



Reconstructing two-decade daily high-resolution seamless global land XCO_2 records using a hybrid Transformer–BiLSTM model

Yu Qu^{1,2}, Xian Shi³, Yulong Fan^{1,3}, Zhihui Wang^{4,5}, and Jing Wei¹

¹MEEKL-AERM, College of Environmental Sciences and Engineering, Institute of Tibetan Plateau, and Center for Environment and Health, Peking University, Beijing 100871, China

²School of Geographical Sciences, South China Normal University, Guangzhou 510631, China

³College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China

⁴University of Science and Technology of China, Hefei 230026, China

⁵Anhui Province Key Laboratory of Optical Quantitative Remote Sensing, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China

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

Received: 20 January 2026 – Discussion started: 13 February 2026

Revised: 14 May 2026 – Accepted: 11 June 2026 – Published: 24 June 2026

Abstract. Accurate and temporally continuous global observations of atmospheric carbon dioxide (XCO_2) are essential for climate monitoring and emission assessment. However, satellite-based XCO_2 observations are often spatially incomplete and temporally discontinuous, while existing products typically suffer from coarse spatial resolution, hindering the detection of fine-scale emission changes. Here, we developed a novel spatiotemporal Transformer–BiLSTM deep-learning network that combines the Transformer’s ability to model long-range spatial dependencies through self-attention mechanisms with the BiLSTM’s ability to capture temporal dynamics. The network assimilates multisource data from satellite observations, meteorological reanalysis, and precursor gases to reconstruct a global, daily, and seamless XCO_2 dataset over land at 0.1° spatial resolution from 2003 to 2022. Independent validation of the data-fused XCO_2 product against Total Carbon Column Observing Network (TCCON) measurements shows excellent agreement, with an R^2 of 0.99, an RMSE of 1.10 ppm, and a mean bias of 0.01 ppm. After bias correction, cross-satellite consistency is further enhanced, achieving a sample-based CV- R^2 of 0.99 and an RMSE of 0.36 ppm. The dataset provides accurate daily XCO_2 estimates over global land surfaces, enabling investigations of spatial heterogeneity and regional-to-local XCO_2 enhancement patterns linked to anthropogenic emissions and biomass-burning events. The record reveals a persistent global increase in atmospheric XCO_2 over the past two decades, with a mean growth rate of 2.24 ppm yr^{-1} ($p < 0.001$). It reliably resolves global XCO_2 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 XCO_2 enhancements associated with major wildfire events, demonstrating its capability to represent both persistent and episodic emission signals. This high-resolution, daily global XCO_2 (GlobalHigh XCO_2) product provides a valuable resource for carbon-cycle research, atmospheric model evaluation, and emission monitoring, and is publicly available at <https://doi.org/10.5281/zenodo.18220962> (Qu and Wei, 2026).

1 Introduction

Carbon dioxide (CO_2) 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 CO_2 concentration reached 419.31 ± 0.15 ppm (parts per million) in 2023, approximately 50 % higher than preindustrial levels, making it a major contributor to ongoing climate change (Friedlingstein et al., 2023). In response to rising CO_2 levels, the international community adopted the Paris Agreement in 2015, aiming to limit global warming to well below $2^\circ C$ compared to preindustrial levels. Owing to its long atmospheric lifetime, CO_2 accumulates 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 CO_2 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 CO_2 (XCO_2) are commonly quantified using ground-based networks and satellite remote sensing (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 XCO_2 monitoring, but it generally has a lower temporal sampling frequency (Buchwitz et al., 2015). Since the early 2000s, a series of satellite missions has progressively advanced XCO_2 observations in accuracy, spatial resolution, and temporal coverage. The SCanning Imaging Absorption spectromETER for Atmospheric CHartographY (SCIAMACHY) aboard the Environmental Satellite (Envisat), launched in 2002, was the first satellite instrument to provide global XCO_2 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 understanding large-scale carbon cycle processes. Despite major advances in satellite XCO_2 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 address spatial gaps in satellite XCO_2 observations, researchers have explored a range of data fusion and reconstruction methods. 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 XCO_2 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 a point-wise regression problem, 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 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 these models to fuse multisource satellite data (Huang et al., 2024a), incorporate temporal sequence features for daily XCO_2 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 XCO_2 dataset 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 (Li et al., 2024b), bias correction is not yet fully integrated into the reconstruction framework. Moreover, 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 and captures long-range spatiotemporal dependencies at the global daily scale remains lacking.

To address these limitations, we develop a hybrid spatiotemporal deep-learning framework based on a Transformer–BiLSTM architecture for global daily XCO_2 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 in atmospheric XCO_2 fields. Leveraging this framework, we generate a seamless global daily XCO_2 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 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 d, retrieved using the Bremen Optimal Estimation DOAS (BESD) algorithm (Bovensmann et al., 1999; Buchwitz et al., 2005; Reuter et al., 2011). GOSAT, launched in 2009, provides near-global coverage at 10.5 km spatial resolution every 3 d 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 $\sim 16 \text{ d}$ using three grating spectrometers (Crisp, 2015). Here, we used the SCIAMACHY BESD products from 2003 to 2009, GOSAT Level 2 XCO_2 Version 9r products from 2010 to 2014, and 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 resolution and a 3 h 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 XCO_2 measurements with an average uncertainty of approximately 0.2 %. The network offers sub-daily temporal sampling, typically including multiple observations per clear-sky day, and serves as a standard reference for the validation of satellite XCO_2 retrievals (Wunch et al., 2017). We used all available 31 stations in the GGG2020 release, distributed across global land areas (Fig. 1), serving as the reference for assessing reconstruction performance and guiding bias correction. To independently evaluate the bias-corrected daily XCO_2 dataset, 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 1 or 2 were retained to ensure consistency with the World Meteorological Organization X2019 calibration scale.

2.1.2 Meteorological and auxiliary variables

Meteorological variables relevant to atmospheric CO_2 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 (TEM), surface net solar radiation (SNSR), total evaporation (ET), and 10 m u - and v -components of wind (WU and WV). Boundary layer height (BLH) and relative humidity (RH) were obtained 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 (NO_2) and carbon monoxide (CO) total columns from the CAMS analysis, as well as column-averaged methane (XCH_4) derived from our previous study (Qu et al., 2025), which were 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 Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) at 90 m resolution. Table 1 provides an overview of all data sources used in this study.

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 spatial and temporal dependencies. The input predictors are first processed by a Transformer encoder (Vaswani

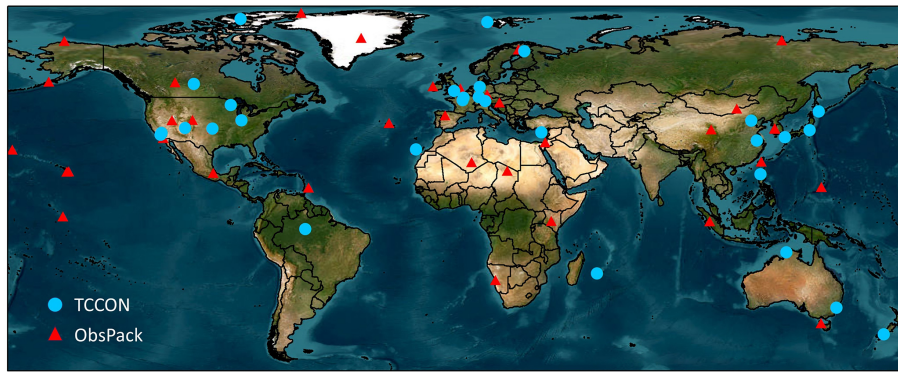


Figure 1. Geographic locations of TCCON (Total Carbon Column Observing Network; blue dots) and ObsPack (Observations Package; red triangles) observation stations.

Table 1. Summary of the datasets used for XCO₂ reconstruction in this study.

Type	Source	Product	Spatial resolution	Temporal resolution	Period
Satellite	SCIAMACHY	XCO ₂	30 × 60 km ²	~ 30 d	2003–2009
	GOSAT	XCO ₂	10.5 km ²	~ 3 d	2010–2014
	OCO-2	XCO ₂	1.29 × 2.25 km ²	~ 16 d	2015–2022
Meteorological variables	ERA5	Surface pressure (SP) 2 m temperature (TEM) 10 m <i>u</i> -component of wind (WU) 10 m <i>v</i> -component of wind (WV) Surface net solar radiation (SNSR) Total evaporation (ET)	0.1°	Hourly	2003–2022
		Relative humidity (RH) Boundary layer height (BLH)			
Precursor	CAMS	XCO ₂	0.75°	3 h	2003–2022
		Nitrogen dioxide (NO ₂) Carbon monoxide (CO)	0.75° 0.75°	3 h 3 h	2003–2022 2003–2022
	Qu et al. (2025)	GlobalHighXCH ₄	0.1°	Daily	2003–2022
Surface-related variables	MODIS	NDVI LST	0.05°	Monthly	2003–2022
	SRTM	Elevation	90 m	–	–

et al., 2017), which models long-range spatial dependencies at regional and global scales through multi-head self-attention and position-wise feed-forward layers, without relying on local receptive fields. Each location's feature vector is projected into query (Q), key (K), and value (V) vectors, and the resulting attention scores quantify the influence of all other locations on that location. The multi-head attention mechanism integrates complementary spatial relationships across multiple feature subspaces, enabling the model to explicitly capture planetary-scale influences, such as atmospheric teleconnections and large-scale weather systems, on regional XCO₂ concentrations.

The spatially encoded features generated by the Transformer are subsequently passed to a Bidirectional Long Short-Term Memory (BiLSTM) network, which learns forward and backward temporal dependencies by integrating both past and future information (Zhang et al., 2020). This sequence-to-sequence formulation effectively captures temporal variability arising from diurnal cycles, synoptic weather variations, and seasonal biospheric processes affecting XCO₂ concentrations, thereby generating comprehensive spatiotemporal feature representations for each grid cell.

The resulting spatiotemporal features are finally passed through a multilayer perceptron (MLP) block, followed by a fully connected linear layer that projects the fused features

to scalar daily XCO₂ estimates for each grid cell. To effectively capture the complex spatiotemporal dynamics of atmospheric XCO₂, we designed a weighted spatiotemporal loss function that accounts for intrinsic temporal continuity and spatial correlation, in contrast to the standard mean squared error (MSE) loss function:

$$L_{XCO_2} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 + \chi_1 \frac{1}{N-1} \sum_{i=2}^N (\Delta y_i - \Delta \hat{y}_i)^2 + \chi_2 L_{\text{spatial}} \quad (1)$$

where y_i and \hat{y}_i denote the observed and predicted XCO₂ values for the i th sample, respectively, and Δy_i and $\Delta \hat{y}_i$ represent the corresponding temporal differences between consecutive time steps. The first term enforces point-wise reconstruction accuracy, the second term promotes temporal consistency, and L_{spatial} introduces a Laplacian-based spatial smoothness constraint to preserve coherence among neighboring 0.1° grid cells. The hyperparameters $\chi_1 = 0.5$ and $\chi_2 = 0.01$, which regulate the temporal and spatial regularization terms, respectively, were determined via a grid search.

