https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

Open Access Earth System Science Data Discussions

Signal-Domain Guided Deep Learning for Gap-Filling of XCO and XCH_4: A Masked Spatio-Temporal Fusion of TROPOMI and GEOS-Chem (2019–2023)

Zhiwei Li [a], Yuan Tian [a, b, *], Peize Lin [a], Bowen Chang [a], Jingkai Xue [c]

[a] Institutes of Physical Science and Information Technology, Anhui University, Hefei 230601, China

[b] National Key Laboratory of Opto-Electronic Information Acquisition and Protection Technology, Institutes of Physical Science and Information Technology, Anhui University, Hefei, Anhui, P.R. China

[c] School of Environmental Science and Optoelectronic Technology, University of Science and Technology of China, Hefei 230026, China

* Corresponding author: Yuan Tian (ytian@ahu.edu.cn)

Abstract

Long-term, high-resolution monitoring of carbon monoxide (CO) and methane (CH_4) is essential for understanding their spatiotemporal variability and guiding climate mitigation strategies. However, satellite observations like TROPOMI are often incomplete, and existing fusion methods have limitations in accuracy and continuity. This study proposes a signal-domain fusion approach combining 3D discrete cosine transform (DCT) and singular value decomposition (SVD) to integrate TROPOMI data with GEOS-Chem simulations. A lightweight residual U-Net is employed to refine the initial reconstruction by learning the residual field using meteorological drivers and model outputs, guided by a masked loss. The method produces global 0.25° and China-specific 0.05° daily gap-free XCO and XCH_4 datasets from 2019 to 2023. The fused results outperform GEOS-Chem and are comparable or superior to TROPOMI, with R^2 values of 0.92 for XCO and 0.85 for XCH_4. Trend analysis reveals regional patterns such as XCO increases in North America and declines in Eastern China, and widespread CH_4 growth. High-resolution data captures enhancements during the 2022 Chongqing wildfires, with average increases of 17.1 ppb in XCO and 24.5 ppb in XCH_4, and reveals lower XCH_4 increases over rice-growing areas compared to

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

Open Access Earth System Science Data Discussions

TROPOMI, with overestimation reduced by 17–26%, and stronger XCO reductions, with satellite underestimations up to 38%. These results highlight agricultural contributions and policy impacts. This approach effectively reconstructs missing observations and enhances the utility of satellite–model data for atmospheric research and emission assessments. The generated daily gap-free datasets are publicly available at https://doi.org/10.5281/zenodo.17936461.

Keywords: CO; CH_4; TROPOMI; GEOS-Chem; Signal Domain Fusion; Deep Learning; Spatiotemporal Reconstruction

1. Introduction

Carbon monoxide (CO) and methane (CH_4) are two critical atmospheric gases that play a key role in air quality monitoring and climate change research. CO is primarily produced through incomplete combustion of fossil fuels and biomass, as well as the oxidation of methane and other hydrocarbons. Its main sink is the reaction with hydroxyl radicals (OH), which contributes to the formation of tropospheric ozone (O_3) and greenhouse gases such as carbon dioxide (CO_2) (Lelieveld et al., 2016; Spivakovsky et al., 2000). Due to its atmospheric lifetime of weeks to months, CO serves as an effective tracer for pollutant transport, facilitating the study of both horizontal and vertical atmospheric movements (Heald et al., 2003). Methane (CH_4) is a powerful greenhouse gas with a significantly longer atmospheric lifespan—around ten years—and a greater global warming potential compared to numerous other gases (Filonchyk et al., 2024; Heilig, 1994). It originates from both natural sources, such as wetlands, permafrost, and wildfires, and anthropogenic activities, including agriculture, livestock digestion, landfills, and fossil fuel extraction (Chai et al., 2016; Jackson et al., 2020). CH_4 plays a crucial role in atmospheric chemistry, influencing the oxidative capacity of the atmosphere and contributing to tropospheric ozone formation. Its rising concentration is a major driver of climate change.

CO provides valuable insights into pollution transport and atmospheric chemistry, while CH_4 monitoring is critical for assessing greenhouse gas emissions and climate impacts. Both CO and CH_4 are essential targets for monitoring efforts due to their distinct roles in atmospheric processes. A comprehensive understanding of these gases is fundamental to improving air quality predictions and formulating climate mitigation policies. The spatial distribution of CO and CH_4 in the

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

atmosphere is now routinely determined through remote sensing observations. The source, transport, and removal processes of pollutants can be effectively tracked and analyzed through global-scale monitoring through satelite platforms. At present, a number of satellite instruments are equipped with the ability to observe CO and CH_4, including the Measurement of Pollutants in the Troposphere (MOPITT) (Deeter et al., 2003), Atmospheric Infrared Sounder (AIRS) (McMillan et al., 2005), Tropospheric Emission Spectrometer (TES) (Rinsland et al., 2006), and Interferometer for the Exploration of the Atmosphere in the Infrared (IASI) (Turquety et al., 2004). These instruments offer critical data support for the global distribution of CO and CH_4 in the atmosphere.

The TROPOMI instrument of the Sentinel-5P mission of the European Space Agency (ESA) provides a greater degree of daily global coverage and spatial resolution than previous observing missions. TROPOMI has been able to monitor the total atmospheric XCO and XCH_4 (column concentrations of atmospheric CO and CH_4, respectively) through spectroscopic measurements in the ultraviolet (UV) to short-wave infrared (SWIR) bands, with a spatial resolution of approximately 7 × 7 km and high radiometric precision, since its successful launch on October 13, 2017. TROPOMI has already been employed in numerous pertinent applications for XCO and XCH_4 investigations, including the calculation of emissions from biomass combustion (Goudar et al., 2023; Griffin et al., 2024). Nevertheless, the TROPOMI observations of XCO and XCH_4 at the surface and in the atmosphere are impeded by the ability of clouds and atmospheric aerosols to block or deflect reflected sunlight observed by satellite sensors. Observations may be incomplete, particularly in regions with elevated levels of air pollution or overcast skies.

The continuous coverage of XCO and XCH_4 products for TROPOMI has been the subject of numerous efforts, which primarily fall into three categories. On one hand, machine learning-based interpolation methods (X. Chen et al., 2022; Hu et al., 2022; Valerio et al., 2025; Wei et al., 2025) can effectively address data gaps. However, their precision is contingent upon the quality and diversity of the training data, and inaccurate predictions may arise from insufficient or unbalanced datasets. In addition to these approaches, model fusion methods have been developed to leverage the complementary strengths of different data sources or modeling frameworks. For instance, statistical data assimilation and hybrid modeling schemes combine outputs from chemical transport models (Fritz et al., 2022; Schneising et al., 2023; Sicard et al., 2021) with satellite retrievals to generate more spatially consistent and temporally continuous products. By integrating physical

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

model constraints with empirical corrections, these methods can effectively reduce retrieval bias and improve gap-filling robustness(Inness et al., 2022; Wang et al., 2023). However, their performance often depends on the accuracy of prior model simulations and the representativeness of assimilated observations.On the other hand, enhanced spectral fitting algorithms (Borsdorff et al., 2019; Guanter et al., 2015; Schneising et al., 2023; Wang et al., 2020) offer an alternative approach. These algorithms can mitigate bias from fluctuations in spectral reflectance by increasing the order of the polynomial fit. Nevertheless, this heightened complexity often demands greater computational resources and leads to longer processing times.

This study aims to develop a robust framework for generating daily global and regional continuous XCO and XCH_4 products over the period 2019–2023, at a spatial resolution of 0.25° globally and 0.05° over China. The core objective is to address the frequent data gaps in satellite observations, especially those caused by cloud coverage and retrieval errors in the TROPOMI sensor, by leveraging complementary information from chemical transport modeling and frequency-domain representations.

To achieve this, we propose a signal-domain guided spatio-temporal fusion method that integrates GEOS-Chem simulations with TROPOMI observations. Our approach consists of two stages: (1) a low-rank signal-domain reconstruction using 3D Discrete Cosine Transform (DCT) (Rao and Yip, 2014) and Singular Value Decomposition (SVD) (Wall et al., 2003), which exploits the shared spatio-temporal structure between model and satellite data to approximate missing values, and (2) a learning-based refinement module that employs a lightweight residual U-Net to predict pixel-level corrections based on GEOS-Chem output, reconstructed fields, and meteorological variables(Ronneberger et al., 2015; Tang, n.d.).

Instead of using the TROPOMI data mask as a direct input, we apply it as a spatial constraint during model training to focus learning only on valid observations(Wei et al., 2022). This masked learning strategy improves generalization while preserving physical consistency.

By fusing model-driven priors and observational constraints in both frequency and spatial domains, our method significantly improves the completeness and accuracy of XCO and XCH_4 datasets. Validation against held-out TROPOMI data shows that the fused outputs outperform GEOS-Chem alone and remain comparable to or better than TROPOMI retrievals in cloud-covered regions. The proposed framework provides an efficient and interpretable solution for large-scale

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

trace gas monitoring and offers new opportunities for atmospheric data assimilation and long-term climate analysis.

2. Measurement and materials

2.1. Data description

2.1.1. TROPOMI XCO and XCH_4 product

This study employs TROPOMI Level-2 data products of column-averaged CO (XCO) and CH_4 (XCH_4). Mounted on the Sentinel-5 Precursor (S5P) satellite, the TROPOspheric Monitoring Instrument (TROPOMI) functions in a polar sun-synchronous orbit. The Sentinel-5 Precursor (S5P) satellite is equipped with the TROPOMI, which operates in a polar sun-synchronous orbit. TROPOMI provides daily global XCO measurements at 13:30 local solar time, thereby facilitating daily global coverage. The pixel resolution of TROPOMI's XCO and XCH_4 data has been enhanced from 7.0 × 7.0 km^2 to 7.0 × 5.5 km^2 as of June 2019.

By measuring the Earth's radiation in the short-wave infrared (SWIR) spectral range (2305-2385 nm), TROPOMI inverts XCO and XCH_4. TROPOMI exhibits a high sensitivity to the tropospheric boundary layer for XCO in clear-sky conditions, but this sensitivity is subject to variation based on the optical path in overcast conditions. TROPOMI employs absorption information in the oxygen-A band (760 nm) and the SWIR spectral range (2305-2385 nm) for inversion of XCH_4.

In order to guarantee the quality of the data, we implemented rigorous data selection criteria and exclusively selected data points with a Quality Value (QV) greater than 0.5 for analysis (Kawka et al., 2021). The data were obtained from a secondary offline and reprocessed product provided by NASA (https://tropomi.gesdisc.eosdis.nasa.gov/data/) for the period 2019–2023. The data regarding XCO and XCH_4 are presented in the form of column-averaged dry air mole fractions in parts per billion (ppb) in this study.

2.1.2. GEOS-Chem Chemical Transport Model

This investigation employs the GEOS-Chem model, a three-dimensional global atmospheric chemistry model that is comprehensive and employs Goddard Earth Observation System (GEOS) meteorological data supplied by the NASA Global Modelling and Assimilation Office (GMAO).

