₂ estimates

We first evaluated the Transformer-BiLSTM model in the data-fusion phase using three 10-fold cross-validation approaches, comparing daily XCO₂ estimates with satellite retrievals from SCIAMACHY, GOSAT, and OCO-2 (Fig. 3). Sample-based CV shows that our model accurately reconstructs daily XCO₂ concentrations across different satellite missions, as evidenced by increasing cross-validation R² (CV- R²) values from 0.91 (SCIAMACHY) to 0.96 (GOSAT) and 0.98 (OCO-2), and decreasing RMSE (MAE) values from 1.39 (1.03) to 0.82 (0.58) and 0.76 (0.53) ppm, reflecting improvements due to higher observation density and retrieval quality in later missions. Temporal-based CV further demonstrates stable predictive performance under interannual extrapolation, with CV-R² rising from 0.72 (SCIAMACHY) to 0.88 (GOSAT) and 0.93 (OCO-2), and RMSE (MAE) decreasing from 2.47 (1.43) to 1.45 (1.09) and 1.42 (1.05) ppm, indicating robust generalization across years and satellite transitions. Spatial-based CV confirms strong spatial predictive ability in under-monitored regions, with CV-R² increasing from 0.86 (SCIAMACHY) to 0.89 (GOSAT) and 0.96 (OCO-2), accompanied by RMSE (MAE) reduction from 1.71 (1.28) to 1.30 (0.98) and 1.10 (0.79) ppm.

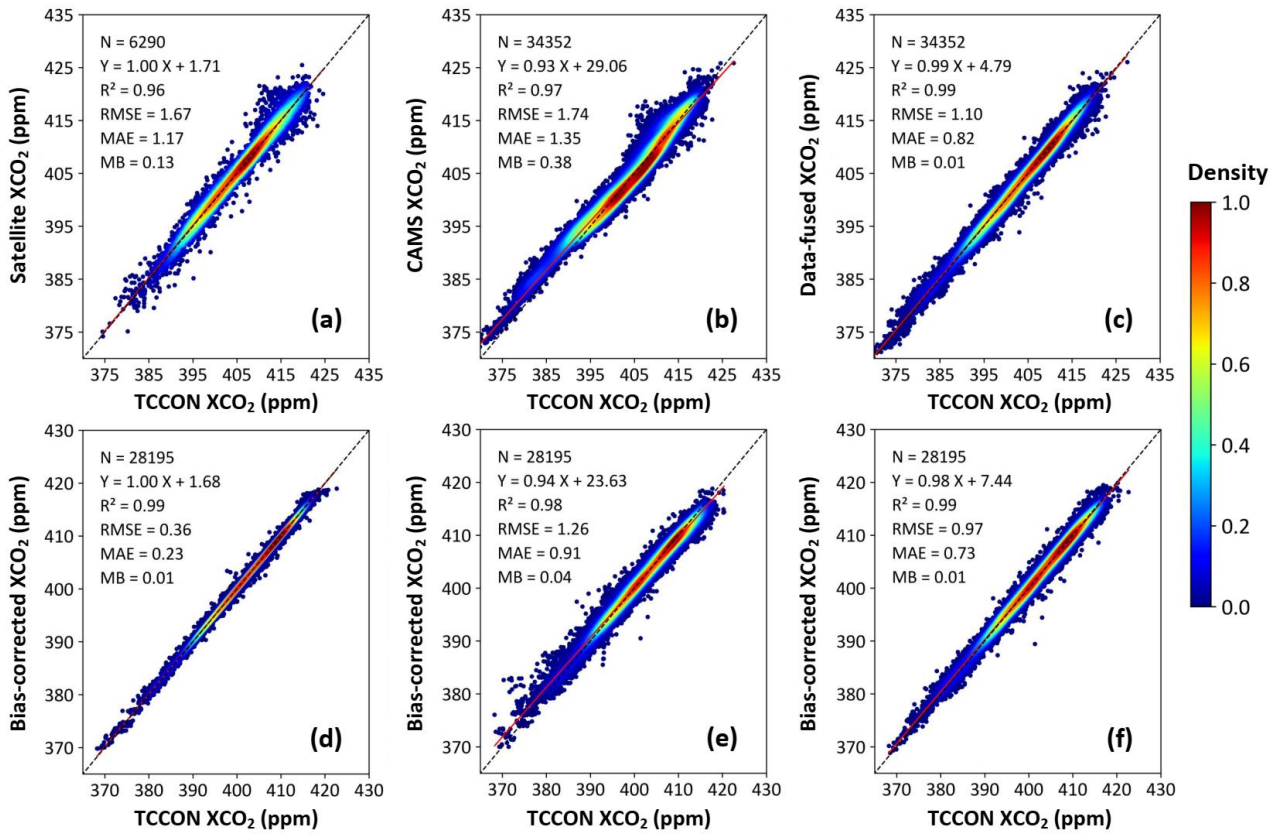


320 **Figure 3.** Density scatter plots of sample-based (left column), temporal-based (middle column), and spatial (right column) 10-fold cross-validation (CV) results for daily XCO₂ estimates during 2003-2009 (SCIAMACHY, top row), 2010-2014 (GOSAT, middle row), and 2015-2022 (OCO-2, bottom row). Dotted black dashed lines represent 1:1 lines, and solid red lines indicate linear regression fits.

325 Furthermore, we independently evaluated our data-fused XCO₂ estimates against TCCON ground-based measurements and compared their performance with satellite retrievals and CAMS reanalysis. Satellite XCO₂ observations show strong agreement with TCCON, yielding an R² of 0.96, an RMSE of 1.67 ppm, and a mean bias of 0.13 ppm (Fig. 4a). CAMS reanalysis exhibits a comparable correlation (R² = 0.97) but larger errors and a more pronounced systematic bias (RMSE = 1.74 ppm, mean bias = 0.38 ppm; Fig. 4b). In contrast, our data-fused XCO₂ achieves the best agreement with

330

TCCON, with an R^2 of 0.99, an RMSE of 1.10 ppm, an MAE of 0.82 ppm, and a near-zero mean bias of 0.01 ppm (Fig. 4c), demonstrating that the fusion framework substantially improves consistency with ground-based observations.



335

Figure 4. Density scatterplots comparing daily XCO₂ concentrations from (a) satellite retrievals, (b) CAMS reanalysis, and (c) our data fusion approach against TCCON ground-based measurements, as well as 10-fold cross-validation (CV) results for our bias-corrected daily XCO₂ concentrations using (d) sample-based, (e) temporal-based, and (f) spatial-based ten-fold CV methods. Dotted black dashed lines represent 1:1 lines, and solid red lines indicate linear regression fits.

340

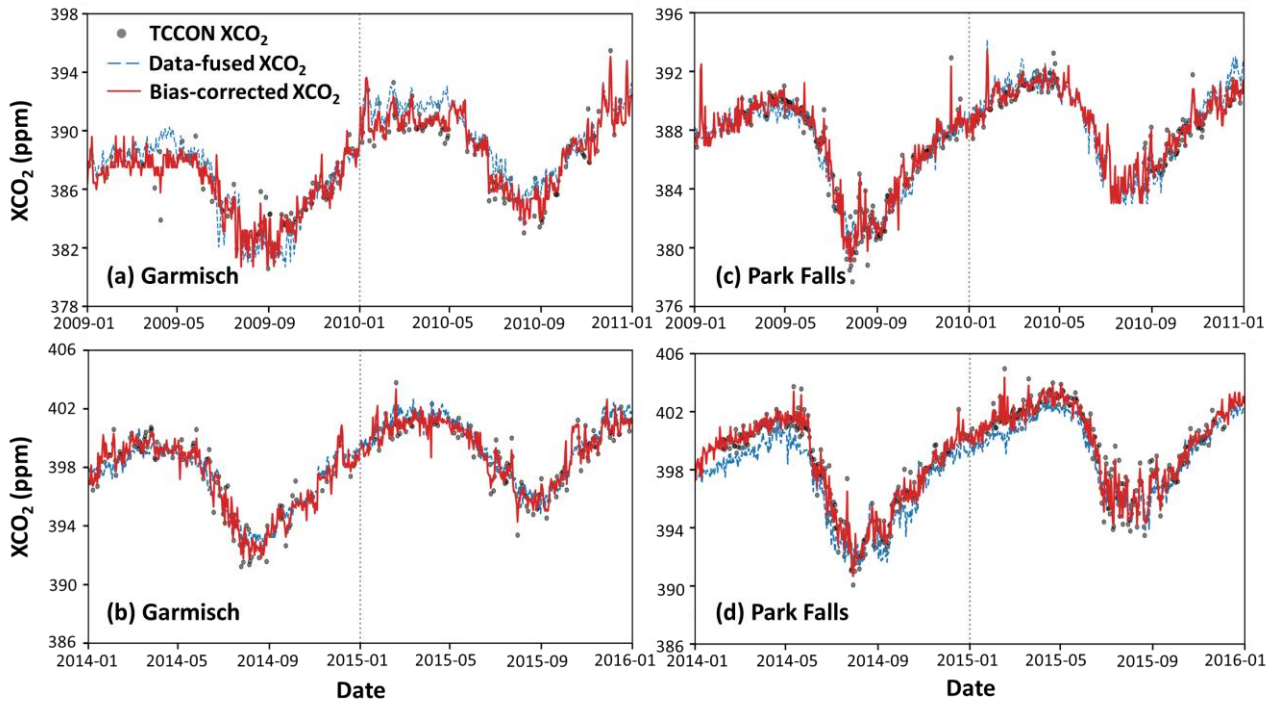
3.1.2 Validation of bias-corrected XCO₂ estimates

The bias-corrected XCO₂ estimates achieve excellent agreement with TCCON observations ($R^2 = 0.99$, RMSE = 1.03 ppm, MAE = 0.79 ppm) and a near-zero mean bias of 0.01 ppm, indicating that systematic biases in the data-fused data are effectively removed (Fig. 4d). Temporal-based CV further demonstrates stable predictive performance under interannual extrapolation: when continuous years are withheld from training, the bias-corrected XCO₂ achieves an R^2 of 0.98, with an RMSE of 1.26 ppm, an MAE of 0.91 ppm, and a mean bias of 0.04 ppm (Fig. 4e), confirming stable temporal generalization across different years. Spatial-based CV confirms strong spatial transferability, with

345

350 an R^2 of 0.99, an RMSE of 0.97 ppm, an MAE of 0.73 ppm, and a mean bias of 0.01 ppm (Fig. 4f), demonstrating reliable performance when extrapolating to under-monitored regions.

Beyond overall statistical accuracy, temporal seamlessness is critical when merging multi-decadal satellite records, particularly across mission transitions where systematic retrieval offsets can
355 introduce artificial discontinuities. To evaluate the temporal stability of the reconstructed dataset across major satellite mission switch-overs (~2010 and ~2015), we conducted targeted diagnostics at long-term TCCON sites, including Garmisch (Europe) and Park Falls (North America) (Fig. 5). Daily XCO_2 time series spanning the SCIAMACHY–GOSAT and GOSAT–OCO-2 transitions reveal clear differences between the initial data-fused product and the bias-corrected reconstruction. At both sites,
360 the uncorrected fused XCO_2 occasionally exhibits small but discernible step changes or systematic offsets coincident with mission switch-overs, indicating residual inter-mission inconsistencies. In contrast, the bias-corrected XCO_2 time series remains temporally smooth and continuous across the transition periods, with no apparent artificial discontinuities. Moreover, the corrected series closely follows TCCON observations in both seasonal phase and amplitude, preserving natural temporal
365 variability while effectively eliminating sensor-induced biases. Together, these results highlight the robustness of the bias correction strategy in ensuring temporal seamlessness of the multi-decadal XCO_2 record.