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 single-layer bidirectional LSTM (BiLSTM) with 64 hidden units per direction, yielding a 128-dimensional output feature. A dropout rate of 0.5 was applied to mitigate overfitting. The Transformer module comprised 4 encoder layers, each with 64-dimensional hidden layers and 4 attention heads, with ReLU activation functions applied between layers to enhance learning and facilitate sparse feature extraction. The Transformer–BiLSTM model was trained with the Adam optimizer, an initial learning rate of 0.001, and a batch size of 256. To optimize convergence, a learning-rate scheduler reduced the learning 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 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 workflow that decouples spatial pattern learning from magnitude calibration (Fig. 2).

Phase 1: data fusion

The objective of this stage is to reconstruct a spatially continuous global XCO₂ field from sparse and discontinuous satellite observations. The Transformer–BiLSTM model uses XCO₂ retrievals from SCIAMACHY, GOSAT, and OCO-2 as training targets and integrates multisource predictors,

including atmospheric variables (XCH₄ from our previous study, Qu et al., 2025, as well as CAMS XCO₂, NO₂, and CO columns), ERA5 meteorology variables (BLH, TEM, SNSR, RH, SP, ET, WU, and WV), and land-surface variables (NDVI, LST, and DEM). In addition, spatiotemporal encoding vectors are incorporated to characterize geographic and temporal dependencies. Specifically, spatial coordinates are represented using three Euclidean spherical functions ($P_S = S1, S2, S3$), while temporal information is encoded using three helix-shaped trigonometric functions ($P_T = T1, T2, T3$), enabling the model to capture spatial heterogeneity and temporal variability (Wei et al., 2023; Qu et al., 2025). The output of this stage is a preliminary global XCO₂ dataset at 0.1° × 0.1° spatial resolution, providing a seamless and spatially continuous field for the subsequent bias-correction stage.

$$XCO_2^{\text{Satellites}} \sim f_{\text{Transformer-BiLSTM}} \left(XCO_2^{\text{CAMS}}, \right. \\ \left. NO_2^{\text{CAMS}}, CO^{\text{CAMS}}, XCH_4, \right. \\ \left. \text{Meteo}, NDVI, LST, DEM, P_S, P_T \right) \quad (2)$$

Phase 2: bias correction

Satellite XCO₂ retrievals often exhibit systematic biases arising from differences in sensor characteristics, retrieval algorithms, and temporal sampling. In this phase, the fused XCO₂ field from Phase 1 (XCO_2^{Fusion}) is further calibrated through a bias-correction procedure using ground-based TCCON measurements as reference observations. The Transformer–BiLSTM model ingests the fused XCO₂ product together with the same auxiliary predictors used in the fusion stage, including atmospheric composition variables, meteorological variables, land-surface characteristics, and spatiotemporal encoding vectors. By learning the complex non-linear relationships between satellite-dependent biases and environmental conditions, the model effectively reduces systematic errors while maintaining robust global generalization capability. The final output is a bias-corrected, seamless daily global XCO₂ dataset at 0.1° × 0.1° spatial resolution with improved consistency across all satellite platforms:

$$XCO_2^{\text{TCCON}} \sim f_{\text{Transformer-BiLSTM}} \left(XCO_2^{\text{Fusion}}, \right. \\ \left. NO_2^{\text{CAMS}}, CO^{\text{CAMS}}, XCH_4, \right. \\ \left. \text{Meteo}, NDVI, LST, DEM, P_S, P_T \right). \quad (3)$$

2.2.3 Innovations of the framework

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

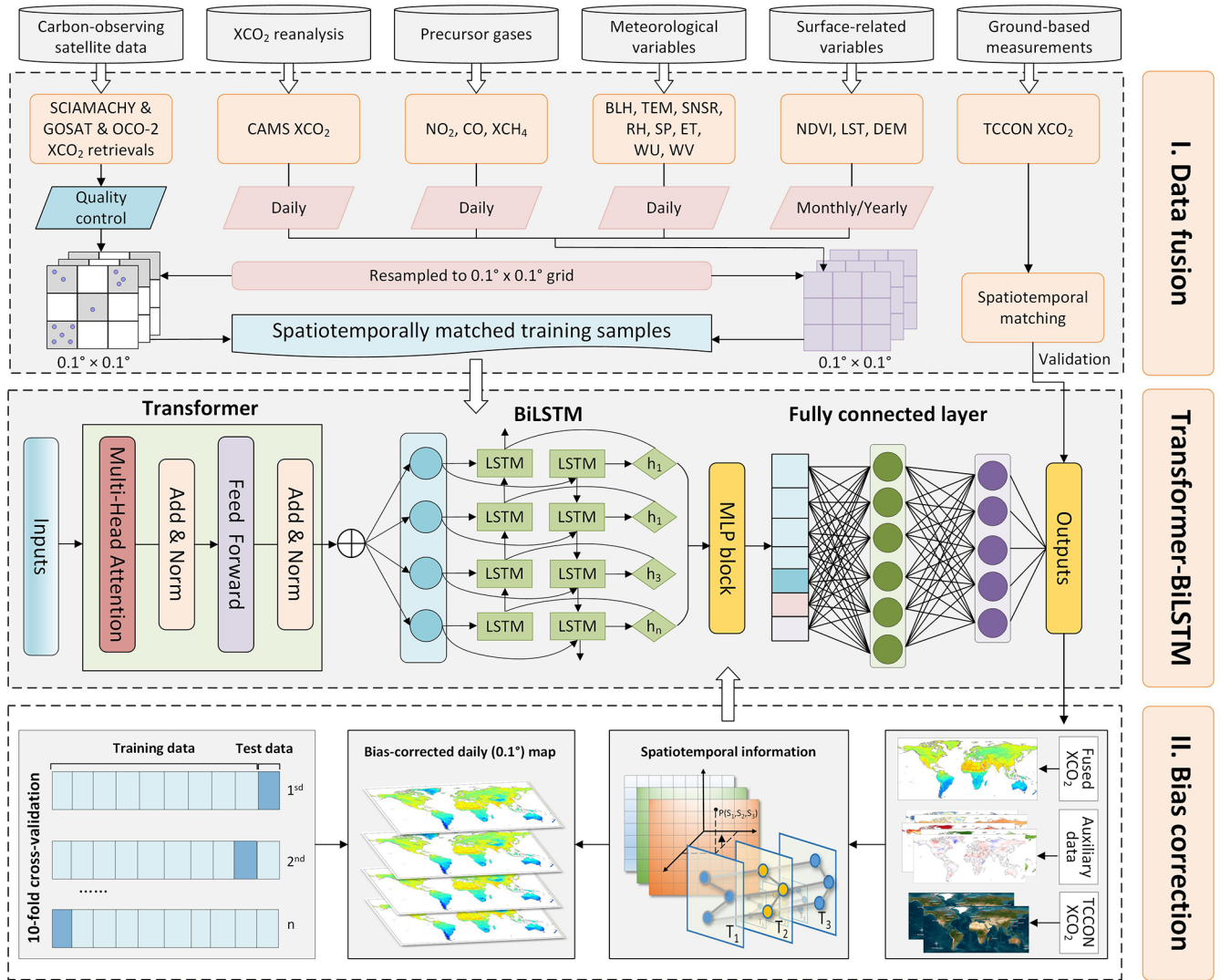


Figure 2. Workflow for generating a global, daily, seamless XCO₂ dataset at a spatial resolution of 0.1° × 0.1° using the developed Transformer–BiLSTM framework.

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: BiLSTM, 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.

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 in-

Table 2. Ablation analysis of different backbone models for XCO₂ reconstruction.

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

consistencies 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 daily XCO₂ reconstruction with improved temporal continuity and cross-mission consistency.

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. The temporal-based CV withheld continuous blocks of daily data to evaluate the model's temporal generalization during periods without ground-based measurements. The spatial-based CV withheld spatially contiguous clusters of grid cells (1° × 1°) to assess the model's capability to generalize to 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 R^2 (CV- R^2) 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 associated with higher observation density and retrieval quality in later missions. Temporal-based CV further demonstrates stable predictive performance under interannual extrapolation, with

CV- R^2 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- R^2 increasing from 0.86 (SCIAMACHY) to 0.89 (GOSAT) and 0.96 (OCO-2), accompanied by corresponding reductions in RMSE (MAE) from 1.71 (1.28) to 1.30 (0.98) and 1.10 (0.79) ppm.

Furthermore, we independently evaluated our data-fused XCO₂ estimates against TCCON ground-based measurements and compared their performance with satellite retrievals and the CAMS reanalysis. Satellite XCO₂ observations show strong agreement with TCCON, yielding an R^2 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 ($R^2 = 0.97$) 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 product achieves the best agreement with 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.

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 = 0.36 ppm, MAE = 0.23 ppm) and a near-zero mean bias of 0.01 ppm, indicating that systematic biases in the data-fused product 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 a CV- 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 a CV- 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 introduce artificial discontinuities. To evaluate the temporal stability of the reconstructed dataset across major satellite mission transitions (~2010 and ~2015), we conducted targeted diagnostics at long-term TCCON sites, including Garmisch (Europe) and Park Falls (North America) (Fig. 5). Daily XCO₂ 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, the un-