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

Global Observation System (GEOS) meteorological data supplied by the NASA Global Modelling and Assimilation Office (GMAO). We employed version 14.1.1 (http://acmg.seas.harvard.edu/geos/), which is powered by GEOS-FP meteorological data. The most recent GEOS-5 meteorological data product to be provided by NASA/GMAO is GEOS-FP ("forward processing") (http://gmao.com/gmao.html). This product has a native horizontal resolution of 0.25° latitude × 0.3125° longitude and a temporal resolution of hourly data and 3-hourly data.

To reduce computational overhead, we employed a 2°×2.5° GEOS-Chem horizontal grid for the global XCO and XCH_4 simulations and a 0.25°×0.3125° GEOS-Chem horizontal grid for the XCO and XCH_4 simulations for the China region in order to achieve more precise results. The model calculates the concentrations of CO and CH_4 by integrating emission inventories from various regions, which are based on surface emissions and chemical reactions.

The following are specific emission inventories: The Air Pollutant Emission Inventory (APEI) v2016 was used to obtain Canadian anthropogenic emissions data, while the National Emission Inventory (NEI) v2015-03 was used to obtain North American regional emissions data. The Diffuse and Inefficient Combustion Emissions in Africa (DICE-Africa) inventory was used to obtain anthropogenic emissions data for the Africa region (Marais and Wiedinmyer, 2016). The MIX v1.1 inventory was used to obtain regional emissions data in Asia (Li et al., 2017) and HTAP. The Community Emissions Data System (CEDS) inventory was used to derive global aircraft and ship emissions data (Hoesly et al., 2018).The Quick Fire Emissions Dataset (QFED): Documentation of Versions 2.1, 2.2, and 2.4 was used to acquire carbon monoxide emissions data from biomass combustion.

The time step for convective and advective transport in the simulation is 300 seconds, while the time step for chemical processes is 600 seconds. The distribution of XCO and XCH_4 concentrations and their alterations from 2019 to 2023 can be efficiently and accurately simulated using the aforementioned setup.

2.1.3. Total Carbon Column Observing Network (TCCON) Measurements

TCCON employs a Fourier Transform Infrared Spectrometer (FTS) to measure direct solar light in order to determine the total column concentrations of greenhouse gases in the atmosphere, including carbon dioxide (CO_2), methane (CH_4), carbon monoxide (CO), and others (Buschmann et

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

al., 2016; Kiel et al., 2016; Sha et al., 2019; Yang et al., 2020). These data are extensively utilized to validate satellite remote sensing data (e.g., G. Chander et al., 2013; Imasu et al., 2023; Lin et al., 2024; Loew et al., 2017; Wu et al., 2019; etc.) and to evaluate the performance of climate models. They are rigorously calibrated and validated with high accuracy and reliability. TCCON's observing stations are situated in numerous regions worldwide, providing a comprehensive understanding of the global greenhouse gas distribution. These stations are capable of observing a diverse array of climates and ecosystems. Harmonized processing is implemented for each site's data. In order to guarantee data consistency and comparability, the data from each station is processed in a consistent manner. TCCON's XCO and XCH_4 Dry Air Mole Fraction (Xgas) in ppb are employed in this investigation. Site-specific and time-scale data are accessible through the official website of TCCON (https://tccon.ornl.gov/), and users may select the data that is most relevant to their research requirements. The validation of XCO and XCH_4 products from TROPOMI and GEOS-Chem is frequently conducted using TCCON (Borsdorff et al., 2019; Cogan et al., 2012; Inness et al., 2022; Schneising et al., 2019). Figure 1 illustrates the global TCCON site location map, which comprises operational, prospective, and former sites, accordingly. The TCCON data version GGG2020 (Laughner et al., 2023) was employed in this experiment. Table 1 provides a list of the sites that were utilized and are cross-correlated with the time frames in our experiment.

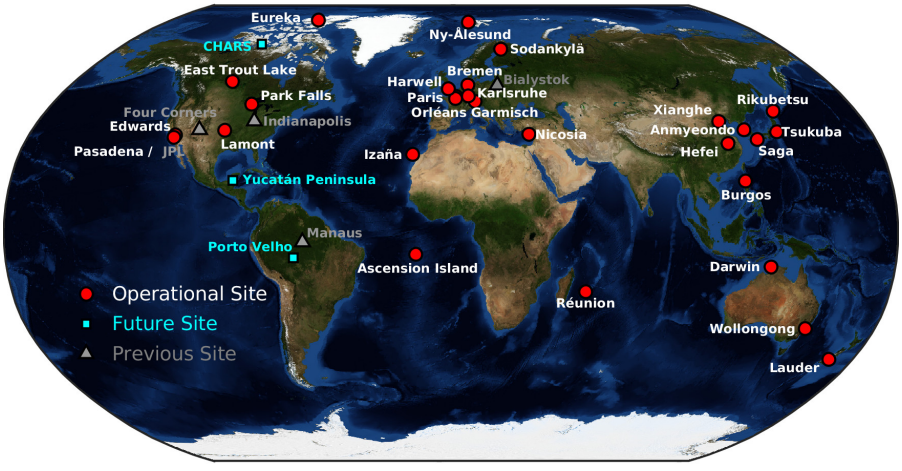

Fig. 1 Map showing global TCCON site locations, including operating sites, potential sites, and past sites, respectively.

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

Table 1. Details of the TCCON sites used in this study, No. is the serial number.

No.	Site name	Latitude	Longitude	Location	Start date	End date
1	bremen01	53.104	8.85	Bremen, Germany	2009/1/6	2021/6/24
2	burgos01	18.5325	120.6496	Burgos, Philippines	2017/3/3	2023/6/23
3	darwin01	-12.425	130.891	Darwin, Australia	2013/1/1	2022/12/27
4	easttroutlake01	54.354	-104.987	East Trout Lake, Canada	2016/10/3	2024/2/15
5	edwards01	34.9599	-117.881	AFRC, Edwards, CA, USA	2013/7/20	2024/2/22
6	garmisch01	47.476	11.063	Garmisch, Germany	2007/7/18	2023/5/4
7	hefei01	31.91	117.17	Hefei, China	2015/11/2	2023/12/25
8	izana01	28.31	-16.5	Izana, Tenerife, Spain	2014/1/2	2023/8/30
9	karlsruhe01	49.103	8.44	Karlsruhe, Germany	2014/1/15	2023/6/26
10	lamont01	36.604	-97.486	Lamont, Oklahoma, USA	2011/4/16	2024/2/25
11	lauder03	-45.038	169.684	Lauder, New Zealand	2018/10/2	2023/12/28
12	nicosia01	35.141	33.381	Nicosia, Cyprus	2019/9/1	2023/5/10
13	orleans01	47.97	2.113	Orleans, France	2009/9/6	2023/6/23
14	parkfalls01	45.945	-90.273	Park Falls, Wisconsin, USA	2004/6/2	2024/2/25
15	pasadena01	34.136	-118.127	Pasadena, California, USA	2012/9/20	2024/2/25
16	reunion01	-20.901	55.485	Reunion Island, France	2015/3/1	2020/7/18
17	rikubetsu01	43.4567	143.7661	Rikubetsu, Hokkaido, Japan	2014/6/24	2021/6/30
18	sodankyla01	67.367	26.631	Sodankylä, Finland	2009/5/16	2023/5/30
19	tsukuba02	36.0513	140.1215	Tsukuba, Ibaraki, Japan, 125HR	2014/3/28	2021/3/31
20	wollongong01	-34.406	150.891	Wollongong, Australia	2013/1/4	2023/6/27
21	xianghe01	39.8	116.96	Xianghe, China	2018/6/14	2023/5/29

2.2. Methodology

In this study, we propose a two-stage fusion framework that integrates physical modeling, signal-domain reconstruction, and deep learning-based residual correction to achieve continuous and accurate global mapping of atmospheric trace gases. The overall workflow is illustrated in Figure 2, which presents the main components and their interconnections. In the first stage, a signal-domain spatio-temporal reconstruction is employed to exploit the low-frequency consistency and spatio-temporal correlations between the TROPOMI observations and GEOS-Chem simulations, effectively filling missing data regions caused by cloud cover or instrument limitations. In the second stage, a residual learning network based on a lightweight residual U-Net is introduced to refine the fused results by learning nonlinear and region-specific discrepancies between the

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

preliminary reconstruction and the true satellite observations.

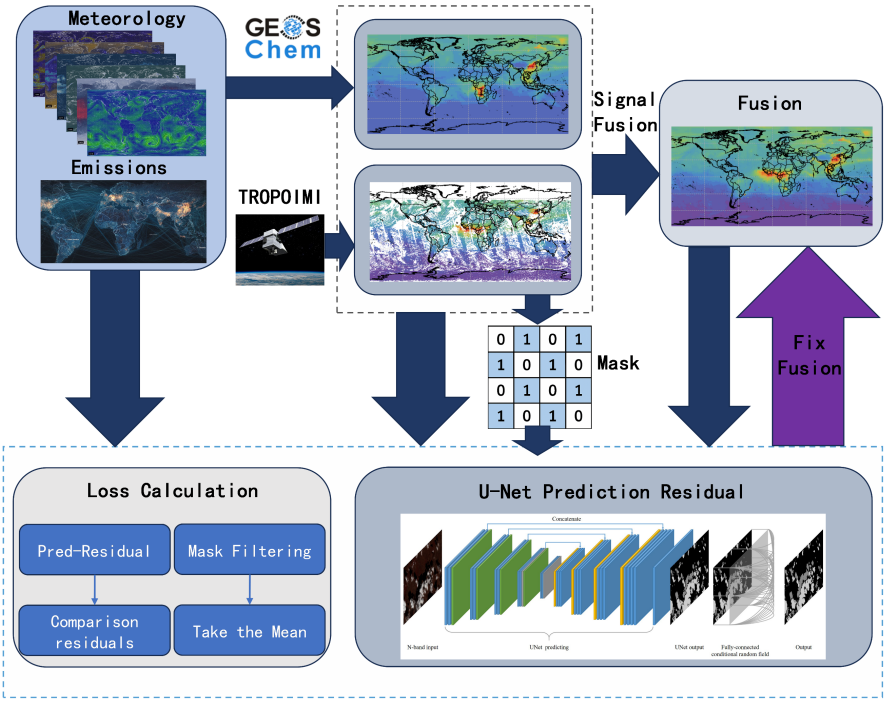

Fig. 2. Overview of the proposed fusion framework combining physical modeling, signal-domain reconstruction, and deep learning-based residual correction. The process includes data preprocessing, DCT&SVD-based spatio-temporal fusion, and residual refinement via a lightweight residual U-Net architecture. Arrows indicate the information flow across different modules.

2.2.1. Data Preprocessing

Data preprocessing is a critical component in guaranteeing the reliability and rationality of the fusion results. In this investigation, values with data quality less than 0.5 in XCO and XCH_4 in TROPOMI are discarded to eliminate inaccurate data and are subsequently aligned to a global coverage of 720 x 1440 (0.25°) horizontal grid through area-weighted aggregation (Wang et al., 2018). The global-scale data from GEOS-Chem simulations should be aligned to the same 0.25° horizontal grid by inverse distance-weighted interpolation, and the TROPOMI data should be maintained at the same resolving power as GEOS-Chem (Setianto and Triandini, 2013). Global-scale data were simulated by GEOS-Chem and aligned to the same 0.25° horizontal grid.