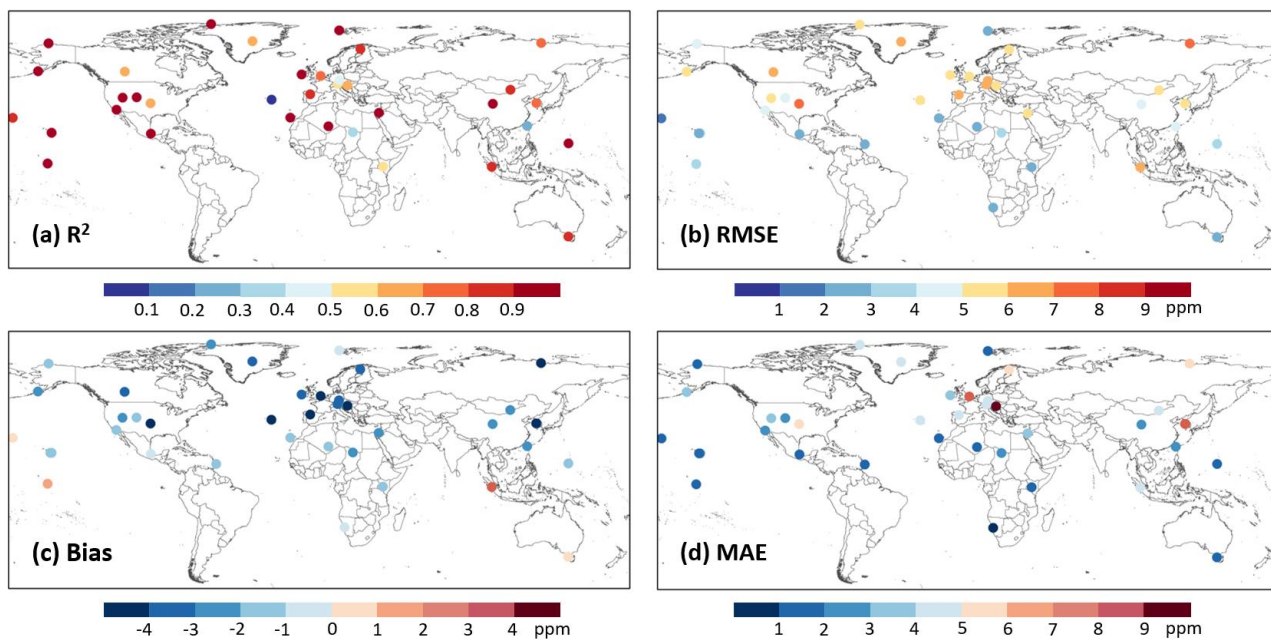


370 **Figure 5.** Time series of TCCON (grey dots), data-fused (blue dashed lines), and bias-corrected (red solid lines) XCO₂ data at the (a-b) Garmisch (Europe) and (c-d) Park Falls (North America) across key satellite transitions: (a) SCIAMACHY to GOSAT (2009-2010) and (b) GOSAT to OCO-2 (2014-2016). The vertical dotted lines mark the approximate times of satellite sensor transitions.

375 3.1.3 Independent validation with ObsPack network

To further evaluate the generalization ability of the bias-corrected daily XCO₂ dataset, we conducted an independent validation against surface flask measurements at 41 ObsPack stations from 2003 to 2022 (Fig. 6). The bias-corrected XCO₂ exhibits strong consistency with ObsPack observations, yielding a global mean R² of 0.86. More than 90% of the stations show high R² values exceeding 0.6, and over half (54%) exceed 0.9, indicating robust performance across diverse geographic and climatic conditions. The low R² 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 R² does not indicate a systematic deficiency of the reconstructed dataset.

380 Spatially, most stations (58%) show RMSE values ranging from 3 to 6 ppm, with larger values observed at stations in North America, Europe, and East Asia, where complex terrain and strong anthropogenic emissions dominate. Nonetheless, the estimated bias and MAE are generally low, with more than 73% and 80% of the stations within ± 3 ppm and below 5.0 ppm, respectively.



390

Figure 6. Global validation of our bias-corrected daily XCO₂ dataset against ObsPack surface flask observations during 2003-2022.

3.2 Spatiotemporal variations of global atmospheric XCO₂

395

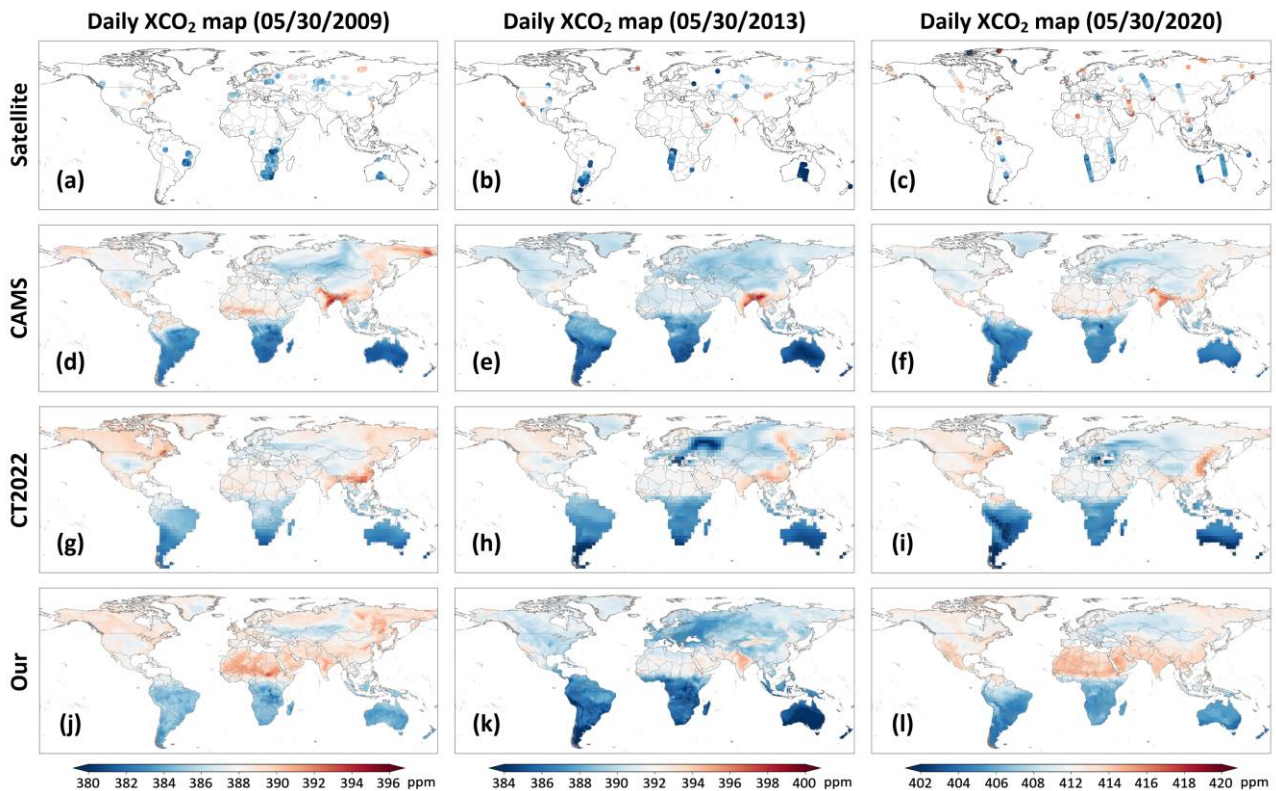
3.2.1 Global daily seamless XCO₂ maps

400

405

We compare global daily XCO₂ data derived from different products for three representative dates: May 30, 2009, 2013, and 2020, where the coarser-resolution products are plotted at their native resolutions for qualitative comparison (Fig. 7). Satellite retrievals provide physically realistic XCO₂ signals but suffer from sparse and discontinuous spatial coverage due to orbital limitations and cloud contamination. CAMS (0.75°×0.75°) offers spatially complete fields but tends to weaken local spatial gradients, attenuating enhancements over major emission regions and underestimating background values in the Southern Hemisphere. CarbonTracker 2022 (CT2022, 3°×2°) is a widely used global CO₂ flux inversion system developed by NOAA to quantify atmospheric CO₂ sources and sinks using transport modeling (Jacobson et al., 2023). It captures large-scale interhemispheric gradients but exhibits regional inconsistencies, including exaggerated concentrations in parts of the Northern Hemisphere mid-latitudes and muted contrasts in tropical source regions. In contrast, our XCO₂ product achieves seamless global coverage while preserving fine-scale spatial heterogeneity and maintaining close consistency with satellite retrievals where available. Representative regional comparisons (highlighted by red circles) further illustrate these differences: In 2009, CT2022

410 overestimates XCO_2 over the eastern United States, while CAMS underestimates regional
 enhancements; our estimates closely match satellite observations. In 2013, both CAMS and CT2022
 underestimate XCO_2 over northern China and southern Mongolia, whereas our product accurately
 captures the observed spatial patterns. In 2020, both CAMS and CT2022 underestimate XCO_2 over
 415 the Middle East, while our reconstruction reflects the enhanced concentrations in agreement with
 satellite observations. These examples demonstrate that our data provides a balanced and physically
 consistent representation of daily global XCO_2 , avoiding CAMS over-smoothing and the regional
 biases evident in CT2022.



420 **Figure 7.** Spatial distribution of global daily XCO_2 concentrations on May 30, 2009, 2013, and 2020
 from (a-c) SCIAMACHY, GOSAT, and OCO-2 satellites, (d-f) CAMS reanalysis, (g-i)
 CarbonTracker (CT2022), and (j-l) our reconstructed product.

3.2.2 Characterization of localized XCO_2 enhancement signals

425 Daily XCO_2 observations provide a distinct advantage for revealing localized and transient regions
 with elevated XCO_2 concentrations that are often obscured in temporally averaged products (Fig. 8).
 Pronounced XCO_2 enhancements are observed over regions with strong anthropogenic influence,
 including major industrial and energy-related areas such as steel production (Yokohama steel factory

in Japan, [f](#)), coal mining (Arizona coal mine in the USA, [a](#)), power generation (electric power plant
430 in Algeria, [i](#)), and chemical processing (Russia, [d](#)). Elevated XCO₂ concentrations are also observed
over industrial clusters in East Asia, Central Asia (Muruntov oltin koni in Uzbekistan, [g](#)), Europe (oil
factory in Poland, [c](#)), North America, and the Middle East (chemical industries in Saudi Arabia, [h](#)).
These patterns are consistent with localized regions of relatively high XCO₂ concentrations associated
with fossil-fuel combustion and extraction-related activities. In addition to stationary sources,
435 localized enhancements are also evident near major transport hubs, such as John F. Kennedy
International Airport in the United States ([b](#)). Beyond anthropogenic sources, the daily XCO₂ dataset
also captures episodic enhancement events associated with natural processes, such as forest fires (e.g.,
Xichang forest fire in China, [e](#)) and large-scale agricultural burning (farmland fire in Brazil, [j](#)), which
can generate short-lived yet spatially coherent XCO₂ anomalies at daily timescales. These examples
440 suggest that the reconstructed daily XCO₂ product can identify typical regions with elevated XCO₂
concentrations associated with both anthropogenic and natural processes, resolve their spatial
footprints, and support the monitoring of regional XCO₂ variability at the global scale. Nevertheless,
these selected examples are intended to illustrate localized elevated XCO₂ signals rather than strict
point-source detections; however, some isolated high-value pixels may partly reflect retrieval
445 uncertainty or noise, especially in regions with limited observations.