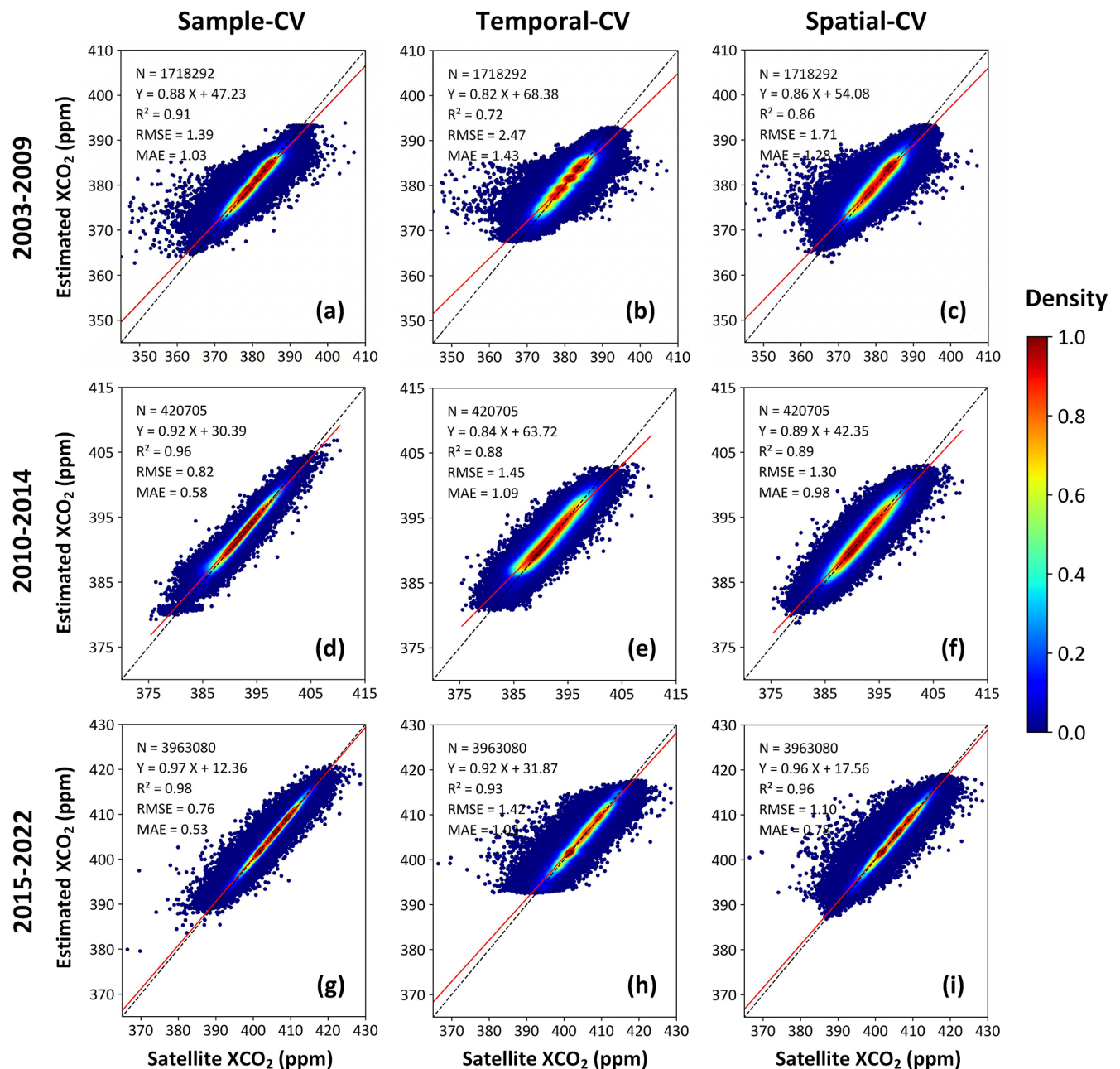


Figure 3. Density scatterplots of sample-based (a, d, g), temporal-based (b, e, h), and spatial-based (c, f, i) ten-fold cross-validation (10-CV) results for daily XCO₂ estimates during 2003–2009 (SCIAMACHY, a–c), 2010–2014 (GOSAT, d–f), and 2015–2022 (OCO-2, g–i). Black dashed lines represent the 1 : 1 relationship, and red solid lines indicate linear regression fits.

corrected fused XCO₂ occasionally exhibits small but discernible step changes or systematic offsets coincident with mission transitions, indicating residual inter-mission inconsistencies. In contrast, the bias-corrected XCO₂ 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 variability while effectively eliminat-

ing sensor-induced biases. Together, these results highlight the robustness of the bias-correction strategy in ensuring the temporal seamlessness of the multi-decadal XCO₂ record.

3.1.3 Independent validation against ObsPack observations

To further evaluate the generalization ability of the bias-corrected daily XCO₂ dataset, we conducted an indepen-

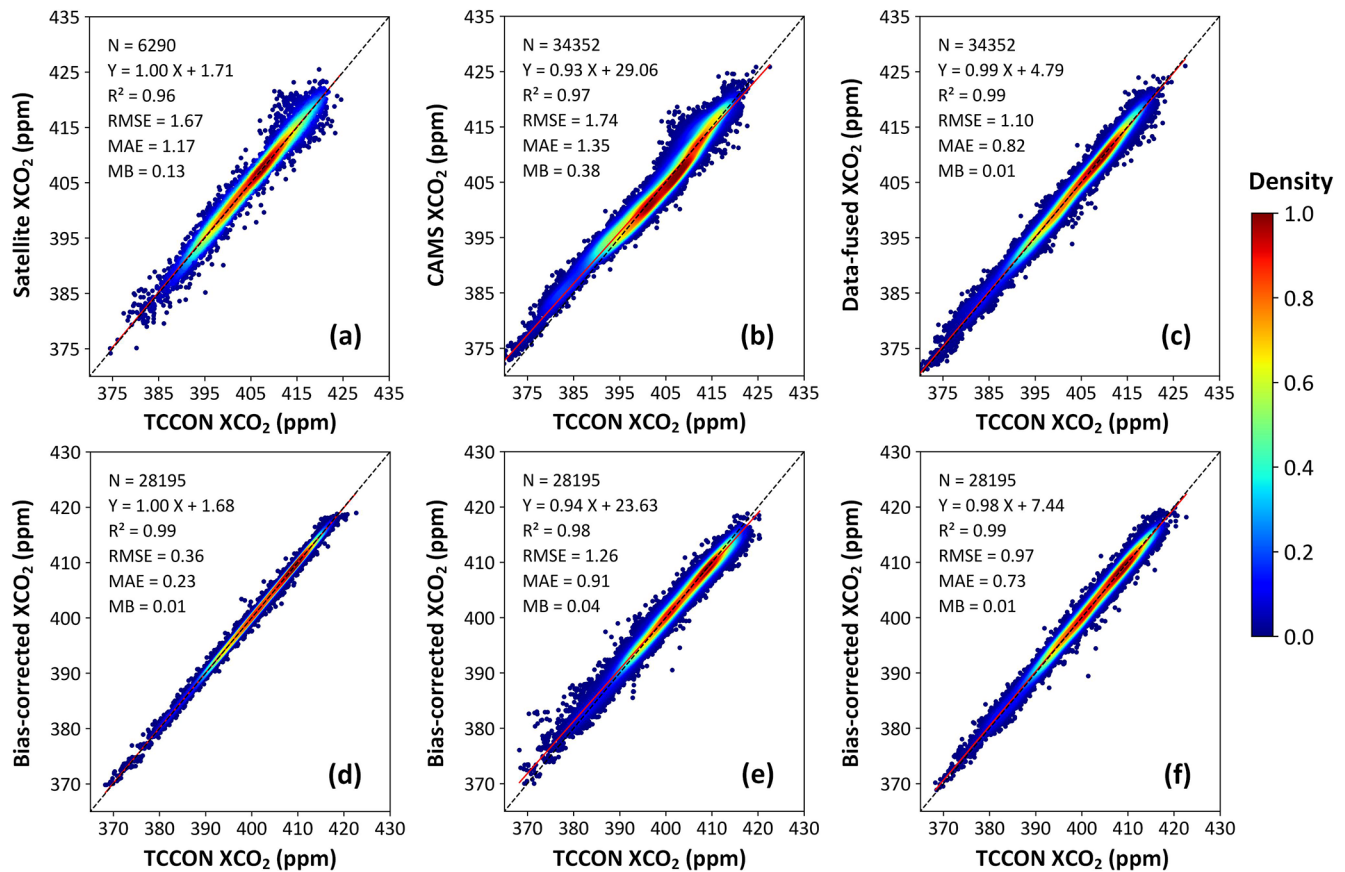


Figure 4. Density scatterplots comparing daily XCO₂ concentrations from (a) satellite retrievals, (b) CAMS reanalysis, and (c) our data-fused approach against TCCON ground-based measurements, as well as ten-fold cross-validation (10-CV) results for our bias-corrected daily XCO₂ concentrations using (d)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

dent 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 R^2 of 0.86. More than 90 % of the stations show high R^2 values exceeding 0.6, and over half (54 %) exceed 0.9, indicating robust performance across diverse geographic and climatic conditions. The low R^2 at the AZR station is mainly due to the limited number of matched observations and the relatively small temporal variability at this site, which can lead to unstable correlation statistics. Importantly, the RMSE and bias at this station remain relatively small, suggesting that the low R^2 does not indicate a systematic deficiency in the reconstructed dataset. 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 having bias values within ± 3 ppm and MAE values below 5.0 ppm, respectively.

3.2 Spatiotemporal variations of global atmospheric XCO₂

3.2.1 Global daily seamless XCO₂ maps

We compare global daily XCO₂ data derived from different products for three representative dates: 30 May 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 ($0.75^\circ \times 0.75^\circ$) 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, $3^\circ \times 2^\circ$) 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 inter-hemispheric gradients but exhibits regional inconsistencies, including exaggerated concentrations in parts of the North-

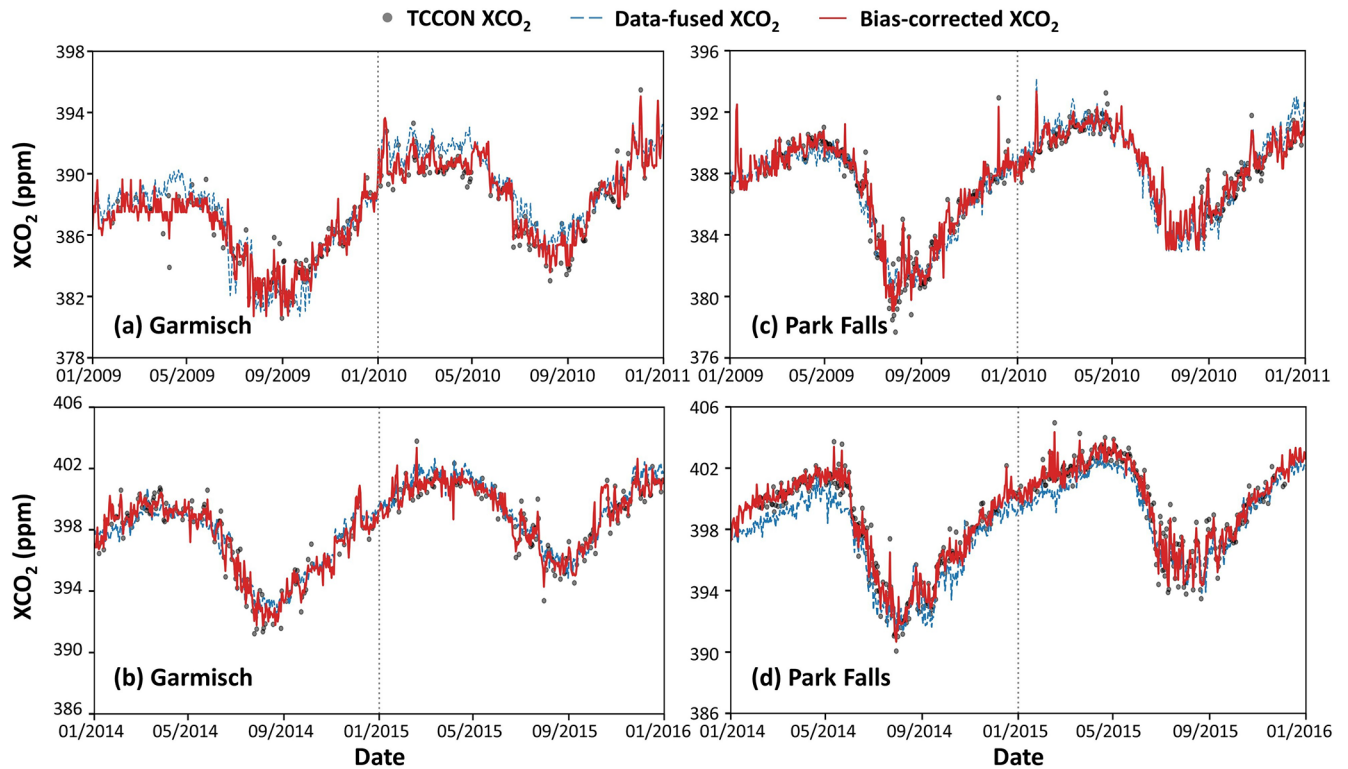


Figure 5. Time series of TCCON (grey dots), data-fused (blue dashed lines), and bias-corrected (red solid lines) XCO_2 concentrations at (a, b) Garmisch (Europe) and (c, d) Park Falls (North America) across major satellite mission transitions: (a, c) SCIAMACHY–GOSAT (2009–2010) and (b, d) GOSAT–OCO-2 (2014–2016). The vertical dotted lines indicate the approximate timing of satellite mission transitions.