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

To reconcile the heterogeneous spatial supports of satellite retrievals and model simulations, the TROPOMI XCO and XCH_4 data are first filtered by discarding pixels with a quality flag below 0.5 and then regridded to a global 0.25°×0.25° grid (720×1440) using area-weighted aggregation (Wang et al., 2018). The GEOS-Chem outputs are interpolated onto the same 0.25° grid via inverse distance weighting to ensure spatial compatibility between the datasets (Setianto and Triandini, 2013). The choice of 0.25° represents a widely adopted trade-off: it is fine enough to preserve mesoscale spatial gradients without excessively increasing data volume or the computational burden for multi-year global fusion analyses (Hu et al., 2024; Wang et al., 2023).

2.2.2. Spatio-Temporal Data Fusion Method Based on Signal Domain Reconstruction

The accuracy of the inventory typically results in an underestimation of GEOS-Chem simulation data, as demonstrated by previous research(Hu et al., 2018; Liang et al., 2023). However, the spatio-temporal consistency of the GEOS-Chem simulation data with TROPOMI data is still satisfactory (i.e., similar increase–decrease patterns over time). The spatio-temporal correlation between GEOS-Chem simulation data and TROPOMI data can be thoroughly leveraged to derive continuous coverage data through data fusion (J. Chen et al., 2022; He et al., 2022; Wang et al., 2021).

We will assume that a spatio-temporal relationship function exists between XCO and XCH_4 for GEOS-Chem and TROPOMI, as shown below:

$$\mathrm{XT} = f(\mathrm{XG},\ \mathrm{Lat},\ \mathrm{Lon},\ \mathrm{Date}) \tag{1}$$

Where Lat, Lon, and Date denote latitude, longitude, and time series, respectively, and XG is the GEOS-Chem XCO and XCH_4 concentration value. That is, the TROPOMI value at a given spatial and temporal coordinate can be determined by taking the modeled concentration value and the spatial and temporal coordinates to which it pertains.

In order to simplify the resolution of this issue, we convert the aforementioned equation to the form of a scalar product of XG with a spacetime transformation matrix M, as follows:

$$\mathrm{XT} = XG * \rho \tag{2}$$

The parameter relationships corresponding to GEOS-Chem and TROPOMI are established at each spatio-temporal coordinate in ρ, a spatio-temporal three-dimensional matrix with the same scale as XT and XG. ρ is a smoothed three-dimensional parameter matrix, inspired by previous work on filling in the vacant values in the spatio-temporal data and enhancing the smoothing of the data

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

using the multidimensional discrete cosine transform (Elharar et al., 2007; Garcia, 2010, 2010; J. Robinson and V. Kecman, 2003; Okolie and Smit, 2022; Peng et al., 2005; Rao and Yip, 2014; Wang et al., 2023). The majority of the known parameters in ρ can be obtained from the valid values in GEOS-Chem and TROPOMI. We suggest a spatio-temporal 3D matrix smoothing algorithm that is based on Singular Value Decomposition (SVD) and Discrete Cosine Transform (DCT) to enhance the smoothness of the data and fill unoccupied values in spatio-temporal data. This algorithm is designed to suit our data. The method effectively manages spatio-temporal data with missing values while maintaining the spatio-temporal correlation of the data by combining spatio-temporal nearest-neighbor interpolation and regularized optimization techniques.

We find the spatio-temporal 3D matrix $\hat{\rho}$ that minimizes Eq. (3) by means of the 3D discrete cosine transform method, as a way to obtain the best estimate of the vacancy value, including the residual term on the left-hand side and the smoothing term on the right-hand side.

$$E(\hat{\rho}) = \left\| \omega^{\frac{1}{2}} * (\hat{\rho} - \rho) \right\|^2 + \varepsilon \|\nabla^2 \hat{\rho}\|^2 \quad (3)$$

where $\|\|$ denotes the Euclidean paradigm, ω is a binary mask indicating the availability of a parameter corresponding to the spatio-temporal location of ρ, ε denotes the smoothing parameter, and ∇^2 denotes the Laplace operator. This satisfied condition $\hat{\rho}$ can be solved by iteration of Eq. (4).

$$\hat{\rho} = \alpha \, \mathrm{IDCT}_3\left(\Gamma^3 * \mathrm{DCT}_3 (\omega * (\rho - \hat{\rho}) + \hat{\rho}) \right) + (1 - \alpha) \hat{\rho} \quad (4)$$

where α is a parametric factor for accelerating convergence, Γ^3 denotes the 3D spatio-temporal filtering matrix associated with the smoothing term, which can be obtained through Eq. (5), and DCT_3 and $IDCT_3$ denote the 3D discrete cosine signal transform and its inverse transform, respectively, with Eqs. (6) and (7) as their transformation rules.

$$\Gamma^3_{i_1,i_2,i_3} = \frac{1}{1+\varepsilon \sum_{k=1}^{3} 2\left[1-\cos\frac{(i_k-1)\pi}{n_k}\right]} \quad (5)$$

Here, i_k denotes the ith value along the kth dimension and n_k denotes the size of ρ along the kth dimension. This means that the value at each position of this three-dimensional spatio-temporal filtering matrix is completely determined by its position, and the closer its position is to the element of the matrix at position (1, 1, 1), the larger (the closer it is to 1) the value is, and vice versa the smaller it is. The value at position (1, 1, 1) is 1. Since the low frequencies of the discrete cosine transformed signal matrix are mainly located close to position (1, 1, 1), with this filtering matrix, it

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

is possible to search for ρ that is suitable for smoothing. In this study, the total number of iterations is empirically set to 100, α is set to 0.75, and ε takes values in the middle of the range from 10^3 to 10^{-1}.

$$F(u,v,w) = \frac{2}{\sqrt{NMP}} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \sum_{z=0}^{P-1} f(x,y,z) \cos\left(\frac{\pi(2x+1)u}{2N}\right) \cos\left(\frac{\pi(2y+1)v}{2M}\right) \cos \quad (6)$$

Three-dimensional discrete cosine signal transform (DCT_3) rule, where u=0, 1, 2, ..., N-1, v=0,1, 2, ..., M-1, w=0, 1, 2, ..., P -1. N, M, and P represent the magnitude of the signal in each of the three dimensions.

$$f(x,y,z) = \frac{2}{\sqrt{NMP}} \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} \sum_{w=0}^{P-1} F(u,v,w) \cos\left(\frac{\pi(2x+1)u}{2N}\right) \cos\left(\frac{\pi(2y+1)v}{2M}\right) \cos \quad (7)$$

Three-dimensional discrete cosine inverse transform ($IDCT_3$) rule, where x=0, 1, 2, ..., N-1, y=0, 1, 2, ..., M-1, and z=0, 1, 2, ..., P -1. N, M, and P represent the magnitude of the signal in each of the three dimensions.

Furthermore, the 3D discrete cosine transform is provided with a complete 3D matrix for signal conversion. Consequently, we interpolate ρ with missing values using the spatio-temporal autocorrelation property of ρ. Subsequently, we employ singular value decomposition (SVD) to retain 80% of the singular value energy, thereby enabling data downscaling and compression. The rules of the singular value decomposition (SVD) are illustrated in Eq. (8) as a method of preserving the primary ρ components for the iteration of Eq. (4).

$$A = U\Sigma V^T \quad (8)$$

where the matrices $A \in \mathbb{R}^{m\times n}$, $U \in \mathbb{R}^{m\times m}$ are orthogonal matrices whose column vectors are called left singular vectors; $\Sigma \in \mathbb{R}^{m\times n}$ is a diagonal matrix whose diagonal elements are the singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0$ (r is the rank of the matrix A). $V \in \mathbb{R}^{n\times n}$ is an orthogonal matrix whose column vectors are called right singular vectors.

Calculate the total energy of the singular values $E_{\text{total}} = \sum_{i=1}^{r} \sigma_i^2$, and find the smallest k such that the first k singular values account for at least 80% of the total energy, i.e., $(\sum_{i=1}^{k} \sigma_i^2)/E_{\text{total}} \geq 0.8$. Setting the last $r-k$ smaller singular values in Σ to zero yields the truncated diagonal matrix $\Sigma_k = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k, 0, \ldots, 0)$, and reconstructing the approximation matrix using the truncated singular value matrix $A_k = U\Sigma_k V^T$.

2.2.3. Deep Residual Refinement via Learning-Based Mask Reconstruction

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

To further enhance the reconstruction quality, we introduce a residual learning module based on a deep neural network to refine the fused product. To evaluate the effectiveness of the residual learning module, we conducted a site-based validation experiment across all TCCON stations, using data from the year 2021, to provide a rigorous comparison of the five reconstruction methods: (1) only DCT fusion(Garcia, 2010), (2) DCT and SVD mixed-signal reconstruction(M. Bengherabi et al., 2008; Majhi and Pal, 2021), (3) residual CNN(Y. Jiang et al., 2024), (4) residual U-Net(Ronneberger et al., 2015; Yan et al., 2022), and (5) residual XGBoost(Naseem et al., 2024). The performance was assessed using the coefficient of determination (R^2), root mean square error (RMSE), and mean bias (μ).

As shown in Tables S1 and S2, the residual U-Net consistently outperforms other methods for both XCO and XCH_4. Specifically, R^2 increased from 0.8227 (only DCT) to 0.845 for XCO and from 0.7056 to 0.7598 for XCH_4, while RMSE decreased correspondingly. These results demonstrate a improvement in the quality of the reconstructed data.

Residual Learning Objective

Let X_{DCT} denote the preliminary fused XCO/XCH_4 field obtained from the DCT/SVD-based reconstruction, and X_{TROPOMI} be the valid observational values from the satellite. The residual between the fusion estimate and the true value (only available at observed locations) is defined as:

$$\Delta X_{\mathrm{true}} = X_{\mathrm{TROPOMI}} - X_{\mathrm{DCT}}, \text{where } M = 1 \quad (9)$$

Here, $M \in \{0,1\}^{H\times W}$ is a binary mask indicating the presence (1) or absence (0) of valid satellite data. The goal is to train a model $\mathcal{F}_\theta(\cdot)$ parameterized by θ to predict the residual ΔX_{pred} across the entire domain:

$$\Delta X_{\mathrm{pred}} = \mathcal{F}_\theta(X_{\mathrm{G}}, X_{\mathrm{DCT}}, \mathbf{A}) \quad (10)$$

Where X_{G} denotes the GEOS-Chem full-coverage simulation data; X_{DCT} denotes the DCT/SVD-reconstructed preliminary fusion; **A** denotes the auxiliary information such as meteorological fields and emission inventories. Precursor meteorological data simulated by GEOS-Chem and emission inventory were used in this study. The final fused product is obtained by correcting the DCT/SVD estimate with the predicted residual:

$$X_{\mathrm{fused}} = X_{\mathrm{DCT}} + \Delta X_{\mathrm{pred}} \quad (11)$$

Loss Function Design

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

The model is trained using only the valid observations, i.e., locations where M=1. The loss function is designed to minimize the residual error at these observed locations:

$$\mathcal{L}(\theta) = \frac{1}{\sum M_{i,j}} \sum_{i,j} M_{i,j} \cdot \left(\Delta X_{\text{pred}}^{(i,j)} - \Delta X_{\text{true}}^{(i,j)}\right)^2 \quad (12)$$

This masked mean squared error ensures that the learning focuses on valid regions, while generalization to unobserved regions is achieved through the spatial context and auxiliary inputs.