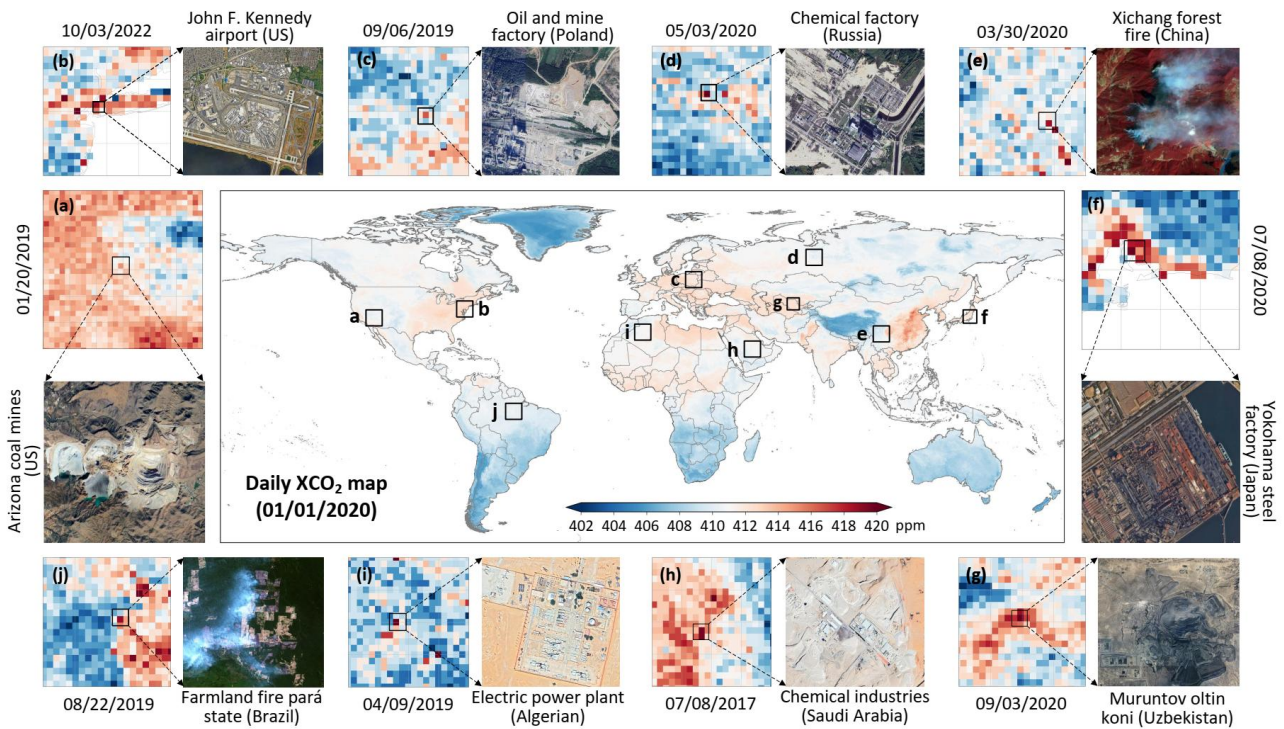
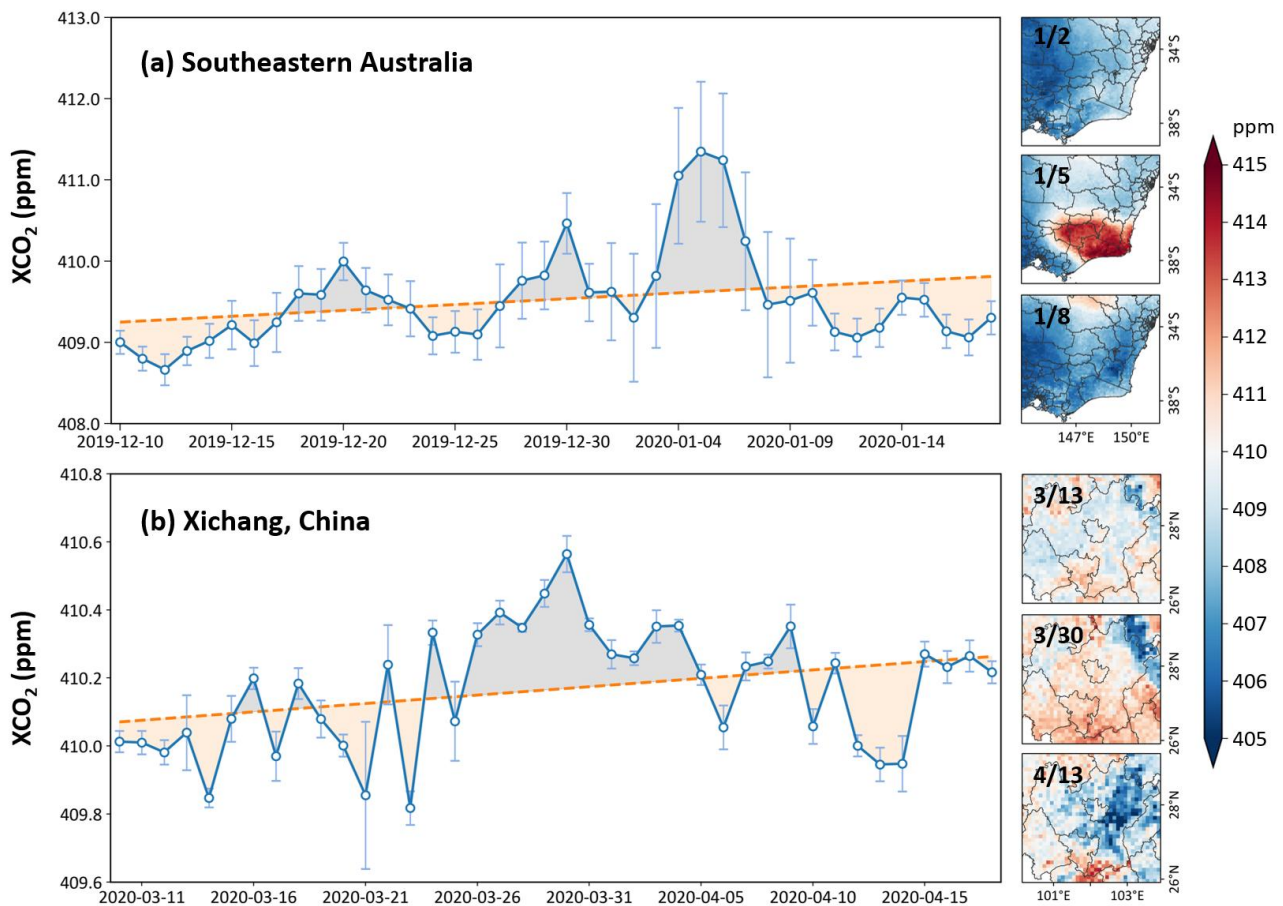


Figure 8. Detection of localized XCO₂ emission hotspots from our reconstructed global daily gapless dataset across (a-j) ten regions of interest on different days, with the central map showing the global distribution on January 1, 2020.

3.2.3 Impacts of biomass burning and ENSO on XCO₂ growth

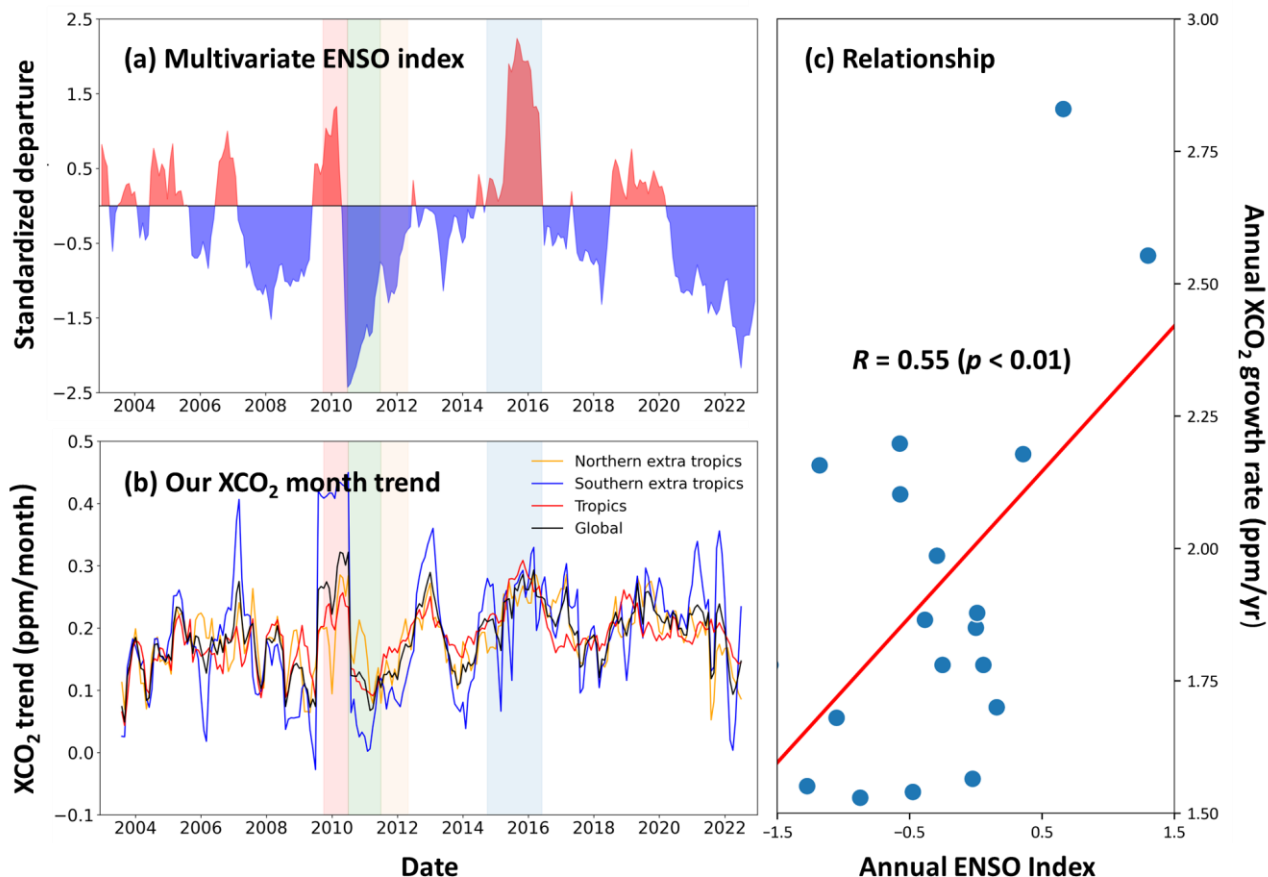
Elevated fire-related CO₂ emissions and corresponding short-term XCO₂ anomalies are clearly captured in our reconstructed dataset during extreme wildfire events. Over southeastern Australia (Fig. 9a), pronounced XCO₂ enhancements are observed from December 27, 2019, to January 8, 2020, coinciding with the peak of the 2019–2020 Australian bushfires. The daily XCO₂ time series exhibits a distinct local maximum on January 5, 2020 (average = 411.35 ± 0.56 ppm), superimposed on a gradually increasing background trend, with spatial distributions revealing significantly elevated concentrations over fire-affected regions. A similar short-term response is detected over Xichang City in southwestern China during the March–April 2020 wildfire episode (Fig. 9b), lasting over two weeks, from March 24 to April 8, 2020, with elevated daily XCO₂ concentrations. A clear local peak of 410.56 ± 0.53 ppm is observed on March 30, coinciding with the reported Xichang forest fire outbreak, followed by a gradual decline after April 5. These event-scale XCO₂ anomalies, resolved at daily resolution, demonstrate the dataset’s ability to capture rapid, short-lived carbon emissions and their impact on regional atmospheric XCO₂ variability.