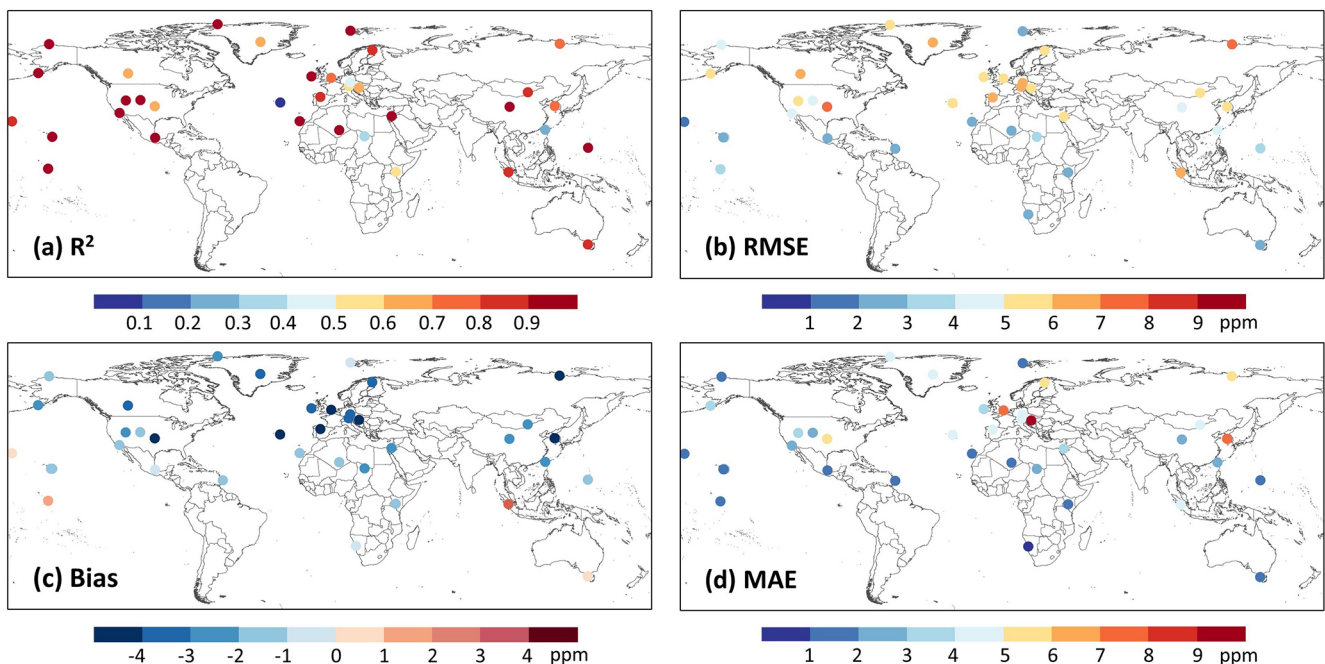


Figure 6. Global validation of our bias-corrected daily XCO_2 dataset against ObsPack surface flask observations during 2003–2022.

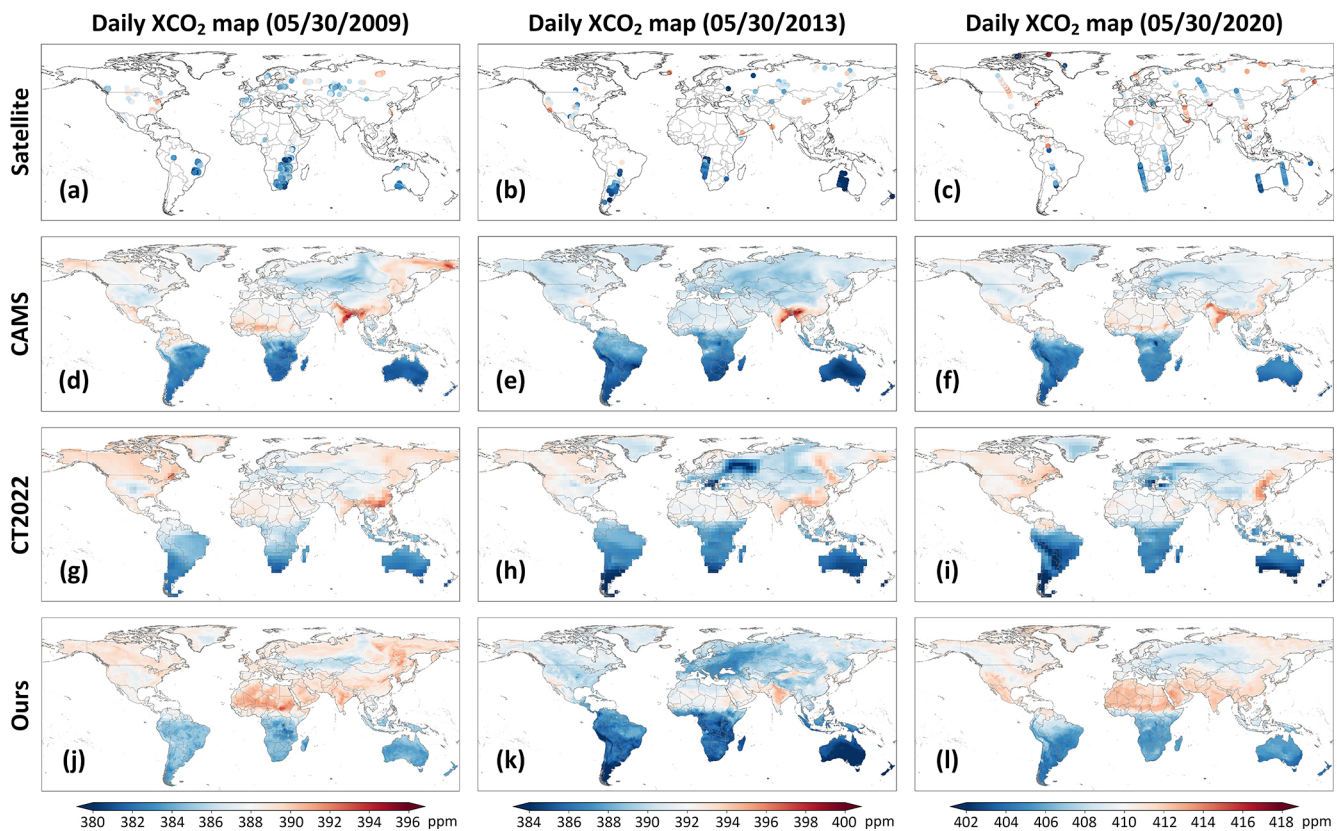


Figure 7. Spatial distributions of global daily XCO₂ concentrations on 30 May 2009, 2013, and 2020 from (a–c) SCIAMACHY, GOSAT, and OCO-2 retrievals, (d–f) CAMS reanalysis, (g–i) CarbonTracker (CT2022), and (j–l) our reconstructed XCO₂ product.

ern 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 overestimated XCO₂ concentrations over the eastern United States, while CAMS underestimated regional XCO₂ enhancements; our estimates closely matched satellite observations. In 2013, both CAMS and CT2022 underestimated XCO₂ concentrations over northern China and southern Mongolia, whereas our product accurately captured the observed spatial patterns. In 2020, both CAMS and CT2022 underestimated XCO₂ concentrations over the Middle East, whereas our reconstruction captured the enhanced concentrations, in agreement with satellite observations. These examples demonstrate that our data provide a balanced and physically consistent representation of daily global XCO₂, avoiding CAMS over-smoothing and the regional biases evident in CT2022.

3.2.2 Characterization of localized XCO₂ enhancement signals

Daily XCO₂ observations provide a distinct advantage for revealing localized and transient regions with elevated XCO₂ concentrations that are often obscured in temporally averaged products (Fig. 8). Pronounced XCO₂ enhancements are observed over regions with strong anthropogenic influence, including major industrial and energy-related areas such as steel production (the Yokohama Steel Plant in Japan), coal mining (the Arizona Coal Mine in the US), power generation (a power plant in Algeria), and chemical processing (a chemical plant in Russia). Elevated XCO₂ concentrations are also observed over industrial clusters in East Asia, Central Asia (the Muruntau Gold Mine in Uzbekistan), Europe (an oil refinery in Poland), North America, and the Middle East (the chemical industry in Saudi Arabia). 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, localized enhancements are also evident near major transport hubs, such as John F. Kennedy International Airport in the US. Beyond anthropogenic sources, the daily XCO₂ dataset also captures episodic enhancement events associated with natural processes, such as forest fires (the Xi-

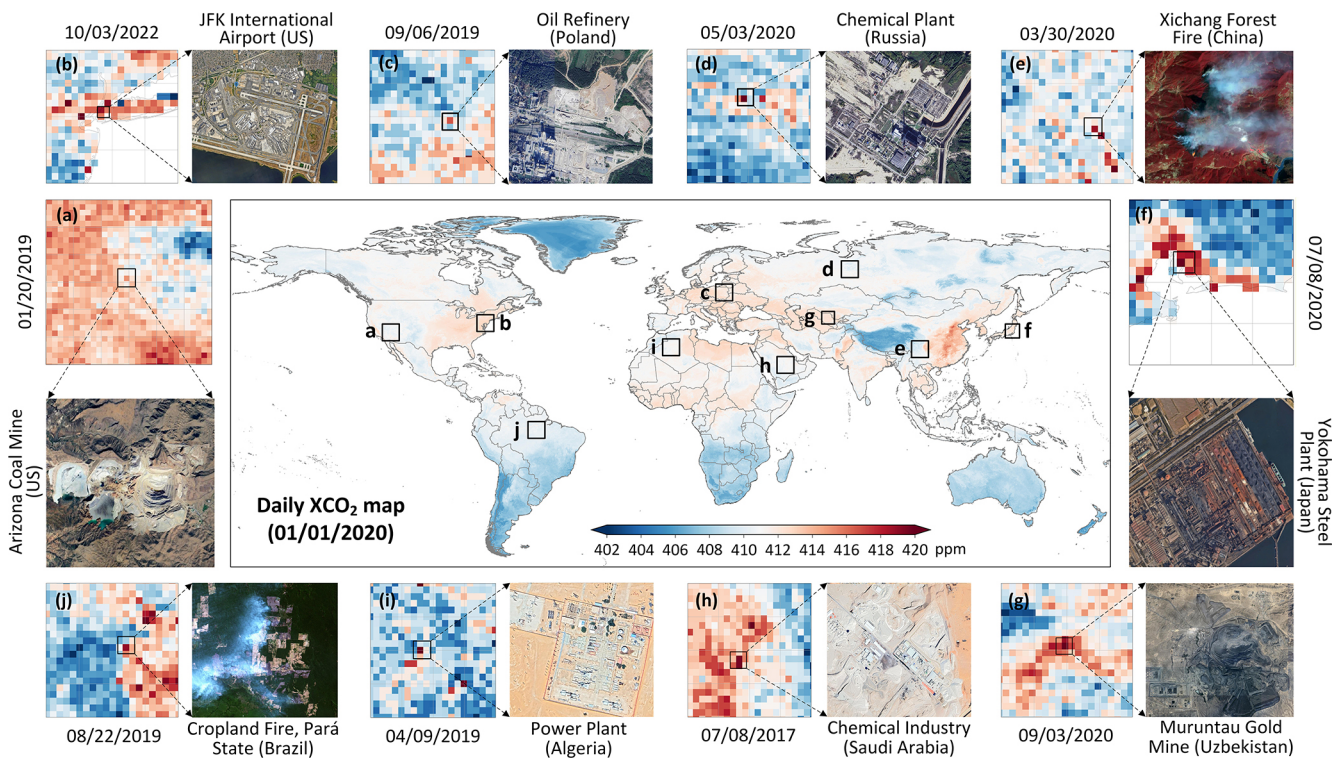


Figure 8. Characterization of localized XCO₂ enhancement identified signals from the reconstructed global daily seamless dataset across (a–j) 10 regions of interest on different days. The central map shows the global distribution of XCO₂ on 1 January 2020.

chang Forest Fire in China) and large-scale agricultural burning (cropland fires in Pará State, Brazil), which can generate short-lived yet spatially coherent XCO₂ anomalies at daily timescales. These examples 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 uncertainty or noise, especially in regions with limited observations.