Model Architecture and Training

In this study, we employ a lightweight residual U-Net architecture to predict the full-domain residual field ΔX_{pred}, which represents the correction from the DCT/SVD reconstruction to the expected TROPOMI observation. The network inputs include the GEOS-Chem simulation, the DCT/SVD reconstruction output, and auxiliary meteorological variables (Wang et al., 2025). The TROPOMI mask is not used as an input but is instead applied during the loss computation to focus learning only on valid pixels.

The residual U-Net consists of an encoder-decoder structure with skip connections and lightweight residual blocks to enhance gradient flow and convergence speed. The input and output spatial dimensions remain unchanged (i.e., 720 × 1440 at 0.25° resolution), enabling pixel-wise learning of spatial residuals. The training data comprises daily global maps from 2019 to 2023.

To ensure stable training and prevent overfitting, we implemented standard regularization strategies by incorporating early stopping, dropout layers (p = 0.3), and input normalization (mean = 0, std = 1). The final output is truncated to avoid physically implausible corrections, by enforcing:

$$|\Delta X_{\text{pred}}| < \gamma \cdot \sigma_{\text{TROPOMI}}, \gamma \in [2,3] \quad (13)$$

where σ_{TROPOMI} denotes the standard deviation of observed valid values, and γ is a hyperparameter controlling confidence bounds.

2.2.4. Evaluation Schem

Our study's evaluation methodology encompasses ground station validation and the assessment of geographical dispersion alongside relevant applications. The ground site validation entails the assessment of GEOS-Chem simulations, TROPOMI, and the fusion-generated XCO and XCH_4 data against TCCON measurements. This process encompasses the validation of overall results across all sites in comparison to individual site evaluations, utilizing the coefficient of determination (R^2), root mean square error (RMSE), mean bias (μ), and standard deviation of the bias (σ) (Karunasingha,

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

2022; Kobayashi and Salam, 2000; Wang et al., 2023). The aforementioned measures were computed using a significant threshold of $p < 0.01$ (Walsh et al., 2014). The assessment of geographical distribution includes the analysis of GEOS-Chem simulations, TROPOMI, and integrated XCO and XCH_4 data over several temporal scales, including multi-year averages, seasonal variations, and yearly metrics. Relevant applications include the evaluation of TROPOMI data acquisition, the integration of XCO and XCH_4 data at extremes, and the comparative analysis of multi-year column concentration increase rates across various geographies.

3. Results and discussion

3.1. TCCON Site Validation

The TCCON sites are uniformly spread worldwide, and the TROPOMI satellite traverses the region at around 13:30 local time; hence, we use the average column concentrations from the TCCON sites at 13:30±1h local time as the validation data for the ground stations. The GEOS-Chem simulation, TROPOMI, and the integrated data are then collected within a 2° radius surrounding each location as the corresponding validation outcomes. Only those cases in which the satellite data exhibit a missing rate exceeding 0.5 at the site are retained for comparison.

Figure 3 and 4 illustrate the time series of daily GEOS-Chem, TROPOMI, fusion data, and TCCON's XCO and XCH_4 data for several different ground stations, respectively. Figure 3 presents a time series plot of carbon monoxide column concentrations from three stations: easttroutlake01, lamont01, and parkfalls01. Figure 4 illustrates a daily time series plot of methane column concentrations, exemplified by edwards01, nicosia01, and pasadena01. In comparison to TCCON, both GEOS-Chem simulated XCO and XCH_4 data exhibit underestimation; nonetheless, they maintain time-varying consistency (i.e., similar increase–decrease patterns over time) with the TCCON site data, suggesting that the model simulation may serve as a reference for temporal and geographic fluctuations in the fused data. The temporal trend of TROPOMI and integrated XCO and XCH_4 data aligns with that of TCCON. The amalgamated data exhibit reduced biases μ (-0.79, -3.58, -1.08 ppb for XCO; -5.43, -2.49, -4.34 ppb for XCH_4) and diminished standard deviations σ (9.93, 4.26, 5.73 ppb for XCO; 5.99, 6.87, 6.77 ppb for XCH_4) in comparison to the TROPOMI data.

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

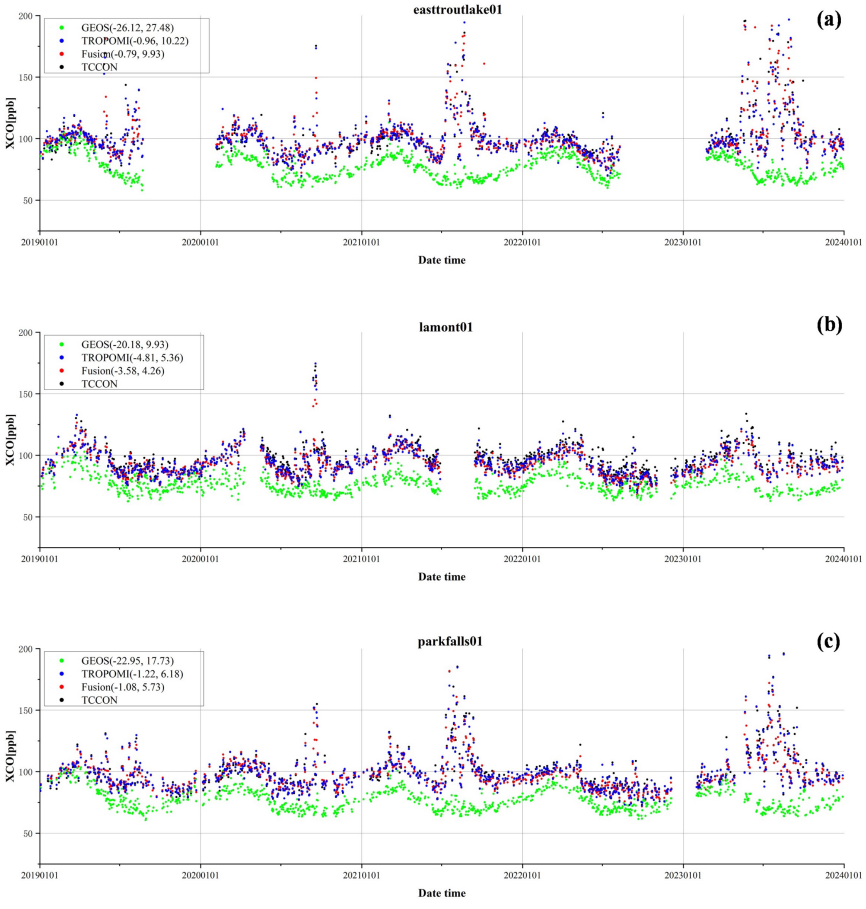

Fig. 3 (a)(b)(c) Time-series scatter plots of daily GEOS-Chem, TROPOMI, fusion data, and TCCON on XCO at three sites, easttroutlake01, lamont01, and parkfalls01, respectively. The first and second numbers in parentheses indicate the deviation (μ) and standard deviation (σ), respectively, both in parts per billion (ppb).

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

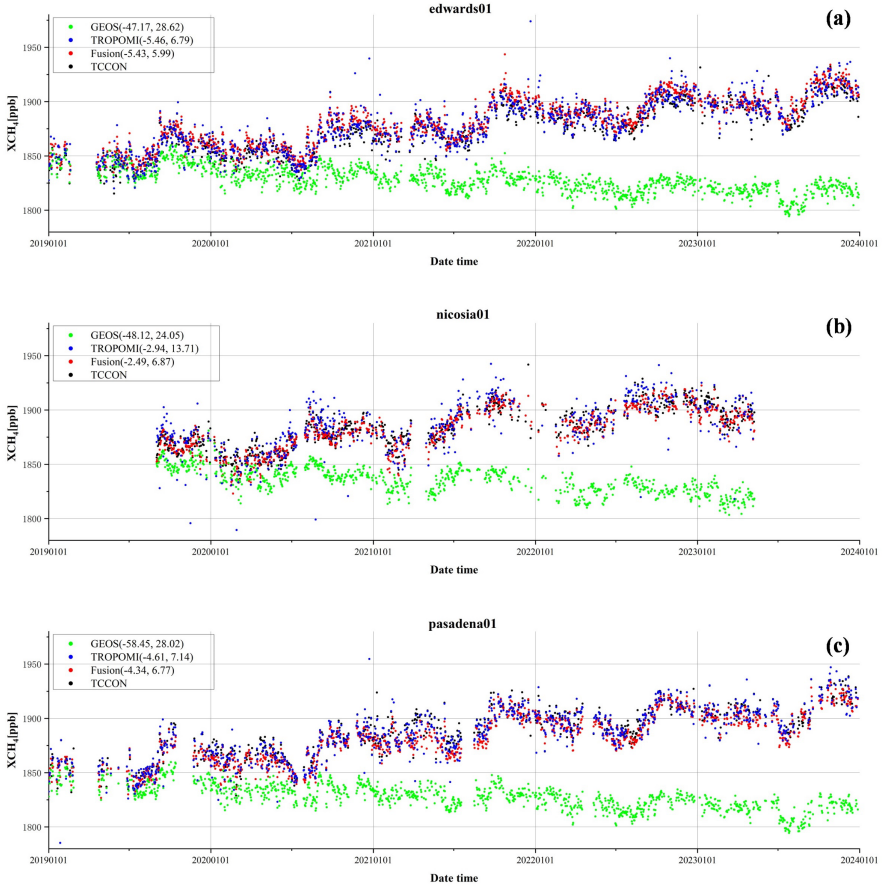

Fig. 4 (a)(b)(c) Scatterplots of daily time series of GEOS-Chem, TROPOMI, fusion data, and TCCON on XCH_4 for the three sites edwards01, nicosia01, and pasadena01, respectively. The first and second numbers in parentheses indicate the deviation (μ) and standard deviation (σ), respectively, both in parts per billion (ppb).

The systematic underestimation of GEOS-Chem model simulations may introduce bias in the correlation analysis when directly compared with TCCON observations. The average bias of the GEOS-Chem simulation data was corrected using a manual correction method, which more accurately reflects the relationship with TCCON observation data, thereby enhancing the reliability of the correlation analysis. This study further enhances scientific validity and relevance by screening TROPOMI satellite observation data samples with a missing rate exceeding 50% at the stations, and comparing these with TCCON data for validation. This screening method effectively addresses uncertainty arising from insufficient data coverage while concentrating on the incremental aspects

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

of the study, thereby enhancing the scientific basis for model optimization and satellite data validation.