470 **Figure 9.** Time series and spatial distributions of daily XCO₂ concentrations associated with extreme wildfire events in (a) southeastern Australia (December 2019–January 2020) and (b) Xichang, China (March–April 2020).

475 **Figure 10** shows that XCO₂ growth rates from our reconstructed dataset consistently respond to ENSO phases, as evidenced by the strong correlation with the multivariate ENSO index (MEI; $R = 0.55$, $p < 0.01$) (Wolter and Timlin, 2011), which explains ~30% of the observed interannual variability (Kim et al., 2016; Chatterjee et al., 2017). ENSO-driven anomalies in temperature and precipitation strongly regulate terrestrial carbon uptake and fire activity (Da Costa et al., 2024). During El Niño conditions, warmer and drier climates suppress ecosystem CO₂ absorption and enhance biomass burning, leading to elevated atmospheric growth rates (Betts et al., 2016; Guan et al., 2023; Ak-Bhd, 2021). This effect was most evident during the extreme 2015–2016 El Niño, which 480 coincided with the highest recorded global XCO₂ growth rate of 3.11 ppm/yr ($p < 0.001$) (Liu et al., 2024). In contrast, La Niña phases are generally cooler and wetter, strengthening terrestrial carbon sinks and reduce atmospheric growth rates, as observed in 2011–2012 and 2022 (Wang et al., 2014; Chatterjee et al., 2017). Regionally, ENSO sensitivity is highest in the tropics, where MEI explains

~45% of the interannual variability, highlighting the vulnerability of tropical forests to climate-induced stress and fire-related carbon losses (Brando et al., 2019; Liu et al., 2022). The Northern Extratropics show moderate sensitivity (~37%), while the Southern Extratropics are less affected (~27%), reflecting hemispheric differences in land–atmosphere coupling and fire regimes.



490 **Figure 10.** Time series of the monthly (a) multivariate ENSO index and (b) detrended XCO₂ trends over major regions: Northern extra tropics (30°N to 90°N), Southern extra tropics (90°S to 30°S), Tropics (30°S to 30°N), and Global. (c) Scatter plots between annual XCO₂ growth rates and the annual ENSO index. The La Niña of 2011 (LN11) is shaded in dark green. The preceding El Niño of 2010 (EN10) and succeeding weak La Niña of 2012 (LN12) are shaded in lighter red and green colors. 495 The El Niño of 2015 (EN15) is shaded in dark blue.

3.2.4 Annual maps and long-term trends in global XCO₂

Figure 11 shows the annual mean global XCO₂ from 2003 to 2022 over land, revealing a pronounced and sustained increase over the past two decades. During the early period (2003–2005), global XCO₂ levels are relatively low, with most regions characterized by concentrations below ~380 ppm. From 2006 to 2010, XCO₂ increases steadily across all continents, while spatial contrasts remain relatively

modest. After approximately 2010, the spatial patterns intensify markedly. From 2011 to 2015, high- XCO_2 regions expand and strengthen, and from 2016 onward, the global increase accelerates further, with XCO_2 exceeding ~ 410 ppm over large portions of the Northern Hemisphere by 2018–2019 (Buchwitz et al., 2018; Lee et al., 2025). By 2022, nearly all continental regions display XCO_2 values above 415 ppm, with the highest concentrations concentrated over major industrialized and densely populated regions in the Northern Hemisphere. Overall, the global mean XCO_2 increased monotonically from 374.54 ppm in 2003 to 416.36 ppm in 2022, corresponding to an increase of approximately 11% over the past 20 years. This substantial growth highlights the rapid accumulation of atmospheric CO_2 and the growing influence of anthropogenic emissions on the global carbon cycle.

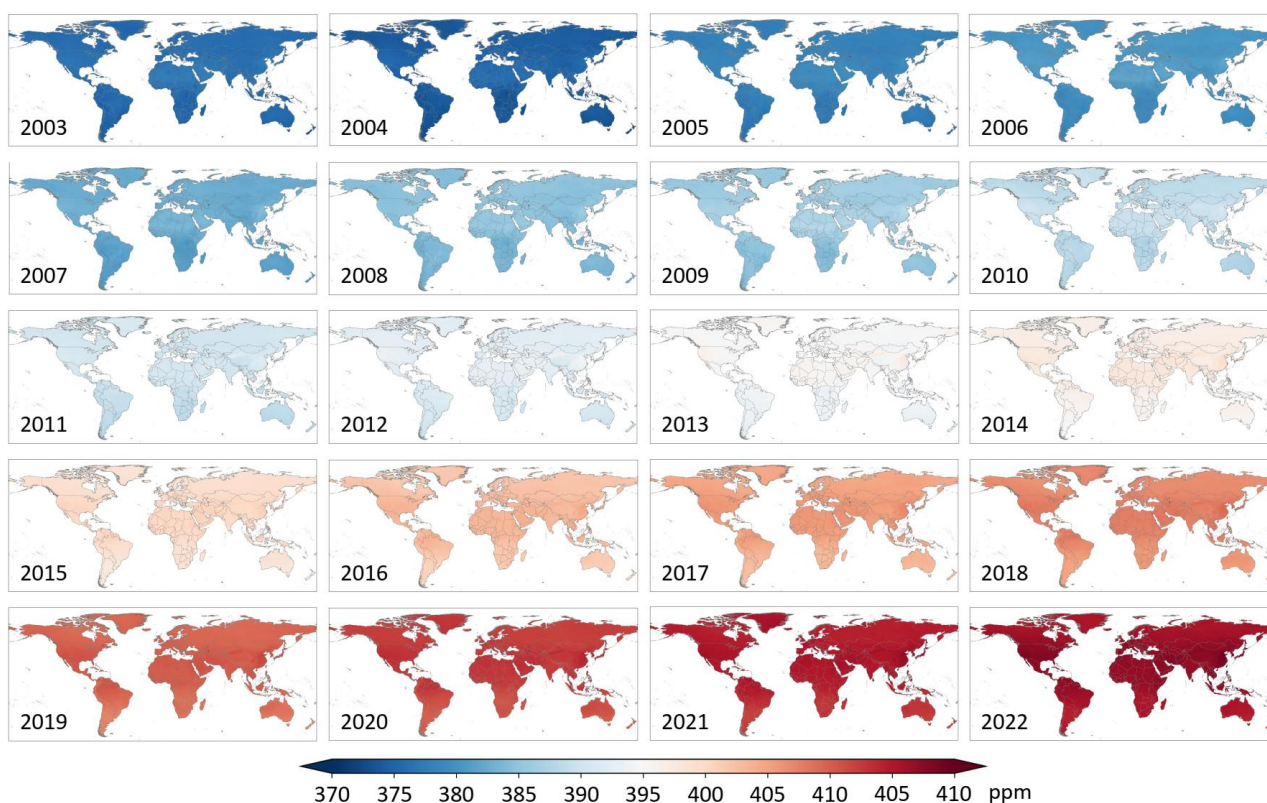


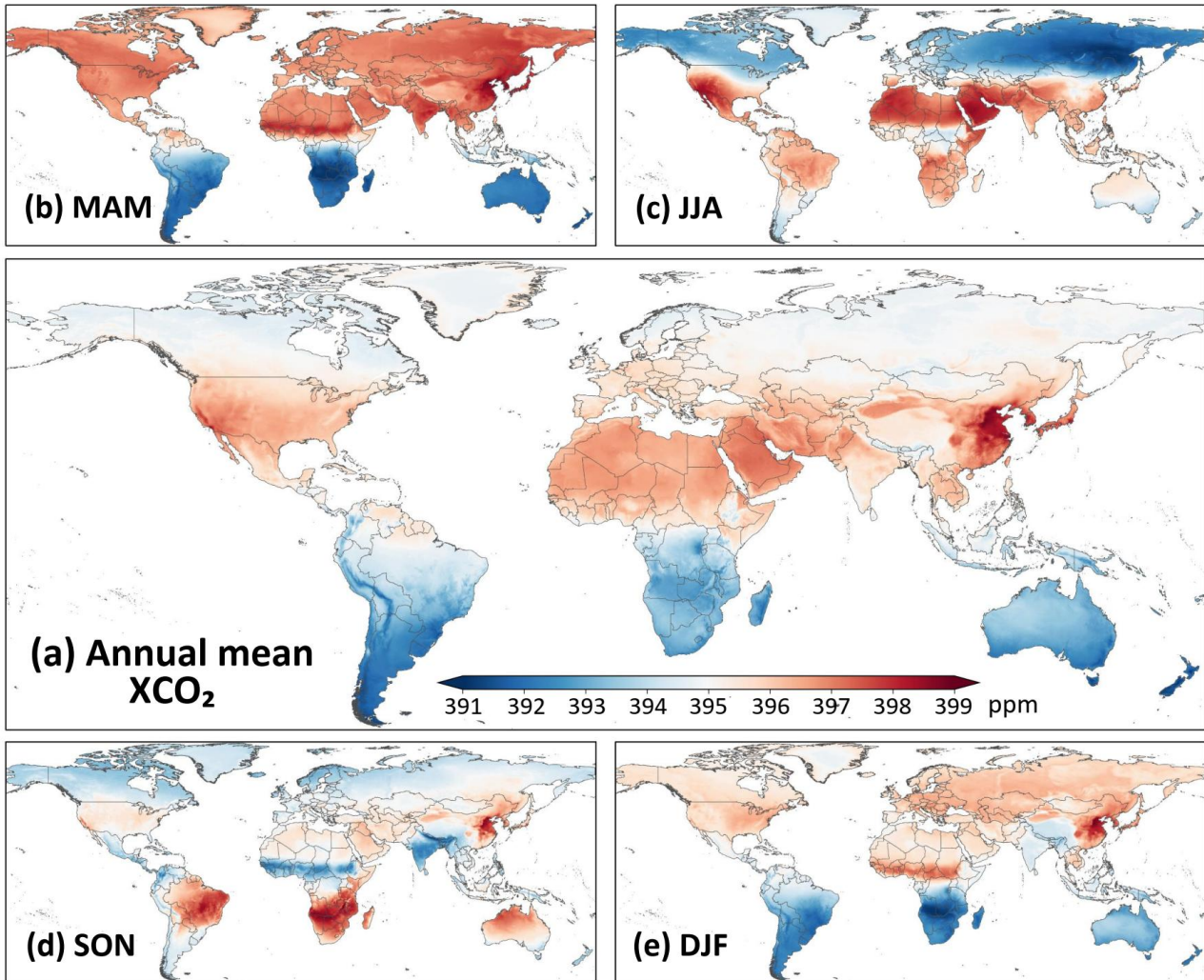
Figure 11. Spatial distribution of annual average XCO_2 concentrations (unit: ppm) over land from 2003 to 2022.