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 27 December 2019, to 8 January 2020, coinciding with the peak of the 2019–2020 Australian bushfires. The daily XCO₂ time series exhibits a distinct local maximum on 5 January 2020 (average = 411.35 ± 0.56 ppm), superimposed on a gradually increasing back-

ground 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 for over two weeks, from 24 March to 8 April 2020, with elevated daily XCO₂ concentrations. A clear local peak of 410.56 ± 0.53 ppm is observed on 30 March, coinciding with the reported Xichang forest fire outbreak, followed by a gradual decline after 5 April. 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.

Figure 10 shows that XCO₂ growth rates from our reconstructed dataset consistently respond to ENSO phases, as evidenced by a strong correlation with the multivariate ENSO index (MEI; $R = 0.55$, $p < 0.01$) (Wolter and Timlin, 2011), which explains $\sim 30\%$ 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, drier climates suppress ecosystem CO₂ uptake 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

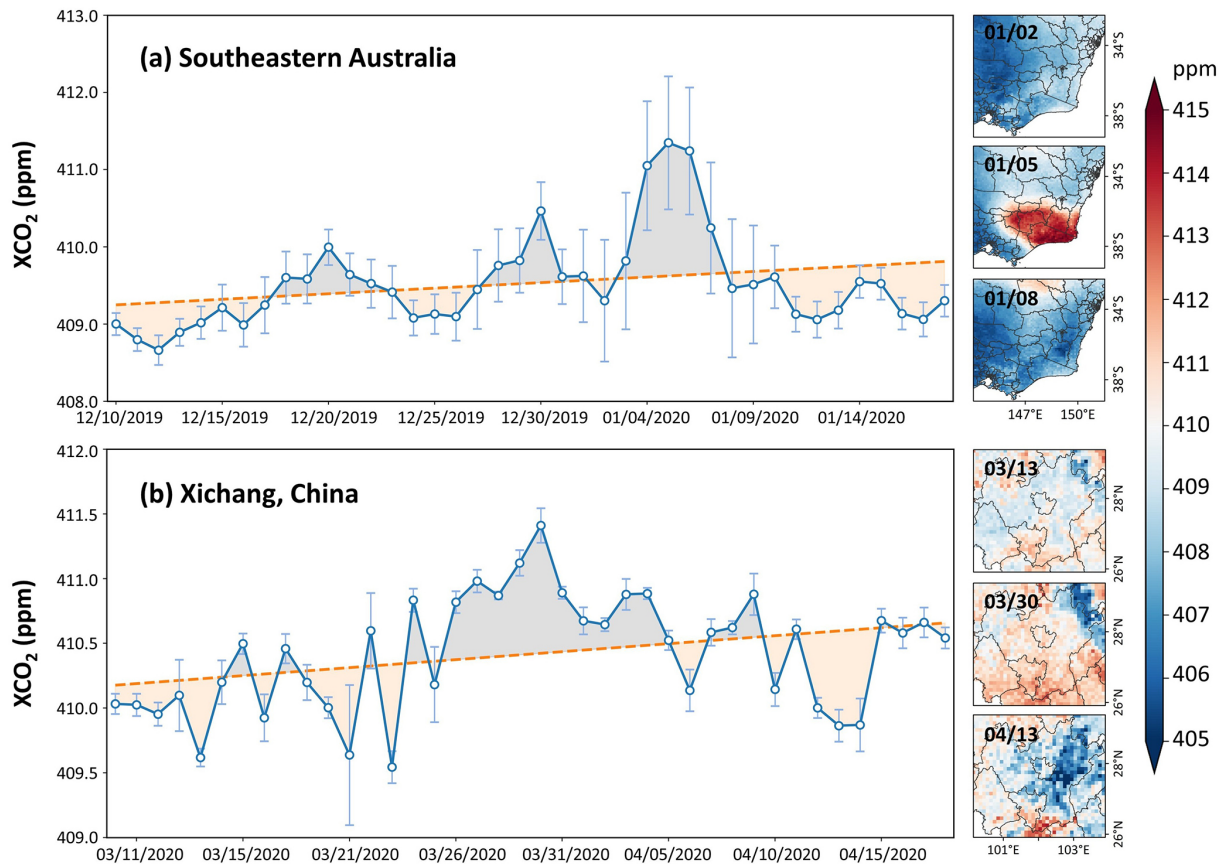


Figure 9. Time series and spatial distributions of daily XCO₂ concentrations associated with extreme wildfire events in (a) southeastern Australia during December 2019–January 2020 and (b) Xichang, China, during March–April 2020.

coincided with the highest recorded global XCO₂ growth rate of 3.11 ppm yr⁻¹ ($p < 0.001$) (Liu et al., 2024). In contrast, La Niña phases are typically associated with cooler and wetter conditions that strengthen terrestrial carbon sinks and reduce atmospheric CO₂ 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 $\sim 45\%$ 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 ($\sim 37\%$), while the Southern Extratropics are less affected ($\sim 27\%$), reflecting hemispheric differences in land–atmosphere coupling and fire regimes.

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 were relatively low, with most regions characterized by concentrations below

~ 380 ppm. From 2006 to 2010, XCO₂ increased steadily across all continents, while spatial contrasts remained relatively modest. After approximately 2010, the spatial patterns intensified markedly. From 2011 to 2015, high-XCO₂ regions expanded and strengthened, and from 2016 onward, the global increase accelerates further, with XCO₂ 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 displayed XCO₂ values above 415 ppm, with the highest concentrations concentrated over major industrialized and densely populated regions in the Northern Hemisphere. Overall, the global mean XCO₂ increased monotonically from 374.54 ppm in 2003 to 416.36 ppm in 2022, corresponding to approximately an 11 % increase over the past 20 years. This substantial growth highlights the rapid accumulation of atmospheric CO₂ and the growing influence of anthropogenic emissions on the global carbon cycle.

Figure 12 presents the long-term (2003–2022) annual mean XCO₂ and seasonal climatology. The long-term global mean XCO₂ 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

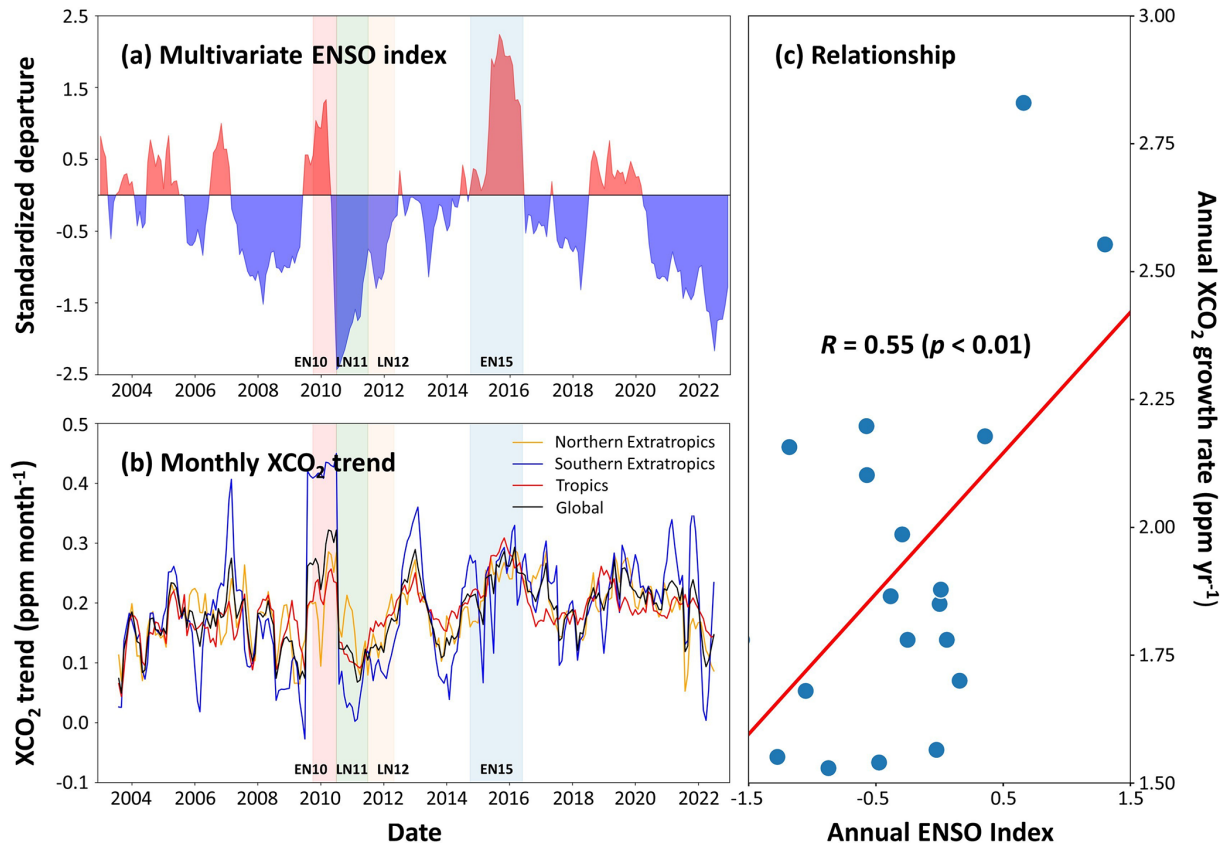


Figure 10. Time series of the monthly (a) multivariate ENSO index and (b) detrended XCO_2 growth rates over major regions, including the Northern Extratropics (30–90° N), Southern Extratropics (90–30° S), Tropics (30° S–30° N), and the globe. (c) Scatter plots of annual XCO_2 growth rates versus the annual ENSO index. The 2011 La Niña event (LN11) is shaded in dark green. The preceding 2010 El Niño event (EN10) and subsequent weak 2012 La Niña event (LN12) are shaded in lighter red and green colors, respectively. The 2015 El Niño event (EN15) is shaded in dark blue.