Representative sites were selected based on data completeness to ensure statistical validity. Specifically, for each TCCON station, we computed the fraction of TROPOMI retrievals with a quality flag $\geqslant$ 0.5 within the 13:30 $\pm$ 1 h local-time window. Stations with the highest availability ratios were retained for analysis. Three sites, burgos01, izana01, and nicosia01, were utilized for XCO validation (refer to Figure 5). The findings indicate that the fused datasets at the burgos01 and izana01 sites demonstrate notable superiority compared to the GEOS-Chem simulation results and TROPOMI satellite observations, as evidenced by higher coefficients of determination (R^2), reduced root-mean-square error (RMSE), diminished bias (μ), and lower standard deviation (σ). This suggests that the fusion methods effectively enhance data accuracy and consistency. The fusion results on the nicosia01 site do not significantly exceed those of TROPOMI, yet they demonstrate comparable performance, thereby further validating the robustness of the fusion method. To validate XCH_4, we examine three sites: burgos01, rikubetsu01, and xianghe01 (refer to Figure 6). The analysis indicates that at the burgos01 and xianghe01 sites, the fused dataset markedly outperforms GEOS-Chem and shows modest but consistent improvements over TROPOMI across key evaluation metrics, including the coefficient of determination (R^2), root mean square error (RMSE), bias (μ), and standard deviation (σ). This finding suggests that the fusion method significantly enhances the quality of XCH_4 data. The fusion results on the rikubetsu01 site do not significantly exceed those of TROPOMI, yet they remain comparable, further illustrating the wide applicability and reliability of the fusion method. Table S3 and Table S4 present the validation results for each valid individual site for XCO and XCH_4, respectively. The fusion results demonstrate superior performance compared to the TROPOMI results at a minimum of 70% of the sites examined.

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

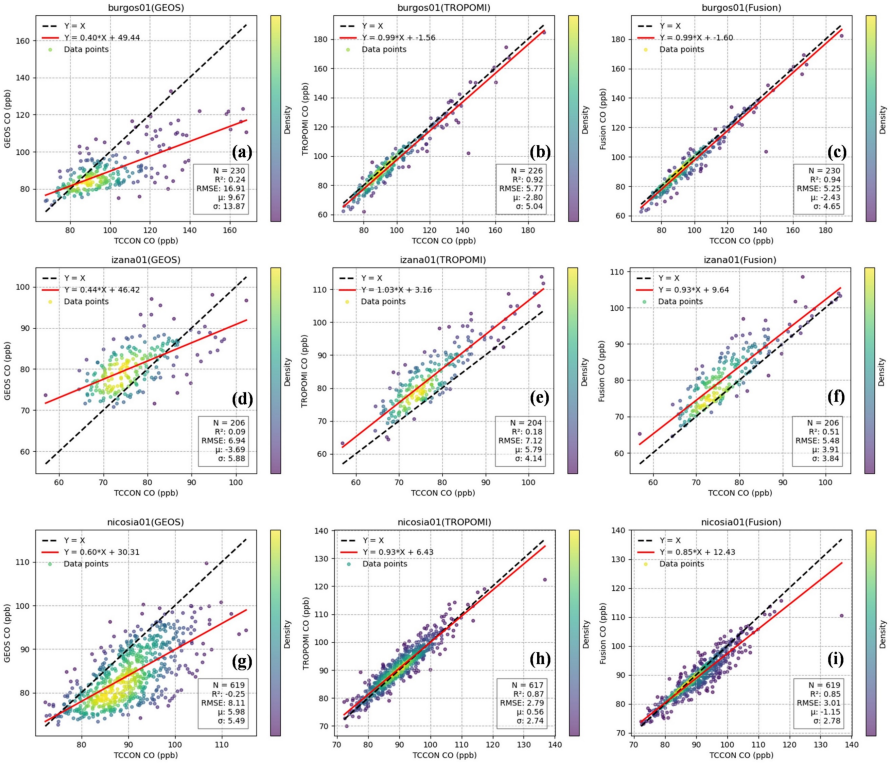

Fig. 5 (a, d, g) Scatterplots of the results of the independent validation of the (a, d, g) GEOS-Chem, (b, e, h) TROPOMI, and (c, f, i) fused XCO data at the burgos01, izana01, and nicosia01 sites, respectively. Black dashed lines and red realizations represent 1:1 and fitted lines, respectively. The x-axis is TCCON data, and the y-axis is GEOS-Chem, TROPOMI, and fusion data, all in parts per billion (ppb).

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

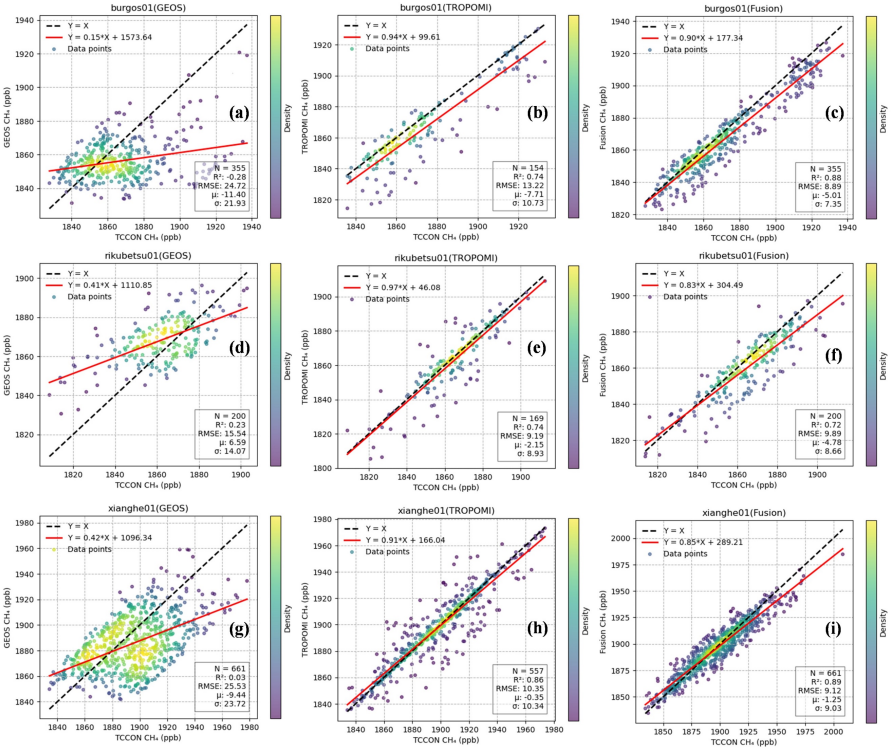

Fig. 6 (a, d, g) Scatterplots of the results of the independent validation of the (a, d, g) GEOS-Chem, (b, e, h) TROPOMI, and (c, f, i) fused XCH_4 data at the burgos01, rikubetsu01, and xianghe01 sites, respectively. Black dashed lines and red realizations represent 1:1 and fitted lines, respectively. The x-axis is TCCON data, and the y-axis is GEOS-Chem, TROPOMI, and fusion data, all in parts per billion (ppb).

Figure 7 illustrates the findings of the comprehensive correlation validation between XCO and XCH_4 across all verified locations. For XCO, the integrated dataset markedly surpasses the GEOS-Chem simulations for data quality and exceeds the TROPOMI satellite observations across numerous critical criteria. The fused data exhibit a coefficient of determination (R^2) of 0.92, markedly surpassing that of TROPOMI, thereby demonstrating superior capacity to encapsulate observational variability. The root mean square error (RMSE) is 4.85 ppb, and the standard deviation (σ) is 4.70 ppb, both of which outperform the corresponding metrics of TROPOMI, further substantiating the enhanced accuracy and consistency of the fused data. Moreover, the fused data exhibit a deviation (μ) index comparable to TROPOMI, indicating that systematic errors have been

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

Open Access Earth System Science Data Discussions

efficiently managed and the overall performance is more resilient. The fused dataset for XCH_4 demonstrates notable advantages, exhibiting a coefficient of determination (R^2) of 0.85, surpassing that of TROPOMI, thereby indicating superior interpretative capability; additionally, it presents a standard deviation (σ) of 12.59 ppb, which is more favorable than TROPOMI's, further substantiating the efficacy of the fused data in minimizing dispersion. While the fused data exhibits comparable performance to TROPOMI for root mean square error (RMSE) and bias (μ) parameters, its overall efficacy reflects a significant level of stability and consistency, particularly in complex contexts. This illustrates the favorable outcomes attained by the aforementioned fusion approach.

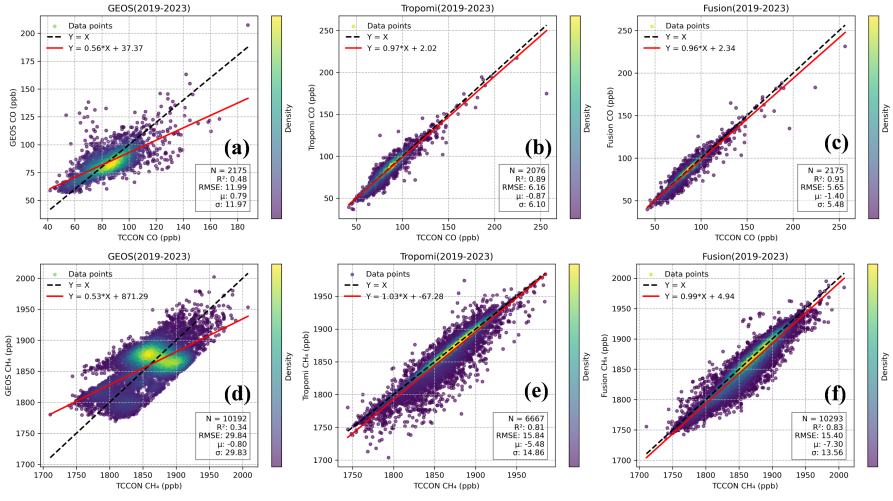

Fig. 7 Scatterplots of the results of the overall validation of (a, d) GEOS-Chem, (b, e) TROPOMI, and (c, f) fused XCO vs. XCH_4 data at all respective validated sites. Black dashed lines and red realizations represent 1:1 and fitted lines, respectively. The x-axis is TCCON data, and the y-axis is GEOS-Chem, TROPOMI, and fusion data, all in parts per billion (ppb).

3.2. Multi-Scale Spatio-Temporal Analysis

Figure S1 illustrates the comparison of worldwide yearly GEOS-Chem, TROPOMI, fused XCO, and XCH_4 for the years 2020 and 2022. The fused data exhibit a similar geographical distribution with TROPOMI. Despite a substantial underestimation of GEOS-Chem, it retains a

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

robust geographic distribution alignment with TROPOMI, providing a critical reference for data fusion, and this underestimation is markedly altered post-fusion.

Figure S2 illustrates the global distribution of fused XCO and XCH_4 for three representative days in 2020 and 2022, respectively. The fusion findings, as seen in the image, offer comprehensive information on atmospheric CO and CH_4, distinctly revealing their worldwide geographical distribution. For comparison, Figure S3 illustrates the global distribution of XCO and XCH_4 observed by TROPOMI on these corresponding days. Meteorological factors have resulted in several gaps in the satellite observations, particularly evident in the XCH_4 data, when compared to Figure S2. Figure S2 illustrates that the integrated data addresses the deficiencies in geographical and temporal information, so improving data continuity while preserving the integrity of the satellite observations.