515

[Figure 12](#) presents the long-term (2003–2022) annual mean XCO_2 and seasonal climatology. The long-term global mean XCO_2 is 394.58 ± 0.76 ppm, with a clear interhemispheric gradient characterized by higher concentrations in the Northern Hemisphere and lower values in the Southern. The highest XCO_2 concentrations are concentrated in the northern low- to mid-latitudes (~ 10 – $45^\circ N$),

520 encompassing East and Southeast Asia, northern Africa and the Middle East, the United States, and parts of Europe and South Asia. These regions coincide with dense population centers, intensive industrial activity, and high fossil fuel consumption (Crippa et al., 2021; Sheng et al., 2021). Conversely, persistently low XCO₂ values occur over South America, central and southern Africa, and Australia, where strong biospheric uptake, sparse anthropogenic emissions, and the influence of
525 clean marine air masses suppress atmospheric CO₂ levels. At the national scale, South Korea exhibits the highest XCO₂ value (396.11 ± 0.69 ppm), followed by Kuwait (395.88 ± 0.79 ppm), consistent with intensive energy use and strong fossil fuel dependence.

Pronounced seasonal variability in XCO₂ is evident across global land (Fig. 12b-d). During boreal
530 spring (MAM; average = 395.36 ± 2.20 ppm), elevated XCO₂ concentrations are widespread across the Northern Hemisphere, reflecting accumulated winter emissions and the delayed onset of biospheric uptake. In boreal summer (JJA; average = 392.74 ± 1.29 ppm), XCO₂ decreases markedly over northern high-latitude continents but increases in the Southern Hemisphere, forming extensive low-concentration bands associated with peak vegetation photosynthesis and strong net carbon uptake.
535 Boreal autumn (SON; average = 393.48 ± 0.54 ppm) is characterized by distinct enhancements over southern Africa, South America, and much of Australia, linked to biomass burning and reduced biospheric uptake. Seasonal amplitudes are particularly large in East Asia and the western United States (e.g., California), where intensive anthropogenic emissions coincide with strong biospheric seasonality (Sheng et al., 2021; Guan et al., 2024). In boreal winter (DJF; average = 394.76 ± 1.37
540 ppm), enhanced fossil fuel combustion and dormant vegetation lead to higher XCO₂ concentrations re-emerge across the Northern Hemisphere, whereas lower values prevail over Southern Hemisphere land and adjacent oceans, resulting in a pronounced hemispheric contrast.



545 **Figure 12.** Multi-year (a) annual mean XCO₂ map (unit: ppm) over land for 2003–2022, and seasonal mean XCO₂ maps for (b) March–April–May (MAM), (c) June–July–August (JJA), (d) September–October–November (SON), and (e) December–January–February (DJF).

550 **Figure 13** presents long-term XCO₂ trends from 2003 to 2022 across different temporal scales. All land regions exhibit significant increasing trends, with an average of 2.24 ppm/yr ($p < 0.001$). The strongest increases (>2.25 ppm/yr) occur over East and Southeast Asia, northern South America, and central Africa, while comparatively weaker growth (<2.15 ppm/yr) is observed over southern South America, Australia, and the Sahara. Seasonal trends reveal distinct patterns (Fig. 13b-e). During MAM (average = 2.16 ppm/yr, $p < 0.001$), large positive trends dominate the northern mid-latitudes, especially across much of Asia and North America, reflecting winter-emission accumulation and delayed biospheric uptake. In JJA (average = 2.23 ppm/yr, $p < 0.001$), trends weaken over boreal and temperate regions but remain pronounced in the tropics, particularly the Amazon basin, central Africa, and Southeast Asia, corresponding to ongoing anthropogenic emissions and biomass burning. SON

(average = 2.25 ppm/yr, $p < 0.001$) shows enhanced trends over Africa and South America, forming clear regional maxima in biomass-burning regions, while northern mid-latitudes experience moderate growth. In DJF (average = 2.19 ppm/yr, $p < 0.001$), trends intensify again across the Northern Hemisphere, notably in East Asia and Europe, driven by wintertime fossil fuel combustion, whereas the Southern Hemisphere shows weaker increases. Although the spring and winter seasons exhibit stronger regional contrasts, the autumn season shows a systematically higher spatially averaged XCO₂ growth trend, indicating a more widespread enhancement rather than localized extremes.

565

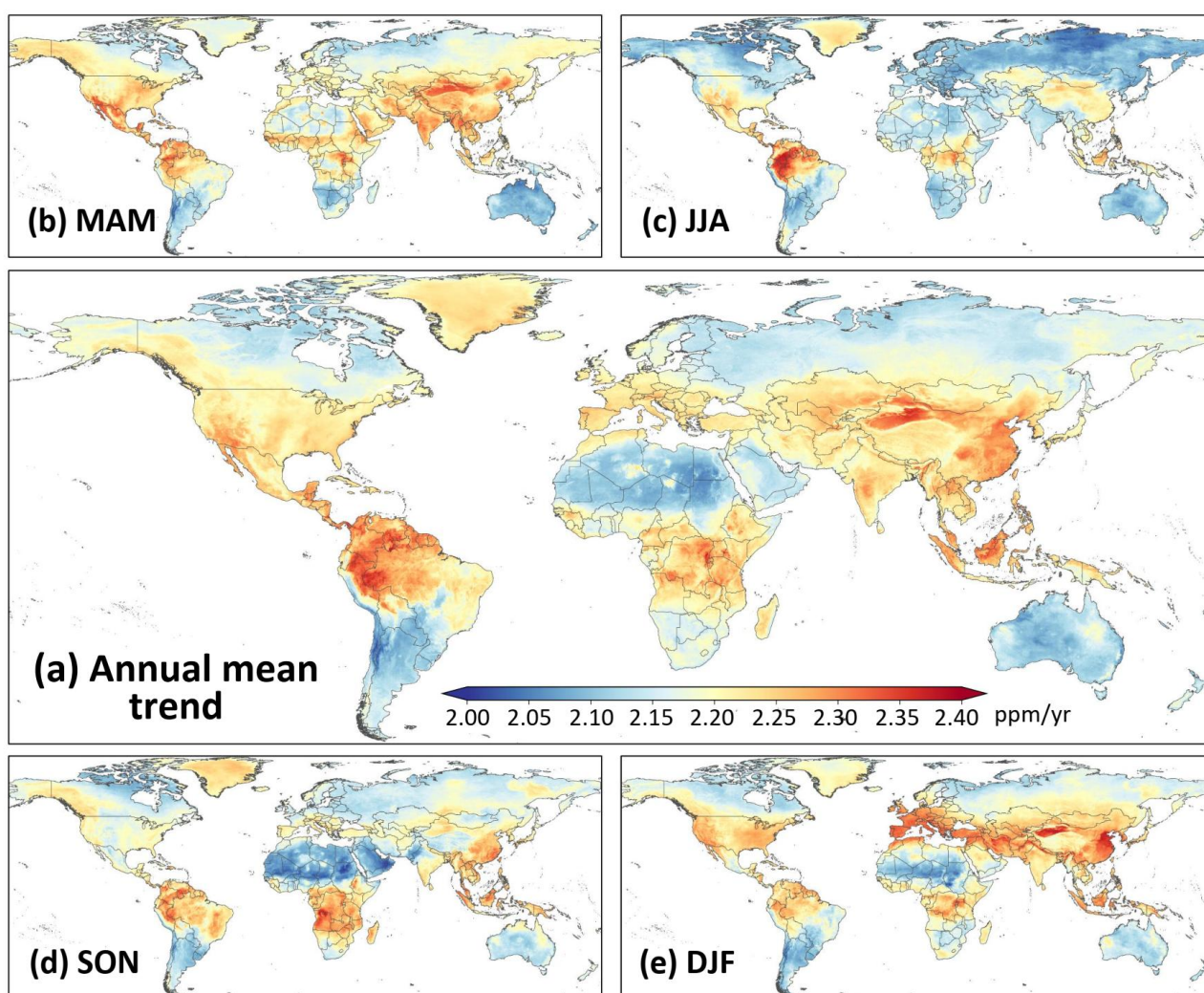


Figure 13. Spatial patterns of long-term annual XCO₂ trends (unit: ppm/yr) from 2003 to 2022, and seasonal XCO₂ trends (unit: ppm/yr) for (b) March–April–May (MAM), (c) June–July–August (JJA), (d) September–October–November (SON), and (e) December–January–February (DJF).