Hemisphere. The highest XCO_2 concentrations are concentrated in the northern low- to mid-latitudes (~ 10 – 45° N), 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 XCO_2 values occur over South America, central and southern Africa, and Australia, where strong biospheric uptake, sparse anthropogenic emissions, and the influence of clean marine air masses suppress atmospheric CO_2 levels. At the national scale, South Korea exhibits the highest XCO_2 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 XCO_2 is evident across global land (Fig. 12b–d). During boreal spring (MAM; average = 395.36 ± 2.20 ppm), elevated XCO_2 concentrations are widespread across the Northern Hemisphere, reflecting accumulated winter emissions and the delayed onset of biospheric uptake. In boreal summer (JJA; average = $392.74 \pm$

1.29 ppm), XCO_2 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. 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; mean = 394.76 ± 1.37 ppm), enhanced fossil fuel combustion and reduced biospheric uptake lead to elevated XCO_2 concentrations across the Northern Hemisphere, whereas lower values prevail over the Southern Hemisphere land and adjacent oceans, resulting in a pronounced interhemispheric gradient.

Figure 13 presents long-term XCO_2 trends from 2003 to 2022 across different temporal scales. All land regions exhibit significant increasing trends, with an average

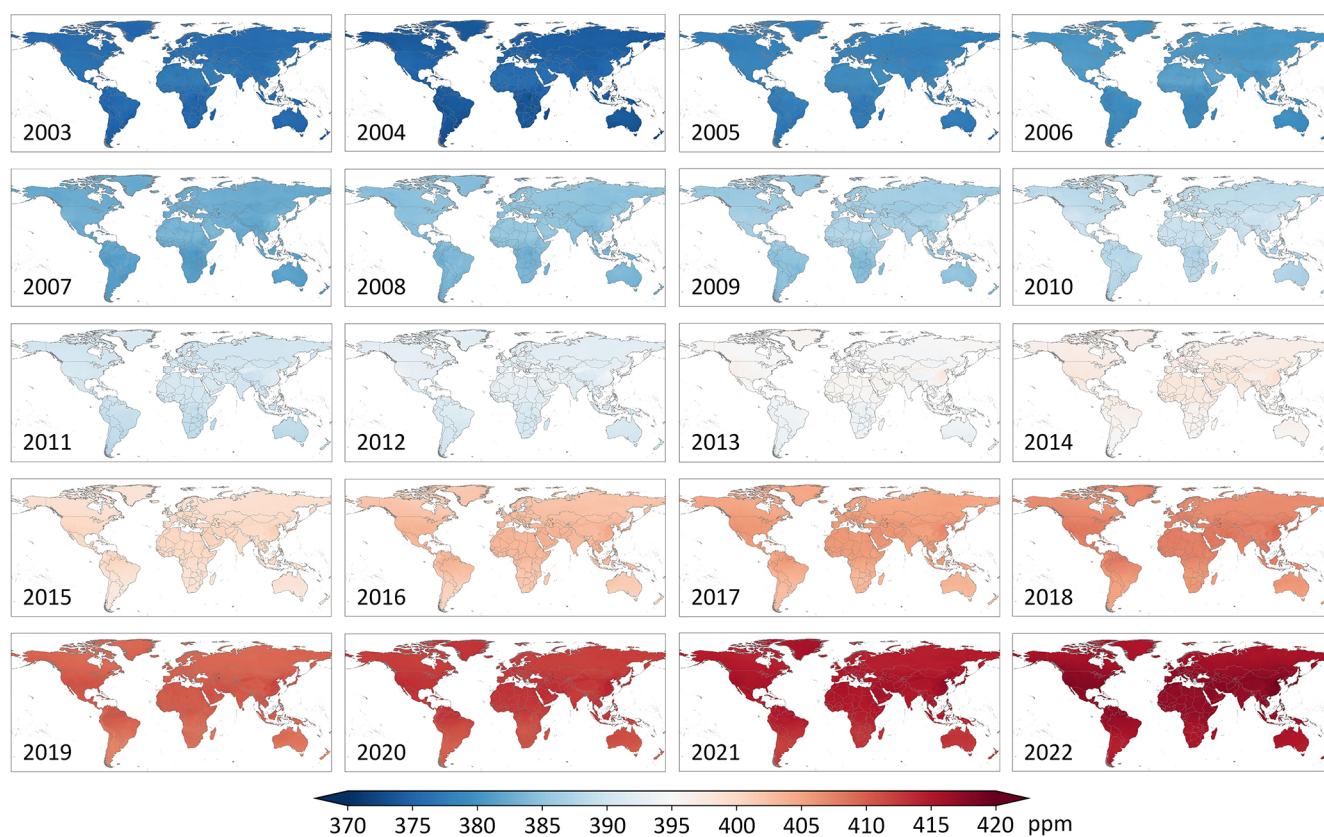


Figure 11. Spatial distribution of annual mean XCO₂ concentrations (unit: ppm) over land from 2003 to 2022.

of 2.24 ppm yr⁻¹ ($p < 0.001$). The strongest increases (> 2.25 ppm yr⁻¹) occur over East and Southeast Asia, northern South America, and central Africa, while comparatively weaker growth (< 2.15 ppm yr⁻¹) 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⁻¹, $p < 0.001$), 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⁻¹, $p < 0.001$), trends weaken over boreal and temperate regions but remain pronounced in the tropics, particularly in the Amazon basin, central Africa, and Southeast Asia, reflecting ongoing anthropogenic emissions and biomass burning. SON (average = 2.25 ppm yr⁻¹, $p < 0.001$) 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⁻¹, $p < 0.001$), 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.

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), 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 for NO₂ suggests that its overall contribution to the global XCO₂ reconstruction is smaller than that

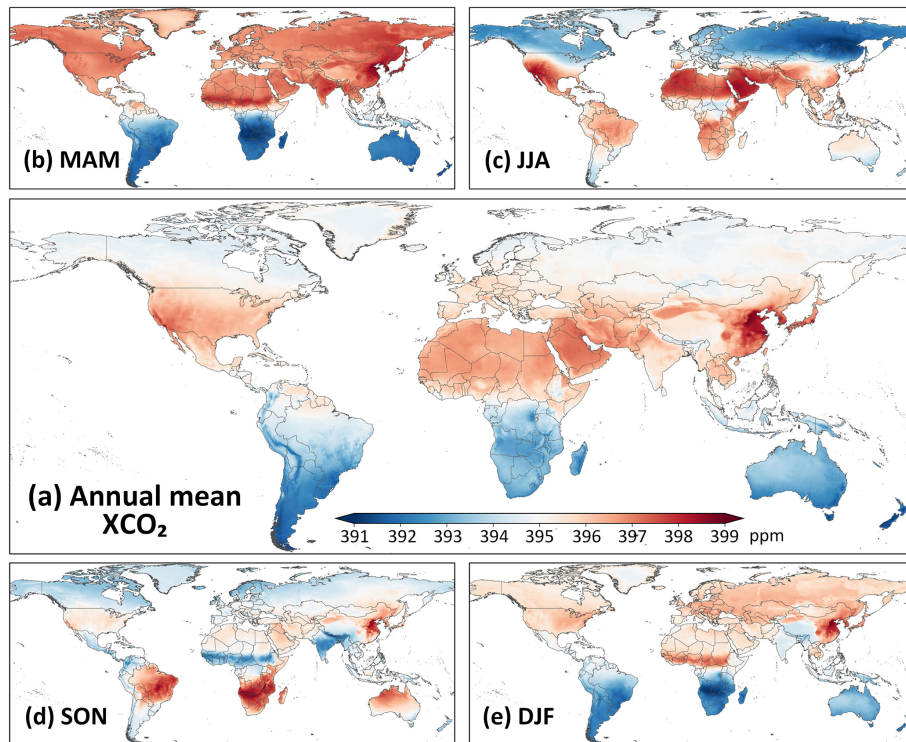


Figure 12. Multi-year (a) annual mean XCO_2 map (unit: ppm) over land for 2003–2022 and seasonal mean XCO_2 maps for (b) March–April–May (MAM), (c) June–July–August (JJA), (d) September–October–November (SON), and (e) December–January–February (DJF).

of variables such as CAMS XCO_2 and meteorological factors. This is expected because NO_2 mainly reflects localized anthropogenic combustion emissions, whereas large-scale atmospheric transport, background concentration fields, and biospheric exchange processes more strongly control XCO_2 variability at the global scale. Nevertheless, NO_2 still provides useful complementary information for identifying regional anthropogenic emission patterns, particularly over urban and industrial areas.

For the bias-correction stage (Fig. 14b), the fused XCO_2 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 XCO_2 structures and reduce systematic biases.

3.3.2 Comparison with previous studies

We compared our reconstructed XCO_2 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 d (Li et al., 2022; Guan et al., 2024), or 3 d (Sheng et al., 2023) scales, which limits their ability to capture short-term (daily) variability. In addition, many studies generated XCO_2 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., 2024b; 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., 2024b; Li et al., 2026; Lee et al., 2025; Yu et al., 2026). More importantly, previous studies with fine temporal resolution (daily or subdaily) exhibit greater fluctuations in overall accuracy, with R^2 ranging from 0.91 to 0.98 and RMSE values from 1.06 to 2.62 ppm. However, a recent study (Wang, 2026) provides a long-term global daily XCO_2 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 XCO_2 dataset at 0.1° spatial resolution over 2003–2022, achieving an R^2 of 0.99 and an RMSE of 1.10 ppm, comparable to or better than those reported in the literature.

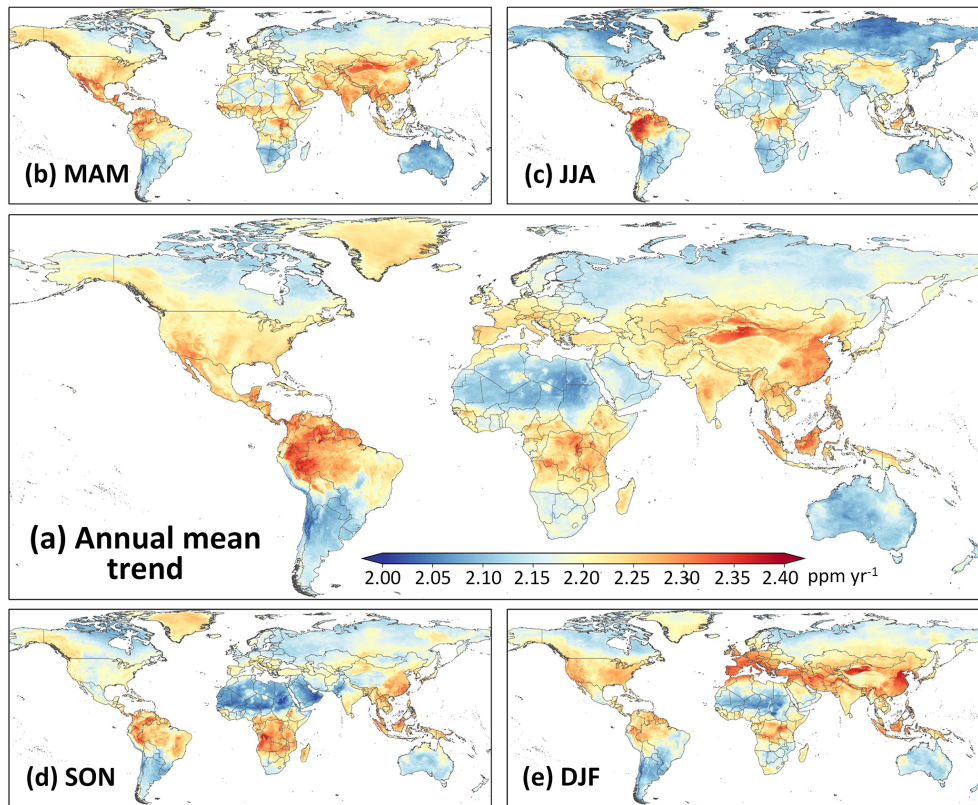


Figure 13. Spatial patterns of long-term (a) annual XCO₂ trends (unit: ppm yr⁻¹) from 2003 to 2022 and seasonal XCO₂ trends (unit: ppm yr⁻¹) for (b) March–April–May (MAM), (c) June–July–August (JJA), (d) September–October–November (SON), and (e) December–January–February (DJF).