Figure 8 illustrates the global multi-year average distributions of fused XCO and XCH_4 for the period 2019–2023. Elevated concentrations of both gases are predominantly observed across Asia, particularly over China and India. For XCO, distinct high-value regions are also evident in Central Africa and northern South America. Figures S4 and S5 illustrate the seasonal averages of the fused global XCO and XCH_4 data from 2019 to 2023. The seasonal variations in the geographical distribution are distinctly captured by the fusion results. Notably, XCO exhibits more pronounced spatiotemporal variability compared to XCH_4. Specifically, CO tends to be spatially concentrated in certain regions, whereas CH_4 displays a relatively more uniform global distribution.

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

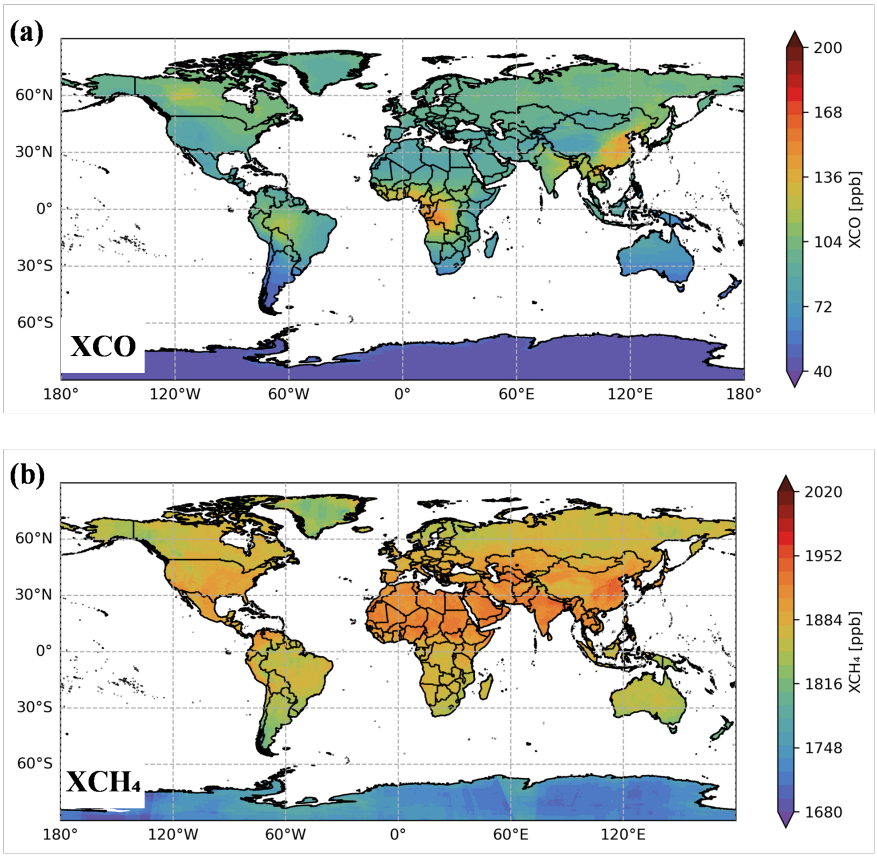

Fig. 8 (a) Global concentration distribution of fused XCO data averaged over multiple years from 2019 to 2023; (b) Global concentration distribution of fused XCH_4 data averaged over multiple years from 2019 to 2023. The color bars indicate the concentrations of XCO and XCH_4 in parts per billion (ppb).

Figures S6 and S7 illustrate the annual trends of global XCO and XCH_4 from 2019 to 2023, alongside their respective regional trajectories. Figure S6 indicates that XCO levels in North America showed a consistent increase from 2019 to 2023, while Central Africa and Eastern China exhibited a declining trend. Global XCH_4 levels demonstrated a steady increase over this period, with a notably higher growth rate observed in Central Africa (Figure S7).

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

Open Access Earth System Science Data Discussions

3.3. Local High-Resolution Data Analysis

To enhance the refinement capability of our fused data, we employed GEOS-Chem to simulate 0.25x0.3125 nested gridded data for the Chinese region. Utilizing the same methodology, we refined the fused data for this region, achieving a grid accuracy of 0.05°, which served as the basis for our analysis of the local area in China.

Figure S8 and S9 illustrate the fluctuations of XCO and XCH_4 in the TROPOMI and fusion datasets, respectively, during the late August 2022 hill fire in Chongqing. This encompasses a comparative analysis of the TROPOMI and fusion data over three days during the fire, along with a histogram depicting the daily variations in mean XCO and XCH_4 concentrations within the fire-affected area during the mid to late months of the event. The Figure demonstrate that the fused data preserves local observational details from the satellite, thereby enhancing the comprehensiveness of the satellite data to a certain degree. The histograms reveal that the fused data reflects an upward trend in XCO and XCH_4 concentrations during the hill fires, with average increases of 17.1 ppb and 24.5 ppb, respectively, during the hill fire period compared to the non-hill fire period at midday.

Figure 9 and 10 depict the fluctuations in XCO and XCH_4 recorded by TROPOMI and integrated data in rice cultivation areas of Northeast and Southeast China, respectively. XCH_4 exhibits a steady increase in both satellite and fused data across these areas. In contrast to the fused data, the yearly growth rate obtained from satellite XCH_4 demonstrates an overestimation in both regions, with discrepancies of 17% and 26% for the Southeast and Northeast, respectively.

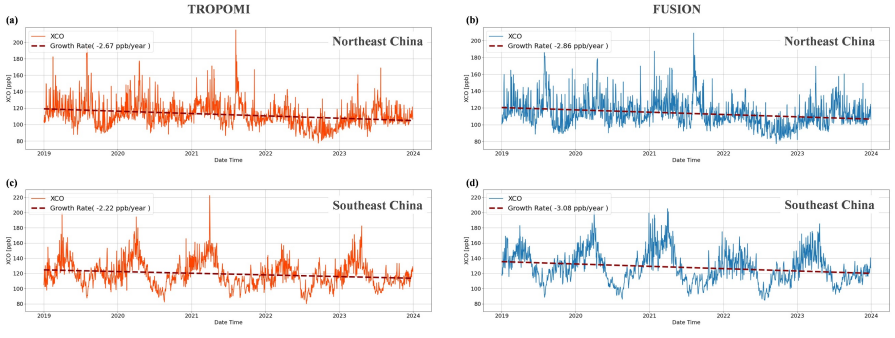

Fig. 9 (a, c) Line plots of daily changes in parts per billion (ppb) for TROPOMI and (b, d) fused XCO data over rice-growing regions of northeastern China and southeastern China, respectively; the red dashed line is the fitted line of annual growth rate, and the annual growth rate is shown in

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

parentheses, in parts per billion per year (ppb/year).

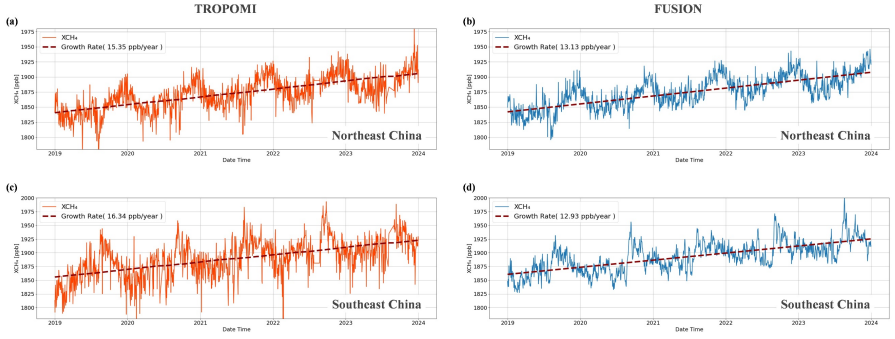

Fig. 10 (a, c) Line plots of daily changes in parts per billion (ppb) for TROPOMI and (b, d) fused XCH_4 data over rice-growing regions in northeastern China versus southeastern China, respectively; red dashed line is the fitted line of the annual growth rate, and in parentheses is the annual growth rate in parts per billion per year (ppb/year).

XCO exhibits a downward trajectory in both satellite and fused data across rice cultivation areas of Northeast and Southeast China. Nonetheless, the annual decline rate of satellite XCO is underestimated relative to the fused data, with discrepancies of 7% and 38% for the Southeast and Northeast regions, respectively. This variance results from absent values in the satellite data, which omit specific low-concentration areas from the computations, thereby inflating the estimation of concentration changes. In contrast, the fused data somewhat alleviate the intensity of these fluctuations.

Rice paddies, as a major source of methane emissions, exhibit a robust association between their increasing cultivation area and the persistent growth in methane emissions. The expansion of rice farming has resulted in a proportional increase in methane emissions, underscoring the considerable influence of agricultural practices on greenhouse gas output(Shen et al., 2024). Straw burning is a key source of carbon monoxide, and the enforcement of pertinent control laws has markedly reduced CO emissions, leading to a steady decline in carbon monoxide levels in both regions(Huang et al., 2021).

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

4. Conclusions

This study introduces a two-stage spatio-temporal fusion approach that integrates GEOS-Chem model outputs with TROPOMI satellite data. The method initially performs a signal-domain reconstruction using three-dimensional discrete cosine transform (DCT) and singular value decomposition (SVD), followed by a residual refinement step based on a lightweight convolutional neural network, enabling accurate gap-filling and spatial smoothing. This method successfully produces global daily continuous coverage of XCO and XCH_4 products at a resolution of 0.25° for the years 2019-2023, achieving high-resolution data fusion at 0.05°. The findings indicate that the fused data surpass the GEOS-Chem simulation results in accuracy and consistency and are comparable to or exceed the TROPOMI satellite observations across numerous critical criteria. The integrated data provide notable improvements in the coefficient of determination (R^2), root mean square error (RMSE), bias (μ), and standard deviation (σ) globally, particularly in addressing missing areas and enhancing smoothness. This study effectively addressed the systematic underestimation issue of the GEOS-Chem model, attributed to inaccuracies in emission inventories, through data fusion, enabling the fused data to more accurately represent the long-term spatial and temporal distributions and seasonal variations of global XCO and XCH_4.

The fused data offer a definitive overview of the geographical distribution patterns of XCO and XCH_4, together with their temporal trends. In regions like Asia, Africa, and North America, the fusion data effectively captured pronounced spatiotemporal variations and regional patterns. The fusion data demonstrates a robust capacity to detect extreme events (e.g., wildfires) and accurately represent short-term significant fluctuations in XCO and XCH_4 concentrations, serving as an excellent instrument for monitoring severe weather and pollution occurrences. This study elucidates the atmospheric distribution and alterations in specific locations of China using high-resolution data fusion. Particularly in the examination of methane emissions from the mountain fire incident and rice cultivation region in Chongqing, the integrated data exhibit elevated spatial resolution and temporal continuity, thereby offering significant data support for regional air quality monitoring and climate change research. All created fusion data may be obtained upon request from the authors for academics and policymakers.

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

This study has yielded significant outcomes in data fusion and product production; nonetheless, several deficiencies require more enhancement and optimization in further research. The GEOS-Chem model exhibits systematic bias in simulating XCO and XCH_4. While the data fusion approach has largely mitigated this issue, additional optimization of the emission inventories and chemical reaction processes is necessary to enhance simulation accuracy. This work primarily utilizes TROPOMI and GEOS-Chem data; however, the incorporation of additional satellite observation data (e.g., GOSAT, OCO-2) and ground observation data in the future might enhance the diversity and precision of data fusion.