570

3.3 Discussion

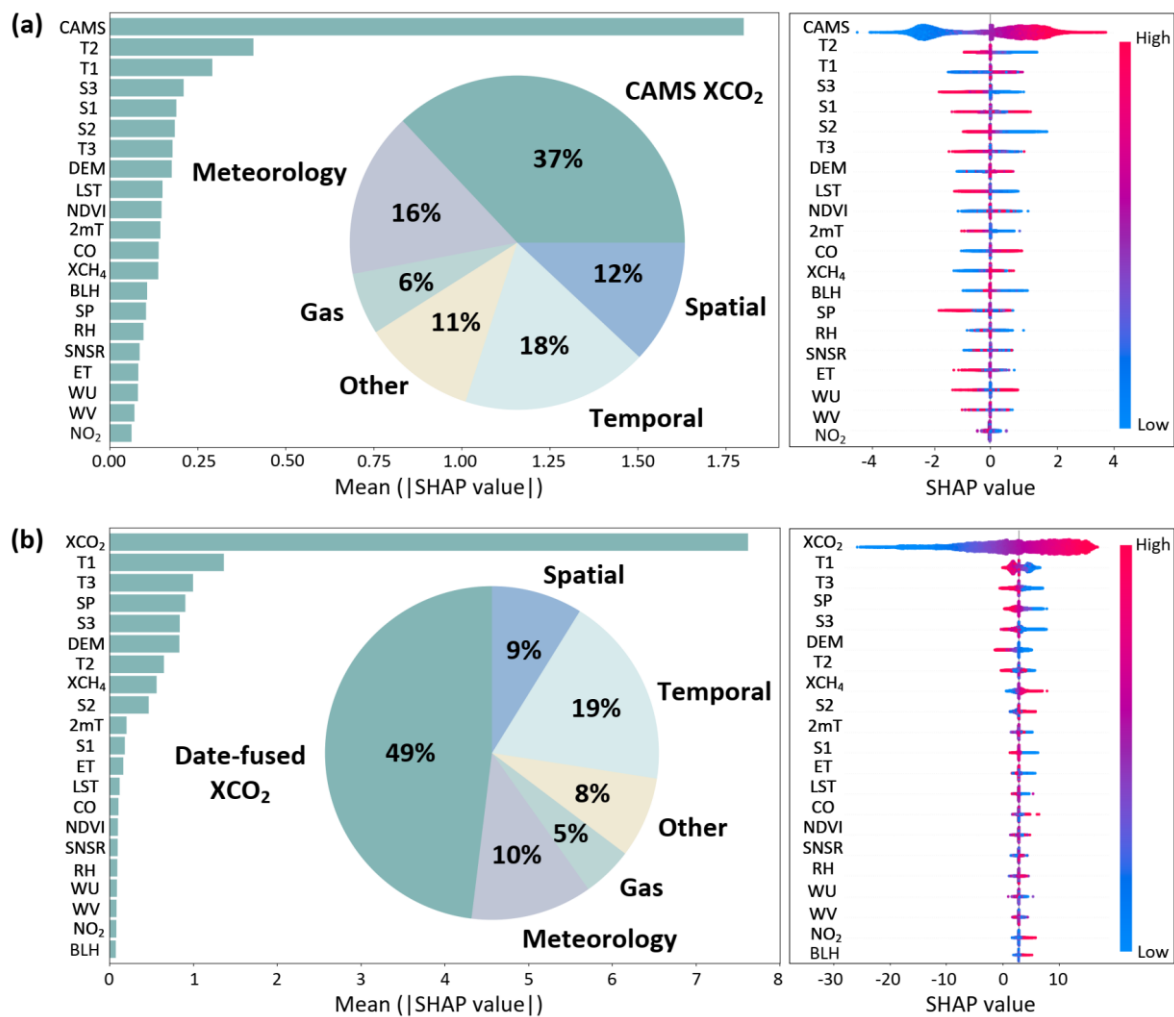
3.3.1 Model interpretation with XAI

To improve model transparency and quantify the relative importance of input variables, we applied SHapley Additive exPlanations (SHAP) to the trained spatiotemporal Transformer-BiLSTM model (Lundberg and Lee, 2017). The SHAP values were calculated to assess the contribution of each predictor to the reconstructed XCO₂ fields. For the data-fusion stage (Fig. 14a), the CAMS XCO₂ emerges as the dominant contributor, accounting for 37% of the total importance, reflecting its role in providing a temporally continuous and spatially coherent background field. Spatiotemporal encoding variables collectively contribute 30%, highlighting the importance of explicitly representing spatial structure and temporal dynamics in global XCO₂ reconstruction. Meteorological variables (16%) and satellite-derived surface variables (11%) provide complementary information that refines regional-scale variability, while precursor gases contribute a smaller but non-negligible share (6%). The relatively small mean SHAP value of NO₂ suggests that its overall contribution to the global XCO₂ reconstruction is smaller compared with variables such as CAMS XCO₂ and meteorological factors. This is expected because NO₂ mainly reflects localized anthropogenic combustion emissions, whereas XCO₂ variability at the global scale is more strongly controlled by large-scale atmospheric transport, background concentration fields, and biospheric exchange processes. Nevertheless, NO₂ still provides useful complementary information for identifying regional anthropogenic emission patterns, particularly over urban and industrial areas.

590

For the bias-correction stage (Fig. 14b), the fused XCO₂ field dominates the model output, contributing 49%, consistent with its role as the primary constraint for harmonizing inter-satellite differences using TCCON observations. Temporal indicators (19%), spatial variables (9%), and meteorological variables (10%) further capture systematic temporal, spatial, and environmental patterns in satellite retrieval biases. Overall, the XAI analysis indicates that the proposed framework is primarily constrained by physically meaningful background information, while auxiliary variables are effectively integrated to refine spatiotemporal XCO₂ structures and reduce systematic biases.

595



600 **Figure 14.** Model interpretability analysis using the SHAP approach, showing the relative importance of each predictive factor for the (a) data-fused stage and (b) bias-corrected stage.

3.3.2 Comparison with previous studies

We compared our reconstructed XCO₂ dataset with previous global studies that applied independent validation against TCCON observations (Table 3). Some studies are limited to coarser temporal resolutions at monthly (Zhang et al., 2023; Wang et al., 2025; Hwang et al., 2026; Yu et al., 2026), 8-day (Li et al., 2022; Guan et al., 2024), or 3-day (Sheng et al., 2023) scales, which limits their ability to capture short-term (daily) variability. In addition, many studies generated XCO₂ at coarser spatial resolutions ($\geq 0.25^\circ$), limiting their representation of localized emission gradients and subregional heterogeneity (Li et al., 2022; Sheng et al., 2023; Jin et al., 2022; Gao et al., 2023; Huang et al., 2024c; Wang et al., 2023). Many products cover only limited periods (usually <10 years) after 2014, constraining analyses of long-term historical trends (Sheng et al., 2023; Huang et al., 2024c; Li et al., 2026; Lee et al., 2025; Yu et al., 2026). More importantly, previous studies on fine-temporal resolution

605

610

(daily or subdaily) exhibit larger fluctuations in overall accuracy, with R^2 ranging from 0.91 to 0.98 and RMSEs from 1.06 to 2.62 ppm. However, a recent study (Wang, 2026) provides a long-term global daily XCO₂ product at a 1 km spatial resolution with considerable accuracy ($R^2 = 0.98$, RMSE = 1.10 ppm). In contrast, our Transformer–BiLSTM framework reconstructs a global, seamless, cross-mission-consistent, daily XCO₂ dataset at a 0.1° spatial resolution over 2003–2022, achieving a stronger or comparable R^2 of 0.99 and RMSE of 1.10 ppm.

Table 3. Comparison of reconstructed global XCO₂ datasets from previous studies.

Model	Temporal	Spatial	R^2	RMSE	Time frame	Literature
DINEOF-BME	Monthly	1°	0.92	1.60	2010-2020	Gao et al. (2023)
LASSO	Monthly	0.5°	0.98	1.08	2003-2022	Yang et al. (2026)
LightGBM	Monthly	0.1°	0.98	1.17	2003-2019	Hwang et al. (2026)
Deep forest	Monthly	0.1°	0.96	0.95	2014-2020	Zhang et al. (2023)
At-Bi-LSTM	Monthly	0.05°	0.92	1.54	2015-2021	Wang et al. (2025)
Extreme tree	8-day	0.01°	0.83	1.79	2014-2018	Li et al. (2022)
Stacking model	8-day	0.05°	0.98	1.03	2000-2020	Guan et al. (2024)
Kriging	3-day	1°	0.97	0.88	2009-2023	Sheng et al. (2023)
M-FRK	Daily	1°	0.93	1.06	2016-2019	Huang et al. (2024c)
Self-supervised	Daily	0.25°	0.93	1.18	2010-2020	Wang et al. (2023)
Random Forest	Daily	0.25°	0.95	1.05	2018-2023	Lee et al. (2025)
Stacking model	Daily	0.1°	0.94	1.08	2016-2020	Li et al. (2026)
DINCAE	Daily	0.1°	0.90	1.44	2018-2022	Antezana Lopez et al. (2025)
Random Forest	Daily	1 × 2 km	0.91	1.09	2020-2023	Yu et al. (2026)
ConvLSTM + Unet	Daily	1 km	0.98	1.10	2003-2023	Wang (2026)
MLE	3-h	0.5°	0.92	2.62	2003-2020	Jin et al. (2022)
Transformer–BiLSTM	Daily	0.1°	0.99	1.10	2003-2022	This study

DINEOF-BME: Data Interpolation Empirical Orthogonal Function- Bayesian Maximum Entropy; At-Bi-LSTM: Attention bidirectional long short-term memory; LASSO: Least Absolute Shrinkage and Selection Operator (LASSO) regression; M-FRK: Multiscale fixed rank kriging; DINCAE: Data Interpolating Empirical Orthogonal Functions and Convolutional Auto-Encoder; MLE: Maximum likelihood estimation.

3.3.3 Strengths

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.

665

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
670 regions and atmospheric conditions.

670

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
675 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.

675

680 **4 Conclusions**

Accurate, high-resolution observations of atmospheric XCO₂ are critical for understanding the global carbon cycle, quantifying anthropogenic emissions, and supporting climate mitigation strategies. This study develops a novel Transformer–BiLSTM framework to reconstruct a global, daily, and spatially seamless XCO₂ dataset over land at 0.1° resolution covering 2003–2022. By jointly integrating multi-
685 mission satellite retrievals, CAMS reanalysis, meteorological variables, and precursor gas information, the framework effectively bridges spatial gaps and temporal discontinuities inherent in individual satellite records. A dedicated spatiotemporal loss formulation enforces continuity across space and time, and a bias-correction strategy is applied to harmonize inter-satellite differences and generate a physically consistent long-term XCO₂ record.

690

Comprehensive validation demonstrates that the reconstructed dataset achieves high accuracy and robustness, as evidenced by independent evaluation against TCCON observations after data fusion ($R^2 = 0.99$, RMSE = 1.10 ppm), and ten-fold cross-validation after bias correction ($CV-R^2 = 0.99$, RMSE = 0.97 ppm). Using the daily XCO₂ data, we are able to capture rapid variations and identify
695 localized XCO₂ emission hotspots from diverse sources, including urban activities, industrial emissions, and biomass burning. The dataset also effectively reflects strong responses of XCO₂ growth to major climate events, such as intense El Niño episodes and large-scale wildfire events. Across temporal scales, XCO₂ exhibits a long-term increasing trend of ~2.24 ppm/yr ($p < 0.001$) from 2003 to 2022, with the larger growth rates observed over northern South America, East and Southeast
700 Asia, and central Africa, and weaker growth over southern South America, Australia, and the Sahara. This high-resolution, daily global land XCO₂ dataset provides a robust tool for analyzing both long-term trends and short-term variability, offering valuable insights for carbon-cycle research, emission monitoring, and climate mitigation assessments. Future improvements may include enhancing XCO₂ retrievals using higher-resolution satellite observations, better representing natural and, in particular,
705 urban emissions, and expanding ground-based validation networks.

Data availability

The SCIAMACHY, GOSAT XCO₂, and OCO-2/3 XCO₂ datasets are available at <https://data.ceda.ac.uk/neodc/esacci/ghg/data>, <https://disc.gsfc.nasa.gov/>, and
710 <https://disc.gsfc.nasa.gov/>, respectively. The Copernicus Climate Data Store provides the CAMS-EGG4 XCO₂ product at <https://ads.atmosphere.copernicus.eu/datasets>, and ERA5 reanalysis data are available at <https://cds.climate.copernicus.eu/datasets>. The global daily high-resolution XCH₄ dataset (GlobalHighXCH₄) is available at <https://doi.org/10.5281/zenodo.15492399>. MODIS products can be accessed at <https://ladsweb.modaps.eosdis.nasa.gov/search>. The global daily high-resolution
715 XCO₂ product (GlobalHighXCO₂) is publicly available at <https://doi.org/10.5281/zenodo.18220961> (Qu and Wei, 2026).