Table 3. Comparison of reconstructed global XCO₂ datasets reported in 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-BiLSTM	Monthly	0.05°	0.92	1.54	2015–2021	Wang et al. (2025)
Stacking model	8 d	0.05°	0.98	1.03	2000–2020	Guan et al. (2024)
Extreme tree	8 d	0.01°	0.83	1.79	2014–2018	Li et al. (2022)
Kriging	3 d	1°	0.97	0.88	2009–2023	Sheng et al. (2023)
M-FRK	Daily	1°	0.93	1.06	2016–2019	Huang et al. (2024b)
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)
DINCAE	Daily	0.1°	0.90	1.44	2018–2022	Antezana Lopez et al. (2025)
Stacking model	Daily	0.1°	0.94	1.08	2016–2020	Li et al. (2026)
Random Forest	Daily	1 km × 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-BiLSTM: 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.

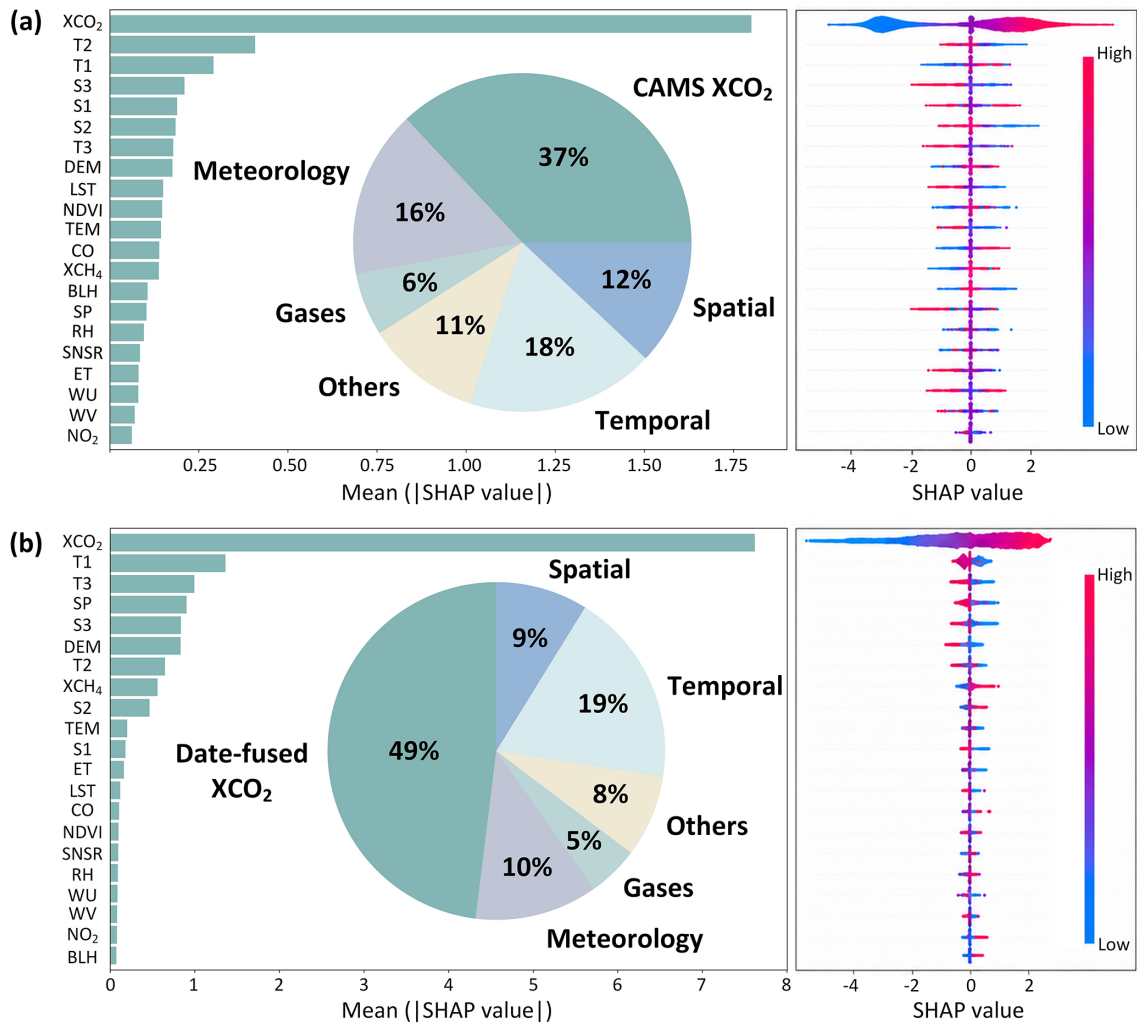


Figure 14. Model interpretability analysis using the SHAP approach, showing the relative importance of each predictor in the (a) data-fusion phase and (b) bias-correction phase.

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 arising from differences in sensor characteristics, orbital sampling, retrieval algorithms, and mission transitions.

These inconsistencies can introduce artificial discontinuities and temporal drift, compromising 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. Compared with the uncorrected data-fused product, the bias-corrected XCO₂ dataset exhibits improved agreement with TCCON, with RMSEs of 0.36, 1.26, and 0.97 ppm under sample-based, temporal-based, and spatial-based ten-fold cross-validation, respectively. These findings highlight the importance of cross-mission harmonization for constructing consistent and 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 Transformer module leverages self-attention mechanisms to characterize non-local spatial relationships and global contextual information, while the BiLSTM module extracts temporal features to preserve daily continuity and capture temporal evolution. 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 seamless daily XCO₂ fields while better capturing temporal variability driven by atmospheric transport, biospheric carbon exchange, and anthropogenic emissions.

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

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

Therefore, the novelty of our study lies not in generating another high-resolution XCO₂ dataset but in developing 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 regard, our dataset complements previous ultra-high-resolution products (Wang, 2026), which emphasize fine-scale spatial mapping, by providing a framework optimized for long-term temporal analyses and climate applications that require stable cross-mission continuity.

4 Data availability

The daily high-resolution seamless global XCO₂ product (GlobalHighXCO₂) is publicly available at <https://doi.org/10.5281/zenodo.18220962> (Qu and Wei, 2026).

5 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-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 harmonizes inter-satellite differences to generate a physically consistent long-term XCO₂ record.

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 sample-based ten-fold cross-validation after bias correction (CV- $R^2 = 0.99$, RMSE = 0.36 ppm). The daily XCO₂ dataset enables the detection of rapid temporal variations and the characterization of localized XCO₂ enhancement patterns associated with a wide range of emission sources, including urban activities, industrial emissions, and biomass burning. The dataset also effectively captures strong responses of XCO₂ growth to major climate events, such as intense El Niño episodes and large-scale wildfires. Across temporal scales, XCO₂ increased significantly at a global mean rate of approximately 2.24 ppm yr⁻¹ ($p < 0.001$) from 2003 to 2022. Spatially, the largest growth rates occurred over northern South America, East and Southeast Asia, and central Africa, while relatively weaker increases were observed 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, urban emissions, and expanding ground-based validation networks.

Author contributions. JW conceived and supervised the study, managed project administration, and acquired funding. YQ, XS, YF, and ZW collected the data, performed the formal analysis and investigation, and wrote the original draft of the manuscript.

Competing interests. At least one of the (co-)authors is a member of the editorial board of *Earth System Science Data*. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. The authors bear the ultimate responsibility for providing appropriate place names. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

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

Financial support. This work was supported by the Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (grant no. JYB2025XDXM906), the Jing-Jin-Ji Regional Integrated Environmental Improvement-National Science and Technology Major Project (grant no. 2026ZD1212900), the National Key Technology and Development Program of Corps (grant no. 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 (grant no. 2025KYLX064).

Review statement. This paper was edited by Yuqiang Zhang and reviewed by two anonymous referees.

References

- Agustí-Panareda, A., Barré, J., Massart, S., Inness, A., Aben, I., Ades, M., Baier, B. C., Balsamo, G., Borsdorff, T., and Bousseres, N.: The CAMS greenhouse gas reanalysis from 2003 to 2020, *Atmos. Chem. Phys.*, 23, 3829–3859, <https://doi.org/10.5194/acp-23-3829-2023>, 2023.
- Ak-Bhd, M.: WMO greenhouse gas bulletin, World Meteorological Organization: Geneva, Switzerland, <https://gaw.kishou.go.jp/publications/bulletin> (last access: 14 May 2026), 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, *Sci. Data*, 12, 268, <https://doi.org/10.1038/s41597-024-04135-w>, 2025.
- Betts, R. A., Jones, C. D., Knight, J. R., Keeling, R. F., and Kennedy, J. J.: El Niño and a record CO₂ rise, *Nat. Clim. 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, *J. Atmos. Sci.*, 56, 127–150, [https://doi.org/10.1175/1520-0469\(1999\)056<0127:SMOAMM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2), 1999.
- 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, *Annu. Rev. Earth Planet. Sci.*, 47, 555–581, <https://doi.org/10.1146/annurev-earth-082517-010235>, 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, *Atmos. Chem. Phys.*, 5, 3313–3329, <https://doi.org/10.5194/acp-5-3313-2005>, 2005.
- 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 Sens. Environ.*, 162, 344–362, <https://doi.org/10.1016/j.rse.2013.04.024>, 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, *Atmos. Chem. Phys.*, 18, 17355–17370, <https://doi.org/10.5194/acp-18-17355-2018>, 2018.
- 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, *Geophys. Res. Lett.*, 38, <https://doi.org/10.1029/2011GL047888>, 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, <https://doi.org/10.1126/science.aam5776>, 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 J. Photogram. Remote Sens.*, 208, 107–120, <https://doi.org/10.1016/j.isprsjprs.2024.01.009>, 2024.
- Crippa, M., Guizzardi, D., Pisoni, E., Solazzo, E., Guion, A., Muntean, M., Florczyk, A., Schiavina, M., Melchiorri, M., and Hufilter, A. F.: Global anthropogenic emissions in urban areas: patterns, trends, and challenges, *Environ. Res. Lett.*, 16, 074033, <https://doi.org/10.1088/1748-9326/ac00e2>, 2021.
- Crisp, D.: Measuring atmospheric carbon dioxide from space with the Orbiting Carbon Observatory-2 (OCO-2), *Proc. SPIE 9607, Earth Observing Systems XX*, 960702, <https://doi.org/10.1117/12.2187291>, 2015.
- 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, *Atmos. Meas. Tech.*, 10, 59–81, <https://doi.org/10.5194/amt-10-59-2017>, 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, *Adv. Space Res.*, 73, 4158–4174, <https://doi.org/10.1016/j.asr.2024.01.016>, 2024.
- Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S.