This work effectively produced high-resolution XCO and XCH_4 products for global and Chinese areas by employing multi-source data fusion and spatio-temporal distribution analysis methodologies, therefore offering significant data assistance for climate change and air quality monitoring. In the future, as model accuracy improves, data fusion methods are optimized, and multi-source data is introduced, the methodologies and findings of this study are anticipated to have a more significant impact across a broader array of application situations. Data Availability Statement The global (0.25°) and China-specific (0.05°) daily gap-free XCO and XCH_4 datasets (2019–2023) generated in this study are openly available in the Zenodo repository at https://doi.org/10.5281/zenodo.17936461.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The primary dataset generated in this work is available at Zenodo via https://doi.org/10.5281/zenodo.17936461(Li et al., 2025).

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

Acknowledgments

This research was jointly funded by the National Key Research and Development Program of China (2022YFC3700103) and the National Natural Science Foundation of China (42205151).

References

Borsdorff, T., aan de Brugh, J., Schneider, A., Lorente, A., Birk, M., Wagner, G., Kivi, R., Hase, F., Feist, D.G., Sussmann, R., 2019. Improving the TROPOMI CO data product: update of the spectroscopic database and destriping of single orbits. Atmospheric Measurement Techniques 12, 5443–5455.

Buschmann, M., Deutscher, N.M., Sherlock, V., Palm, M., Warneke, T., Notholt, J., 2016. Retrieval of xCO 2 from ground-based mid-infrared (NDACC) solar absorption spectra and comparison to TCCON. Atmospheric Measurement Techniques 9, 577–585.

Chai, X., Tonjes, D.J., Mahajan, D., 2016. Methane emissions as energy reservoir: Context, scope, causes and mitigation strategies. Progress in Energy and Combustion Science 56, 33–70. https://doi.org/10.1016/j.pecs.2016.05.001

Chen, J., Shen, H., Li, X., Li, T., Wei, Y., 2022. Ground-level ozone estimation based on geo-intelligent machine learning by fusing in-situ observations, remote sensing data, and model simulation data. International Journal of Applied Earth Observation and Geoinformation 112, 102955. https://doi.org/10.1016/j.jag.2022.102955

Chen, X., Huang, Y., Nie, C., Zhang, S., Wang, G., Chen, S., Chen, Z., 2022. A long-term reconstructed TROPOMI solar-induced fluorescence dataset using machine learning algorithms. Scientific Data 9, 427. https://doi.org/10.1038/s41597-022-01520-1

Cogan, A.J., Boesch, H., Parker, R.J., Feng, L., Palmer, P.I., Blavier, J.-F., Deutscher, N.M., Macatangay, R., Notholt, J., Roehl, C., 2012. Atmospheric carbon dioxide retrieved from the Greenhouse gases Observing SATellite (GOSAT): Comparison with ground-based TCCON observations and GEOS-Chem model calculations. Journal of Geophysical Research: Atmospheres 117.

Deeter, M.N., Emmons, L.K., Francis, G.L., Edwards, D.P., Gille, J.C., Warner, J.X., Khattatov, B., Ziskin, D., Lamarque, J. -F., Ho, S. -P., Yudin, V., Attié, J. -L., Packman, D., Chen, J., Mao, D., Drummond, J.R., 2003. Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument. J. Geophys. Res. 108, 2002JD003186. https://doi.org/10.1029/2002JD003186

Elharar, E., Stern, A., Hadar, O., Javidi, B., 2007. A Hybrid Compression Method for Integral Images Using Discrete Wavelet Transform and Discrete Cosine Transform. J. Display Technol. 3, 321–325.

Filonchyk, M., Peterson, M.P., Zhang, L., Hurynovich, V., He, Y., 2024. Greenhouse gases emissions and global climate change: Examining the influence of CO2, CH4, and N2O. Science of The Total Environment 935, 173359. https://doi.org/10.1016/j.scitotenv.2024.173359

Fritz, T.M., Eastham, S.D., Emmons, L.K., Lin, H., Lundgren, E.W., Goldhaber, S., Barrett, S.R.H., Jacob, D.J., 2022. Implementation and evaluation of the GEOS-Chem chemistry module

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

version 13.1.2 within the Community Earth System Model v2.1. Geosci. Model Dev. 15, 8669–8704. https://doi.org/10.5194/gmd-15-8669-2022

G. Chander, T. J. Hewison, N. Fox, X. Wu, X. Xiong, W. J. Blackwell, 2013. Overview of Intercalibration of Satellite Instruments. IEEE Transactions on Geoscience and Remote Sensing 51, 1056–1080. https://doi.org/10.1109/TGRS.2012.2228654

Garcia, D., 2010. Robust smoothing of gridded data in one and higher dimensions with missing values. Computational Statistics & Data Analysis 54, 1167–1178. https://doi.org/10.1016/j.csda.2009.09.020

Goudar, M., Anema, J.C.S., Kumar, R., Borsdorff, T., Landgraf, J., 2023. Plume detection and emission estimate for biomass burning plumes from TROPOMI carbon monoxide observations using APE v1.1. Geosci. Model Dev. 16, 4835–4852. https://doi.org/10.5194/gmd-16-4835-2023

Griffin, D., Chen, J., Anderson, K., Makar, P., McLinden, C.A., Dammers, E., Fogal, A., 2024. Biomass burning CO emissions: exploring insights through TROPOMI-derived emissions and emission coefficients. Atmos. Chem. Phys. 24, 10159–10186. https://doi.org/10.5194/acp-24-10159-2024

Guanter, L., Aben, I., Tol, P., Krijger, J.M., Hollstein, A., Köhler, P., Damm, A., Joiner, J., Frankenberg, C., Landgraf, J., 2015. Potential of the TROPOspheric Monitoring Instrument (TROPOMI) onboard the Sentinel-5 Precursor for the monitoring of terrestrial chlorophyll fluorescence. Atmos. Meas. Tech. 8, 1337–1352. https://doi.org/10.5194/amt-8-1337-2015

He, C., Ji, M., Grieneisen, M.L., Zhan, Y., 2022. A review of datasets and methods for deriving spatiotemporal distributions of atmospheric CO2. Journal of Environmental Management 322, 116101. https://doi.org/10.1016/j.jenvman.2022.116101

Heald, C.L., Jacob, D.J., Fiore, A.M., Emmons, L.K., Gille, J.C., Deeter, M.N., Warner, J., Edwards, D.P., Crawford, J.H., Hamlin, A.J., Sachse, G.W., Browell, E.V., Avery, M.A., Vay, S.A., Westberg, D.J., Blake, D.R., Singh, H.B., Sandholm, S.T., Talbot, R.W., Fuelberg, H.E., 2003. Asian outflow and trans-Pacific transport of carbon monoxide and ozone pollution: An integrated satellite, aircraft, and model perspective. J. Geophys. Res. 108, 2003JD003507. https://doi.org/10.1029/2003JD003507

Heilig, G.K., 1994. The greenhouse gas methane (CH4): Sources and sinks, the impact of population growth, possible interventions. Population and Environment 16, 109–137. https://doi.org/10.1007/BF02208779

Hoesly, R.M., Smith, S.J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J.J., Vu, L., Andres, R.J., Bolt, R.M., Bond, T.C., Dawidowski, L., Kholod, N., Kurokawa, J.-I., Li, M., Liu, L., Lu, Z., Moura, M.C.P., O'Rourke, P.R., Zhang, Q., 2018. Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS). Geosci. Model Dev. 11, 369–408. https://doi.org/10.5194/gmd-11-369-2018

Hu, J., Jia, J., Ma, Y., Liu, L., Yu, H., 2022. A Reconstructed Global Daily Seamless SIF Product at 0.05 Degree Resolution Based on TROPOMI, MODIS and ERA5 Data. Remote Sensing 14, 1504. https://doi.org/10.3390/rs14061504

Hu, K., Liu, Z., Shao, P., Ma, K., Xu, Y., Wang, S., Wang, Y., Wang, H., Di, L., Xia, M., 2024. A review of satellite-based CO2 data reconstruction studies: Methodologies, challenges, and advances. Remote Sensing 16.

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

685 Hu, L., Keller, C.A., Long, M.S., Sherwen, T., Auer, B., Da Silva, A., Nielsen, J.E., Pawson, S.,
686 Thompson, M.A., Trayanov, A.L., Travis, K.R., Grange, S.K., Evans, M.J., Jacob, D.J.,
687 2018. Global simulation of tropospheric chemistry at 12.5 km resolution: performance and
688 evaluation of the GEOS-Chem chemical module (v10-1) within the NASA GEOS Earth
689 system model (GEOS-5 ESM). Geosci. Model Dev. 11, 4603–4620.
690 https://doi.org/10.5194/gmd-11-4603-2018
691 Huang, L., Zhu, Y., Wang, Q., Zhu, A., Liu, Z., Wang, Y., Allen, D.T., Li, L., 2021. Assessment of
692 the effects of straw burning bans in China: Emissions, air quality, and health impacts.
693 Science of The Total Environment 789, 147935.
694 https://doi.org/10.1016/j.scitotenv.2021.147935
695 Imasu, R., Matsunaga, T., Nakajima, M., Yoshida, Y., Shiomi, K., Morino, I., Saitoh, N., Niwa, Y.,
696 Someya, Y., Oishi, Y., Hashimoto, M., Noda, H., Hikosaka, K., Uchino, O., Maksyutov, S.,
697 Takagi, H., Ishida, H., Nakajima, T.Y., Nakajima, T., Shi, C., 2023. Greenhouse gases
698 Observing SATellite 2 (GOSAT-2): mission overview. Progress in Earth and Planetary
699 Science 10, 33. https://doi.org/10.1186/s40645-023-00562-2
700 Inness, A., Aben, I., Ades, M., Borsdorff, T., Flemming, J., Jones, L., Landgraf, J., Langerock, B.,
701 Nedelec, P., Parrington, M., Ribas, R., 2022. Assimilation of S5P/TROPOMI carbon
702 monoxide data with the global CAMS near-real-time system. Atmos. Chem. Phys. 22,
703 14355–14376. https://doi.org/10.5194/acp-22-14355-2022
704 J. Robinson, V. Kecman, 2003. Combining support vector machine learning with the discrete cosine
705 transform in image compression. IEEE Transactions on Neural Networks 14, 950–958.
706 https://doi.org/10.1109/TNN.2003.813842
707 Jackson, R.B., Saunois, M., Bousquet, P., Canadell, J.G., Poulter, B., Stavert, A.R., Bergamaschi, P.,
708 Niwa, Y., Segers, A., Tsuruta, A., 2020. Increasing anthropogenic methane emissions arise
709 equally from agricultural and fossil fuel sources. Environmental Research Letters 15,
710 071002. https://doi.org/10.1088/1748-9326/ab9ed2
711 Karunasingha, D.S.K., 2022. Root mean square error or mean absolute error? Use their ratio as well.
712 Information Sciences 585, 609–629. https://doi.org/10.1016/j.ins.2021.11.036
713 Kawka, M., Struzewska, J., Kaminski, J.W., 2021. Spatial and Temporal Variation of NO2 Vertical
714 Column Densities (VCDs) over Poland: Comparison of the Sentinel-5P TROPOMI
715 Observations and the GEM-AQ Model Simulations. Atmosphere 12.
716 https://doi.org/10.3390/atmos12070896
717 Kiel, M., Wunch, D., Wennberg, P.O., Toon, G.C., Hase, F., Blumenstock, T., 2016. Improved
718 retrieval of gas abundances from near-infrared solar FTIR spectra measured at the
719 Karlsruhe TCCON station. Atmospheric Measurement Techniques 9, 669–682.
720 Kobayashi, K., Salam, M.U., 2000. Comparing Simulated and Measured Values Using Mean
721 Squared Deviation and its Components. Agronomy Journal 92, 345–352.
722 https://doi.org/10.2134/agronj2000.922345x
723 Laughner, J.L., Roche, S., Kiel, M., Toon, G.C., Wunch, D., Baier, B.C., Biraud, S., Chen, H., Kivi,
724 R., Laemmel, T., McKain, K., Quéhé, P.-Y., Rousogenous, C., Stephens, B.B., Walker, K.,
725 Wennberg, P.O., 2023. A new algorithm to generate a priori trace gas profiles for the
726 GGG2020 retrieval algorithm. Atmos. Meas. Tech. 16, 1121–1146.
727 https://doi.org/10.5194/amt-16-1121-2023
728 Lelieveld, J., Gromov, S., Pozzer, A., Taraborrelli, D., 2016. Global tropospheric hydroxyl