Financial support

This work was supported by the Jing-Jin-Ji Regional Integrated Environmental Improvement-
720 National Science and Technology Major Project (2026ZD1212900), the Fundamental and Interdisci

plinary Disciplines Breakthrough Plan of the Ministry of Education of China (JYB2025XD XM906), the National Key Technology and Development Program of Corps (2025AA001), the Fundamental Research Funds for the Central Universities, Peking University, and the Scientific Research Innovation Project of Graduate School of South China Normal University (2025KYLX064).

725

Acknowledgments

The authors acknowledge the data providers whose datasets were used in this study.

Author contributions

730 YQ, XS, YF, ZW collected the data, conducted the formal analysis, performed the research, and wrote the initial draft of this paper. JW managed project administration, supervised the project, acquired funding.

Competing interests

735 At least one of the (co-)authors is a member of the editorial board of *Earth System Science Data*.

References:

- Agustí-Panareda, A., Barré, J., Massart, S., Inness, A., Aben, I., Ades, M., Baier, B. C., Balsamo, G., Borsdorff, T., and Bousserez, N.: The CAMS greenhouse gas reanalysis from 2003 to 2020, Atmospheric Chemistry and Physics, 23, 3829-3859, 2023.
- 740 Ak-Bhd, M.: WMO greenhouse gas bulletin, World Meteorological Organization: Geneva, Switzerland, 2021.
- Antezana Lopez, F. P., Zhou, G., Jing, G., Zhang, K., Chen, L., Chen, L., and Tan, Y.: Global Daily Column Average CO₂ at 0.1° × 0.1° Spatial Resolution Integrating OCO-3, GOSAT, CAMS with EOF and Deep Learning, Scientific Data, 12, 268, 2025.
- 745 Betts, R. A., Jones, C. D., Knight, J. R., Keeling, R. F., and Kennedy, J. J.: El Niño and a record CO₂ rise, Nature Climate Change, 6, 806-810, 2016.
- Bovensmann, H., Burrows, J., Buchwitz, M., Frerick, J., Noel, S., Rozanov, V., Chance, K., and Goede, A.: SCIAMACHY: Mission objectives and measurement modes, Journal of the atmospheric sciences, 56, 127-150, 1999.
- 750 Brando, P. M., Paolucci, L., Ummenhofer, C. C., Ordway, E. M., Hartmann, H., Cattau, M. E., Rattis, L., Medjibe, V., Coe, M. T., and Balch, J.: Droughts, wildfires, and forest carbon cycling: A pantropical synthesis, Annual Review of Earth and Planetary Sciences, 47, 555-581, 2019.
- Buchwitz, M., De Beek, R., Noël, S., Burrows, J., Bovensmann, H., Bremer, H., Bergamaschi, P., Körner, S., and Heimann, M.: Carbon monoxide, methane and carbon dioxide columns retrieved from SCIAMACHY by WFM-DOAS: year 2003 initial data set, Atmospheric Chemistry and Physics, 5, 3313-3329, 2005.
- 755 Buchwitz, M., Reuter, M., Schneising, O., Boesch, H., Guerlet, S., Dils, B., Aben, I., Armante, R.,

- Bergamaschi, P., and Blumenstock, T.: The Greenhouse Gas Climate Change Initiative (GHG-CCI): Comparison and quality assessment of near-surface-sensitive satellite-derived CO₂ and CH₄ global data sets, *Remote Sensing of Environment*, 162, 344-362, 2015.
- Buchwitz, M., Reuter, M., Schneising, O., Noël, S., Gier, B., Bovensmann, H., Burrows, J. P., Boesch, H., Anand, J., and Parker, R. J.: Computation and analysis of atmospheric carbon dioxide annual mean growth rates from satellite observations during 2003–2016, *Atmospheric Chemistry and Physics*, 18, 17355-17370, 2018.
- Budget, G. C.: Global carbon budget 2023, 2023.
- Butz, A., Guerlet, S., Hasekamp, O., Schepers, D., Galli, A., Aben, I., Frankenberg, C., Hartmann, J. M., Tran, H., and Kuze, A.: Toward accurate CO₂ and CH₄ observations from GOSAT, *Geophysical Research Letters*, 38, 2011.
- Chatterjee, A., Gierach, M., Sutton, A., Feely, R., Crisp, D., Eldering, A., Gunson, M., O'Dell, C., Stephens, B., and Schimel, D.: Influence of El Niño on atmospheric CO₂ over the tropical Pacific Ocean: Findings from NASA's OCO-2 mission, *Science*, 358, eaam5776, 2017.
- Chen, J., Hu, R., Chen, L., Liao, Z., Che, L., and Li, T.: Multi-sensor integrated mapping of global XCO₂ from 2015 to 2021 with a local random forest model, *ISPRS Journal of Photogrammetry and Remote Sensing*, 208, 107-120, 2024.
- Crippa, M., Guizzardi, D., Pisoni, E., Solazzo, E., Guion, A., Muntean, M., Florczyk, A., Schiavina, M., Melchiorri, M., and Hutfilter, A. F.: Global anthropogenic emissions in urban areas: patterns, trends, and challenges, *Environmental Research Letters*, 16, 074033, 2021.
- Crisp, D.: Measuring atmospheric carbon dioxide from space with the Orbiting Carbon Observatory-2 (OCO-2), *Earth observing systems xx*, 960702,
- Crisp, D., Pollock, H. R., Rosenberg, R., Chapsky, L., Lee, R. A., Oyafuso, F. A., Frankenberg, C., O'Dell, C. W., Bruegge, C. J., and Doran, G. B.: The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products, *Atmospheric Measurement Techniques*, 10, 59-81, 2017.
- da Costa, L. M., de Araújo Santos, G. A., de Mendonça, G. C., de Souza Maria, L., da Silva Jr, C. A., Panosso, A. R., and La Scala Jr, N.: Exploring CO₂ anomalies in Brazilian biomes combining OCO-2 & 3 data: Linkages to wildfires patterns, *Advances in Space Research*, 73, 4158-4174, 2024.
- Gao, Z., Jiang, Y., He, J., and Wu, J.: Spatiotemporal variation analysis of global XCO₂ concentration during 2010–2020 based on DINEOF-BME framework and wavelet function, *Science of The Total Environment*, 892, 164750, 2023.
- Guan, X., Sun, Z., Chu, D., Xie, G., Wang, Y., and Shen, H.: Long-term (2000–2020) global 0.05° continuous atmospheric carbon dioxide mapping combining OCO-2 observations and model simulations, *Science of The Total Environment*, 957, 177051, 2024.
- Guan, Y., Keppel-Aleks, G., Doney, S. C., Petri, C., Pollard, D., Wunch, D., Hase, F., Ohyama, H., Morino, I., and Notholt, J.: Characteristics of interannual variability in space-based XCO₂ global observations, *Atmospheric Chemistry and Physics*, 23, 5355-5372, 2023.
- Hammerling, D. M., Michalak, A. M., and Kawa, S. R.: Mapping of CO₂ at high spatiotemporal resolution using satellite observations: Global distributions from OCO-2, *Journal of Geophysical Research: Atmospheres*, 117, 2012.
- He, C., Ji, M., Li, T., Liu, X., Tang, D., Zhang, S., Luo, Y., Grieneisen, M. L., Zhou, Z., and Zhan, Y.: Deriving full-coverage and fine-scale XCO₂ across China based on OCO-2 satellite retrievals

- and CarbonTracker output, *Geophysical Research Letters*, 49, e2022GL098435, 2022.
- 805 He, Q., Ye, T., Chen, X., Dong, H., Wang, W., Liang, Y., and Li, Y.: Full-coverage mapping high-resolution atmospheric CO₂ concentrations in China from 2015 to 2020: Spatiotemporal variations and coupled trends with particulate pollution, *Journal of Cleaner Production*, 428, 139290, 2023a.
- 810 He, S., Yuan, Y., Wang, Z., Luo, L., Zhang, Z., Dong, H., and Zhang, C.: Machine learning model-based estimation of XCO₂ with high spatiotemporal resolution in china, *Atmosphere*, 14, 436, 2023b.
- He, Z., Fan, G., Li, X., Gong, F.-Y., Liang, M., Gao, L., and Zhou, M.: Spatio-temporal modeling of satellite-observed CO₂ columns in China using deep learning, *International Journal of Applied Earth Observation and Geoinformation*, 129, 103859, 2024.
- 815 He, Z., Lei, L., Zhang, Y., Sheng, M., Wu, C., Li, L., Zeng, Z.-C., and Welp, L. R.: Spatio-temporal mapping of multi-satellite observed column atmospheric CO₂ using precision-weighted kriging method, *Remote Sensing*, 12, 576, 2020.
- Hua, Y., Zhao, X., Sun, W., and Sun, Q.: Satellite-Based Reconstruction of Atmospheric CO₂ Concentration over China Using a Hybrid CNN and Spatiotemporal Kriging Model, *Remote Sensing*, 16, 2433, 2024.
- 820 Huang, N., Gu, L., Black, T. A., Wang, L., and Niu, Z.: Remote sensing-based estimation of annual soil respiration at two contrasting forest sites, *Journal of Geophysical Research: Biogeosciences*, 120, 2306-2325, 2015.
- Huang, X., Deng, Z., Jiang, F., Zhou, M., Lin, X., Liu, Z., and Peng, M.: Improved consistency of satellite XCO₂ retrievals based on machine learning, *Geophysical Research Letters*, 51, e2023GL107536, 2024a.
- 825 Huang, X., Yang, H., Lv, Q., Fan, H., Cui, L., Qiao, Y., Yao, Y., and Feng, G.: A full-coverage daily XCO₂ dataset in China from 2015 to 2020 based on DSC-DF-LGB, *Earth System Science Data Discussions*, 2024, 1-22, 2024b.
- 830 Huang, Y., Wang, R., Ju, M., Zhu, X., and Xie, Y.: Reconstructing global daily XCO₂ at 1× 1 spatial resolution from 2016 to 2019 with multisource satellite observation data, *Journal of Applied Remote Sensing*, 18, 028502-028502, 2024c.
- Hwang, S., Choi, H., Kang, Y., and Im, J.: Reconstructing long-term (2003–2019) global high-resolution XCO₂: bridging observational gaps with machine learning, *GIScience & Remote Sensing*, 63, 2627042, 2026.
- 835 Jacobson, A., Schuldt, K., Tans, P., Andrews, A., Miller, J., Oda, T., Mund, J., Weir, B., Ott, L., and Aalto, T.: CarbonTracker CT2022, NOAA Global Monitoring Laboratory, 2023.
- Jin, C., Xue, Y., Jiang, X., Zhao, L., Yuan, T., Sun, Y., Wu, S., and Wang, X.: A long-term global XCO₂ dataset: Ensemble of satellite products, *Atmospheric Research*, 279, 106385, 2022.
- 840 Kemp, L., Xu, C., Depledge, J., Ebi, K. L., Gibbins, G., Kohler, T. A., Rockström, J., Scheffer, M., Schellnhuber, H. J., and Steffen, W.: Climate endgame: Exploring catastrophic climate change scenarios, *Proceedings of the National Academy of Sciences*, 119, e2108146119, 2022.
- Kim, J.-S., Kug, J.-S., Yoon, J.-H., and Jeong, S.-J.: Increased atmospheric CO₂ growth rate during El Niño driven by reduced terrestrial productivity in the CMIP5 ESMs, *Journal of Climate*, 29, 8783-8805, 2016.
- 845 Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T.: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for