- R., Anthoni, P., Barbero, L., Bates, N. R., Becker, M., Bellouin, N., Decharme, B., Bopp, L., Brasika, I. B. M., Cadule, P., Chamberlain, M. A., Chandra, N., Chau, T.-T.-T., Chevallier, F., Chini, L. P., Cronin, M., Dou, X., Enyo, K., Evans, W., Falk, S., Feely, R. A., Feng, L., Ford, D. J., Gasser, T., Ghattas, J., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Heinke, J., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jacobson, A. R., Jain, A., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Joos, F., Kato, E., Keeling, R. F., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Körtzinger, A., Lan, X., Lefèvre, N., Li, H., Liu, J., Liu, Z., Ma, L., Marland, G., Mayot, N., McGuire, P. C., McKinley, G. A., Meyer, G., Morgan, E. J., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O'Brien, K. M., Olsen, A., Omar, A. M., Ono, T., Paulsen, M., Pierrot, D., Pockock, K., Poulter, B., Powis, C. M., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T. M., Schwinger, J., Séférian, R., Smallman, T. L., Smith, S. M., Sospedra-Alfonso, R., Sun, Q., Sutton, A. J., Sweeney, C., Takao, S., Tans, P. P., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., van Ooijen, E., Wanninkhof, R., Watanabe, M., Wimart-Rousseau, C., Yang, D., Yang, X., Yuan, W., Yue, X., Zaehle, S., Zeng, J., and Zheng, B.: Global Carbon Budget 2023, *Earth Syst. Sci. Data*, 15, 5301–5369, <https://doi.org/10.5194/essd-15-5301-2023>, 2023.
- 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, *Sci. Total Environ.*, 892, 164750, <https://doi.org/10.1016/j.scitotenv.2023.164750>, 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, *Sci. Total Environ.*, 957, 177051, <https://doi.org/10.1016/j.scitotenv.2024.177051>, 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, *Atmos. Chem. Phys.*, 23, 5355–5372, <https://doi.org/10.5194/acp-23-5355-2023>, 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, *J. Geophys. Res.-Atmos.*, 117, <https://doi.org/10.1029/2011JD017015>, 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, *Geophys. Res. Lett.*, 49, e2022GL098435, <https://doi.org/10.1029/2022GL098435>, 2022.
- 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, *J. Clean. Product.*, 428, 139290, <https://doi.org/10.1016/j.jclepro.2023.139290>, 2023a.
- 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, <https://doi.org/10.3390/atmos14030436>, 2023b.
- 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 Sens.*, 12, 576, <https://doi.org/10.3390/rs12030576>, 2020.
- 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, *Int. J. Appl. Earth Obs. Geoinf.*, 129, 103859, <https://doi.org/10.1016/j.jag.2024.103859>, 2024.
- 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 Sens.*, 16, 2433, <https://doi.org/10.3390/rs16132433>, 2024.
- 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, *J. Geophys. Res.-Biogeophys.*, 120, 2306–2325, <https://doi.org/10.1002/2015JG003060>, 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, *Geophys. Res. Lett.*, 51, e2023GL107536, <https://doi.org/10.1029/2023GL107536>, 2024a.
- 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, *J. Appl. Remote Sens.*, 18, 028502, <https://doi.org/10.1117/1.JRS.18.028502>, 2024b.
- Hwang, S., Choi, H., Kang, Y., and Im, J.: Reconstructing long-term (2003–2019) global high-resolution XCO₂: bridging observational gaps with machine learning, *GISci. Remote Sens.*, 63, 2627042, <https://doi.org/10.1080/15481603.2026.2627042>, 2026.
- 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, <https://gml.noaa.gov/ccgg/carbontracker/> (last access: 14 May 2026), 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, *Atmos. Res.*, 279, 106385, <https://doi.org/10.1016/j.atmosres.2022.106385>, 2022.
- 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, *P. Natl. Acad. Sci. USA*, 119, e2108146119, <https://doi.org/10.1073/pnas.2108146119>, 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, *J. Climate*, 29, 8783–8805, <https://doi.org/10.1175/JCLI-D-14-00672.1>, 2016.
- 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, *Appl. Optics*, 48, 6716–6733, <https://doi.org/10.1364/AO.48.006716>, 2009.
- 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, in: Contribution of working groups I, II and III to the sixth assessment report of the intergov-

- ernmental panel on climate change, edited by: core writing team, Lee, H., and Romero, J., IPCC, Geneva, Switzerland, <https://doi.org/10.59327/IPCC/AR6-9789291691647.001>, 2023.
- 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, *J. Geophys. Res.-Atmos.*, 130, e2024JD042809, <https://doi.org/10.1029/2024JD042809>, 2025.
- 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, *Int. J. Appl. Earth Obs. Geoinf.*, 108, 102743, <https://doi.org/10.1016/j.jag.2022.102743>, 2022.
- 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, *Int. J. Appl. Earth Obs. Geoinf.*, 146, 105092, <https://doi.org/10.1016/j.jag.2026.105092>, 2026.
- 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 Sens. Environ.*, 304, 114039, <https://doi.org/10.1016/j.rse.2024.114039>, 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 T. Geosci. Remote*, 62, 4111411, <https://doi.org/10.1109/TGRS.2024.3483776>, 2024b.
- 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 T. Geosci. Remote*, 63, 4102716, <https://doi.org/10.1109/TGRS.2025.3540289>, 2025.
- Liang, A., Pang, R., Chen, C., and Xiang, C.: XCO₂ Fusion algorithm based on multisource greenhouse gas satellites and carbontracker, *Atmosphere*, 14, 1335, <https://doi.org/10.3390/atmos14091335>, 2023.
- 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, *Ecol. Inform.*, 83, 102806, <https://doi.org/10.1016/j.ecoinf.2024.102806>, 2024.
- 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, *Nat. Geosci.*, 15, 615–620, <https://doi.org/10.1038/s41561-022-00965-8>, 2022.
- Lundberg, S. M. and Lee, S.-I.: A unified approach to interpreting model predictions, *arXiv [preprint]* <https://doi.org/10.48550/arXiv.1705.07874>, 2017.
- 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 T. Geosci. Remote*, 59, 8084–8093, <https://doi.org/10.1109/TGRS.2021.3052215>, 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*, 67, 28452, <https://doi.org/10.3402/tellusb.v67.28452>, 2015.
- Qu, Y. and Wei, J.: GlobalHighXCO₂: Global Daily Seamless 10 km XCO₂ Dataset Over Land (2003–Present), Zenodo [data set], <https://doi.org/10.5281/zenodo.18220962>, 2026.
- 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 T. Geosci. Remote*, 63, 4109012, <https://doi.org/10.1109/TGRS.2025.3593486>, 2025.
- 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, *J. Geophys. Res.-Atmos.*, 116, <https://doi.org/10.1029/2010JD015047>, 2011.
- 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: collocated regional CO₂ and NO₂ enhancements observed by the OCO-2 and S5P satellites, *Atmos. Chem. Phys.*, 19, 9371–9383, <https://doi.org/10.5194/acp-19-9371-2019>, 2019.
- 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, https://doi.org/10.1007/978-3-031-84841-4_8, 2025.
- Romanov, V.: *Greenhouse Gases and Clay Minerals: Enlightening Down-to-Earth Road Map to Basic Science of Clay-Greenhouse Gas Interfaces*, Springer, <https://doi.org/10.1007/978-3-319-12661-6>, 2017.
- 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 Sens.*, 13, 3524, <https://doi.org/10.3390/rs13173524>, 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, <https://doi.org/10.1080/20964471.2022.2033149>, 2023.
- Siabi, Z., Falahatkar, S., and Alavi, S. J.: Spatial distribution of XCO₂ using OCO-2 data in growing seasons, *J. Environ. Manage.*, 244, 110–118, <https://doi.org/10.1016/j.jenvman.2019.05.049>, 2019.
- Tian, W., Zhang, L., Yu, T., Wu, Y., Zhang, W., Wang, Z., and Zhu, H.: Using multisource data and time series features to construct a global terrestrial CO₂ coverage by deep learning, *IEEE T. Geosci. Remote*, 62, 4109814, <https://doi.org/10.1109/TGRS.2024.3462589>, 2024.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I.: Attention is all you need, *arXiv [preprint]*, <https://doi.org/10.48550/arXiv.1706.03762>, 2017.
- Wang, J.: Global daily 1 km gapless XCO₂ (2003–2023) derived from multi-satellite observations and a spatiotemporal deep learning framework, *Environ. Imp. Assess. Rev.*, 117, 108146, <https://doi.org/10.1016/j.eiar.2025.108146>, 2026.
- Wang, J., Liu, Z., Zeng, N., Jiang, F., Wang, H., and Ju, W.: Spaceborne detection of XCO₂ enhancement induced by

- Australian mega-bushfires, *Environ. Res. Lett.*, 15, 124069, <https://doi.org/10.1088/1748-9326/abc846>, 2020.
- 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, <https://doi.org/10.1038/nature12915>, 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 (CAMS-EGG4) with a spatiotemporally self-supervised fusion method, *Earth System Science Data*, 15, 3597–3622, <https://doi.org/10.5194/essd-15-3597-2023>, 2023.
- 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° resolution from 2015 to 2021 for exploring global carbon dynamics, *Earth Syst. Sci. Data*, 17, 5355–5375, <https://doi.org/10.5194/essd-17-5355-2025>, 2025.
- Wei, J., Li, Z., Chen, X., Li, C., Sun, Y., Wang, J., Lya-pustin, 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, *Environ. Sci. Technol.*, 57, 18282–18295, <https://doi.org/10.1021/acs.est.3c00272>, 2023.
- 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 Sens. Environ.*, 315, 114404, <https://doi.org/10.1016/j.rse.2024.114404>, 2024.
- Wolter, K. and Timlin, M. S.: El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI. ext), *Int. J. Climatol.*, 31, 1074–1087, <https://doi.org/10.1002/joc.2336>, 2011.
- 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, *Sci. Total Environ.*, 954, 176171, <https://doi.org/10.1016/j.scitotenv.2024.176171>, 2024.
- 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, *Atmos. Meas. Tech.*, 10, 2209–2238, <https://doi.org/10.5194/amt-10-2209-2017>, 2017.
- Yang, D., Liu, Y., Cai, Z., Chen, X., Yao, L., and Lu, D.: First global carbon dioxide maps produced from TanSat measurements, *Adv. Atmos. Sci.*, 35, 621–623, <https://doi.org/10.1007/s00376-018-7312-6>, 2018.
- Yang, H., Li, T., Wu, J., Wang, Y., Zhang, L., and Hu, R.: Global 20-year XCO₂ mapping through synergy of multi-satellite observations, *Geosci. Front.*, 17, 102333, <https://doi.org/10.1016/j.gsf.2026.102333>, 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 T. Geosci. Remote*, 60, 4103612, <https://doi.org/10.1109/TGRS.2021.3107542>, 2021.
- 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, *Atmos. Res.*, 334, 108747, <https://doi.org/10.1016/j.atmosres.2026.108747>, 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, *Environ. Model. Softw.*, 124, 104600, <https://doi.org/10.1016/j.envsoft.2019.104600>, 2020.
- 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, *Environ. Int.*, 178, 108057, <https://doi.org/10.1016/j.envint.2023.108057>, 2023.
- Zhang, M. and Liu, G.: Mapping contiguous XCO₂ by machine learning and analyzing the spatio-temporal variation in China from 2003 to 2019, *Sci. Total Environ.*, 858, 159588, <https://doi.org/10.1016/j.scitotenv.2022.159588>, 2023.