- greenhouse gases monitoring, *Applied optics*, 48, 6716-6733, 2009.
- 850 Lee, H., Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P., Trisos, C., Romero, J., Aldunce, P., and Barret, K.: IPCC, 2023: Climate change 2023: Synthesis report, summary for policymakers. Contribution of working groups I, II and III to the sixth assessment report of the intergovernmental panel on climate change [core writing team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, 2023.
- 855 Lee, J., Jeong, S., Kim, Y. J., Roh, S., Kim, J., and Jin, H.: Synergy of multiple-satellite measurements to fill the gap of global XCO₂, *Journal of Geophysical Research: Atmospheres*, 130, e2024JD042809, 2025.
- Li, J., Zhang, Z., Li, T., Yuan, Q., and Zhang, L.: 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, 2026.
- 860 Li, J., Jia, K., Wei, X., Xia, M., Chen, Z., Yao, Y., Zhang, X., Jiang, H., Yuan, B., and Tao, G.: High-spatiotemporal resolution mapping of spatiotemporally continuous atmospheric CO₂ concentrations over the global continent, *International Journal of Applied Earth Observation and Geoinformation*, 108, 102743, 2022.
- 865 Li, K., Bai, K., Jiao, P., Chen, H., He, H., Shao, L., Sun, Y., Zheng, Z., Li, R., and Chang, N.-B.: Developing unbiased estimation of atmospheric methane via machine learning and multiobjective programming based on TROPOMI and GOSAT data, *Remote Sensing of Environment*, 304, 114039, 2024a.
- Li, R., Zhou, X., Cheng, T., Tao, Z., Wang, N., Zhang, H., and Lv, T.: Improving Satellite XCO₂ Measurements Accuracy: A Bayesian Bias Correction Approach Considering Spatiotemporal Bias Characteristics, *IEEE Transactions on Geoscience and Remote Sensing*, 2024b.
- 870 Li, Y., Yan, J., Zhong, L., Bao, D., Sun, L., and Li, G.: Full-Coverage Mapping of Daily High-Resolution XCO₂ across China from 2015 to 2020 by Deep Learning-Based Spatio-Temporal Fusion, *IEEE Transactions on Geoscience and Remote Sensing*, 2025.
- Liang, A., Pang, R., Chen, C., and Xiang, C.: XCO₂ Fusion algorithm based on multi-source greenhouse gas satellites and carbontracker, *Atmosphere*, 14, 1335, 2023.
- 875 Liu, W., Li, R., Cao, J., Huang, C., Zhang, F., and Zhang, M.: Mapping high-resolution XCO₂ concentrations in China from 2015 to 2020 based on spatiotemporal ensemble learning model, *Ecological Informatics*, 83, 102806, 2024.
- 880 Liu, Z., Deng, Z., Zhu, B., Ciais, P., Davis, S. J., Tan, J., Andrew, R. M., Boucher, O., Arous, S. B., and Canadell, J. G.: Global patterns of daily CO₂ emissions reductions in the first year of COVID-19, *Nature Geoscience*, 15, 615-620, 2022.
- Lundberg, S. M. and Lee, S.-I.: A unified approach to interpreting model predictions, *Advances in neural information processing systems*, 30, 2017.
- 885 Ma, X., Zhang, H., Han, G., Mao, F., Xu, H., Shi, T., Hu, H., Sun, T., and Gong, W.: A regional spatiotemporal downscaling method for CO₂ columns, *IEEE Transactions on Geoscience and Remote Sensing*, 59, 8084-8093, 2021.
- Petzold, A., Thouret, V., Gerbig, C., Zahn, A., Brenninkmeijer, C. A., Gallagher, M., Hermann, M., Pontaud, M., Ziereis, H., and Boulanger, D.: Global-scale atmosphere monitoring by in-service aircraft—current achievements and future prospects of the European Research Infrastructure IAGOS, *Tellus B: Chemical and Physical Meteorology*, 67, 28452, 2015.
- 890 Qu, Y., Wei, J., Xing, H., Shi, X., Ao, Z., and Meng, X.: Global estimates of daily gapless atmospheric

- XCH₄ concentrations from satellite and reanalysis data during 2003–2020, *IEEE Transactions on Geoscience and Remote Sensing*, 2025.
- 895 Reuter, M., Buchwitz, M., Schneising, O., Krautwurst, S., O'Dell, C. W., Richter, A., Bovensmann, H., and Burrows, J. P.: Towards monitoring localized CO₂ emissions from space: co-located regional CO₂ and NO₂ enhancements observed by the OCO-2 and S5P satellites, *Atmospheric Chemistry and Physics*, 19, 9371-9383, 2019.
- Reuter, M., Bovensmann, H., Buchwitz, M., Burrows, J., Connor, B., Deutscher, N. M., Griffith, D., Heymann, J., Keppel-Aleks, G., and Messerschmidt, J.: Retrieval of atmospheric CO₂ with enhanced accuracy and precision from SCIAMACHY: Validation with FTS measurements and comparison with model results, *Journal of Geophysical Research: Atmospheres*, 116, 2011.
- 900 Rodrigues, A., Albuquerque Sardinha, R., and Pita, G.: Fundamentals of global carbon budgets and climate change, in: *Fundamental Principles of Environmental Physics*, Springer, 303-351, 2025.
- Romanov, V.: *Greenhouse Gases and Clay Minerals: Enlightening Down-to-Earth Road Map to Basic Science of Clay-Greenhouse Gas Interfaces*, Springer 2017.
- 905 Sheng, M., Lei, L., Zeng, Z.-C., Rao, W., and Zhang, S.: Detecting the responses of CO₂ column abundances to anthropogenic emissions from satellite observations of GOSAT and OCO-2, *Remote Sensing*, 13, 3524, 2021.
- Sheng, M., Lei, L., Zeng, Z.-C., Rao, W., Song, H., and Wu, C.: Global land 1° mapping dataset of XCO₂ from satellite observations of GOSAT and OCO-2 from 2009 to 2020, *Big Earth Data*, 7, 170-190, 2023.
- 910 Siabi, Z., Falahatkar, S., and Alavi, S. J.: Spatial distribution of XCO₂ using OCO-2 data in growing seasons, *Journal of environmental management*, 244, 110-118, 2019.
- Tian, W., Zhang, L., Yu, T., Wu, Y., Zhang, W., Wang, Z., and Zhu, H.: Using multi-source data and time series features to construct a global terrestrial CO₂ coverage by deep learning, *IEEE Transactions on Geoscience and Remote Sensing*, 2024.
- 915 Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I.: Attention is all you need, *Advances in neural information processing systems*, 30, 2017.
- Wang, J.: Global daily 1 km gapless XCO₂ (2003– 2023) derived from multi-satellite observations and a spatiotemporal deep learning framework, *Environmental Impact Assessment Review*, 117, 108146, 2026.
- 920 Wang, J., Liu, Z., Zeng, N., Jiang, F., Wang, H., and Ju, W.: Spaceborne detection of XCO₂ enhancement induced by Australian mega-bushfires, *Environmental Research Letters*, 15, 124069, 2020.
- 925 Wang, X., Piao, S., Ciais, P., Friedlingstein, P., Myneni, R. B., Cox, P., Heimann, M., Miller, J., Peng, S., and Wang, T.: A two-fold increase of carbon cycle sensitivity to tropical temperature variations, *Nature*, 506, 212-215, 2014.
- Wang, Y., Yuan, Q., Li, T., Yang, Y., Zhou, S., and Zhang, L.: 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 (CAMSEGG4) with a spatiotemporally self-supervised fusion method, *Earth System Science Data*, 15, 3597-3622, 2023.
- 930 Wang, Z., Zhang, C., Shi, K., Shangguan, Y., Hu, B., Chen, X., Wei, D., Chen, S., Atkinson, P. M., and Zhang, Q.: A full-coverage satellite-based global atmospheric CO₂ dataset at 0.05°

- 935 resolution from 2015 to 2021 for exploring global carbon dynamics, *Earth System Science Data*, 17(10), 5355-5375, 2025.
- Wei, J., Wang, Z., Li, Z., Li, Z., Pang, S., Xi, X., Cribb, M., and Sun, L.: Global aerosol retrieval over land from Landsat imagery integrating Transformer and Google Earth Engine, *Remote Sensing of Environment*, 315, 114404, 2024.
- 940 Wei, J., Li, Z., Chen, X., Li, C., Sun, Y., Wang, J., Lyapustin, A., Brasseur, G. P., Jiang, M., and Sun, L.: Separating daily 1 km PM_{2.5} inorganic chemical composition in China since 2000 via deep learning integrating ground, satellite, and model data, *Environmental science & technology*, 57, 18282-18295, 2023.
- Wolter, K. and Timlin, M. S.: El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI. ext), *International Journal of Climatology*, 31, 1074-1087, 2011.
- 945 Wu, C., Yang, S., Jiao, D., Chen, Y., Yang, J., and Huang, B.: Estimation of daily XCO₂ at 1 km resolution in China using a spatiotemporal ResNet model, *Science of The Total Environment*, 954, 176171, 2024.
- 950 Wunch, D., Wennberg, P. O., Osterman, G., Fisher, B., Naylor, B., Roehl, C. M., O'Dell, C., Mandrake, L., Viatte, C., and Kiel, M.: Comparisons of the orbiting carbon observatory-2 (OCO-2) XCO₂ measurements with TCCON, *Atmospheric Measurement Techniques*, 10, 2209-2238, 2017.
- Yang, D., Liu, Y., Cai, Z., Chen, X., Yao, L., and Lu, D.: First global carbon dioxide maps produced from TanSat measurements, 2018.
- 955 Yang, H., Li, T., Wu, J., Wang, Y., Zhang, L., and Hu, R.: Global 20-year XCO₂ mapping through synergy of multi-satellite observations, *Geoscience Frontiers*, 102333, 2026.
- Yeom, J.-M., Jeong, S., Ha, J.-S., Lee, K.-H., Lee, C.-S., and Park, S.: Estimation of the hourly aerosol optical depth from GOCI geostationary satellite data: deep neural network, machine learning, and physical models, *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-12, 2021.
- 960 Yu, Y., Tian, W., Zhang, L., Yu, T., Wu, Y., and Cheng, T.: MCF-XCO₂: A cross-mission consistency and fusion framework for integrating multi-satellite XCO₂ observations, *Atmospheric Research*, 108747, 2026.
- Zhang, B., Zhang, H., Zhao, G., and Lian, J.: Constructing a PM_{2.5} concentration prediction model by combining auto-encoder with Bi-LSTM neural networks, *Environmental Modelling & Software*, 124, 104600, 2020.
- 965 Zhang, L., Li, T., Wu, J., and Yang, H.: Global estimates of gap-free and fine-scale CO₂ concentrations during 2014–2020 from satellite and reanalysis data, *Environment International*, 178, 108057, 2023.
- Zhang, M. and Liu, G.: Mapping contiguous XCO₂ by machine learning and analyzing the spatio-temporal variation in China from 2003 to 2019, *Science of The Total Environment*, 858, 159588, 2023.
- 970