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

distribution, budget and reactivity. Atmos. Chem. Phys. 16, 12477–12493. https://doi.org/10.5194/acp-16-12477-2016

Li, M., Zhang, Q., Kurokawa, J.-I., Woo, J.-H., He, K., Lu, Z., Ohara, T., Song, Y., Streets, D.G., Carmichael, G.R., Cheng, Y., Hong, C., Huo, H., Jiang, X., Kang, S., Liu, F., Su, H., Zheng, B., 2017. MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP. Atmos. Chem. Phys. 17, 935–963. https://doi.org/10.5194/acp-17-935-2017

Li, Z., Tian, Y., Lin, P., Chang, B., Xue, J., 2025. Signal-Domain Guided Deep Learning for Gap-Filling of XCO and XCH_4: A Masked Spatio-Temporal Fusion of TROPOMI and GEOS-Chem (2019–2023). https://doi.org/10.5281/zenodo.17936462

Liang, R., Zhang, Y., Chen, W., Zhang, P., Liu, J., Chen, C., Mao, H., Shen, G., Qu, Z., Chen, Z., Zhou, M., Wang, P., Parker, R.J., Boesch, H., Lorente, A., Maasakkers, J.D., Aben, I., 2023. East Asian methane emissions inferred from high-resolution inversions of GOSAT and TROPOMI observations: a comparative and evaluative analysis. Atmos. Chem. Phys. 23, 8039–8057. https://doi.org/10.5194/acp-23-8039-2023

Lin, P., Tian, Y., Borsdorff, T., Landgraf, J., Li, Z., Wu, H., Xue, J., Ding, D., Ye, H., Zhu, Y., Liu, C., 2024. Tropomi Unravels Transboundary Transport Pathways of Atmospheric Carbon Monoxide in Tibet, China. https://doi.org/10.2139/ssrn.4756008

Loew, A., Bell, W., Brocca, L., Bulgin, C.E., Burdanowitz, J., Calbet, X., Donner, R.V., Ghent, D., Gruber, A., Kaminski, T., Kinzel, J., Klepp, C., Lambert, J., Schaepman-Strub, G., Schröder, M., Verhoelst, T., 2017. Validation practices for satellite-based Earth observation data across communities. Reviews of Geophysics 55, 779–817. https://doi.org/10.1002/2017rg000562

M. Bengherabi, L. Mezai, F. Harizi, A. Guessoum, M. Cheriet, 2008. Score Fusion of SVD and DCT-RLDA for Face Recognition, in: 2008 First Workshops on Image Processing Theory, Tools and Applications. Presented at the 2008 First Workshops on Image Processing Theory, Tools and Applications, pp. 1–8. https://doi.org/10.1109/IPTA.2008.4743776

Majhi, M., Pal, A.K., 2021. An image retrieval scheme based on block level hybrid dct-svd fused features. Multimedia Tools and Applications 80, 7271–7312. https://doi.org/10.1007/s11042-020-10005-5

Marais, E.A., Wiedinmyer, C., 2016. Air Quality Impact of Diffuse and Inefficient Combustion Emissions in Africa (DICE-Africa). Environ. Sci. Technol. 50, 10739–10745. https://doi.org/10.1021/acs.est.6b02602

McMillan, W.W., Barnet, C., Strow, L., Chahine, M.T., McCourt, M.L., Warner, J.X., Novelli, P.C., Korontzi, S., Maddy, E.S., Datta, S., 2005. Daily global maps of carbon monoxide from NASA's Atmospheric Infrared Sounder. Geophysical Research Letters 32, 2004GL021821. https://doi.org/10.1029/2004GL021821

Naseem, S., Mahmood, T., Khan, A.R., Farooq, U., Nawazish, S., Alamri, F.S., Saba, T., 2024. Image Fusion Using Wavelet Transformation and XGboost Algorithm. Computers, Materials & Continua 79.

Okolie, C.J., Smit, J.L., 2022. A systematic review and meta-analysis of Digital elevation model (DEM) fusion: pre-processing, methods and applications. ISPRS Journal of Photogrammetry and Remote Sensing 188, 1–29. https://doi.org/10.1016/j.isprsjprs.2022.03.016

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

Peng, J., Kim, C.-S., Jay Kuo, C.-C., 2005. Technologies for 3D mesh compression: A survey. Journal of Visual Communication and Image Representation 16, 688–733. https://doi.org/10.1016/j.jvcir.2005.03.001

Rao, K.R., Yip, P., 2014. Discrete cosine transform: algorithms, advantages, applications. Academic press.

Rinsland, C.P., Luo, M., Logan, J.A., Beer, R., Worden, H., Kulawik, S.S., Rider, D., Osterman, G., Gunson, M., Eldering, A., Goldman, A., Shephard, M., Clough, S.A., Rodgers, C., Lampel, M., Chiou, L., 2006. Nadir measurements of carbon monoxide distributions by the Tropospheric Emission Spectrometer instrument onboard the Aura Spacecraft: Overview of analysis approach and examples of initial results. Geophysical Research Letters 33, 2006GL027000. https://doi.org/10.1029/2006GL027000

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, Cham, pp. 234–241.

Schneising, O., Buchwitz, M., Hachmeister, J., Vanselow, S., Reuter, M., Buschmann, M., Bovensmann, H., Burrows, J.P., 2023. Advances in retrieving XCH_4 and XCO from Sentinel-5 Precursor: improvements in the scientific TROPOMI/WFMD algorithm. Atmos. Meas. Tech. 16, 669–694. https://doi.org/10.5194/amt-16-669-2023

Schneising, O., Buchwitz, M., Reuter, M., Bovensmann, H., Burrows, J.P., Borsdorff, T., Deutscher, N.M., Feist, D.G., Griffith, D.W.T., Hase, F., Hermans, C., Iraci, L.T., Kivi, R., Landgraf, J., Morino, I., Notholt, J., Petri, C., Pollard, D.F., Roche, S., Shiomi, K., Strong, K., Sussmann, R., Velazco, V.A., Warneke, T., Wunch, D., 2019. A scientific algorithm to simultaneously retrieve carbon monoxide and methane from TROPOMI onboard Sentinel-5 Precursor. Atmos. Meas. Tech. 12, 6771–6802. https://doi.org/10.5194/amt-12-6771-2019

Setianto, A., Triandini, T., 2013. Comparison of kriging and inverse distance weighted (IDW) interpolation methods in lineament extraction and analysis. Journal of Applied Geology 5.

Sha, M.K., De Mazière, M., Notholt, J., Blumenstock, T., Chen, H., Dehn, A., Griffith, D.W., Hase, F., Heikkinen, P., Hermans, C., 2019. Intercomparison of low and high resolution infrared spectrometers for ground-based solar remote sensing measurements of total column concentrations of CO 2, CH 4 and CO. Atmospheric Measurement Techniques Discussions 2019, 1–67.

Shen, N., Tan, J., Wang, W., Xue, W., Wang, Y., Huang, L., Yan, G., Song, Y., Li, L., 2024. Long-term changes of methane emissions from rice cultivation during 2000 – 2060 in China: Trends, driving factors, predictions and policy implications. Environment International 191, 108958. https://doi.org/10.1016/j.envint.2024.108958

Sicard, P., Crippa, P., De Marco, A., Castruccio, S., Giani, P., Cuesta, J., Paoletti, E., Feng, Z., Anav, A., 2021. High spatial resolution WRF-Chem model over Asia: Physics and chemistry evaluation. Atmospheric Environment 244, 118004. https://doi.org/10.1016/j.atmosenv.2020.118004

Spivakovsky, C.M., Logan, J.A., Montzka, S.A., Balkanski, Y.J., Foreman-Fowler, M., Jones, D.B.A., Horowitz, L.W., Fusco, A.C., Brenninkmeijer, C.A.M., Prather, M.J., Wofsy, S.C., McElroy, M.B., 2000. Three-dimensional climatological distribution of tropospheric OH:

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

Update and evaluation. J. Geophys. Res. 105, 8931–8980. https://doi.org/10.1029/1999JD901006

Tang, Z., n.d. Gap-filling using machine learning: implementations and applications in remote sensing.

Turquety, S., Hadji-Lazaro, J., Clerbaux, C., Hauglustaine, D.A., Clough, S.A., Cassé, V., Schlüssel, P., Mégie, G., 2004. Operational trace gas retrieval algorithm for the Infrared Atmospheric Sounding Interferometer. J. Geophys. Res. 109, 2004JD004821. https://doi.org/10.1029/2004JD004821

Valerio, A., Chen, Y.-C., Liu, C.-Y., Chen, Y.-Y., Lin, C.-Y., 2025. A Hybrid Regression–Kriging–Machine Learning Framework for Imputing Missing TROPOMI NO2 Data over Taiwan. Remote Sensing 17. https://doi.org/10.3390/rs17122084

Wall, M.E., Rechtsteiner, A., Rocha, L.M., 2003. Singular Value Decomposition and Principal Component Analysis, in: Berrar, D.P., Dubitzky, W., Granzow, M. (Eds.), A Practical Approach to Microarray Data Analysis. Springer US, Boston, MA, pp. 91–109. https://doi.org/10.1007/0-306-47815-3_5

Walsh, M., Srinathan, S.K., McAuley, D.F., Mrkobrada, M., Levine, O., Ribic, C., Molnar, A.O., Dattani, N.D., Burke, A., Guyatt, G., Thabane, L., Walter, S.D., Pogue, J., Devereaux, P.J., 2014. The statistical significance of randomized controlled trial results is frequently fragile: a case for a Fragility Index. Journal of Clinical Epidemiology 67, 622–628. https://doi.org/10.1016/j.jclinepi.2013.10.019

Wang, M., Jiang, Z., Chen, X., Han, W., Zhu, L., He, T.-L., Shen, Y., 2025. Daily seamless dataset of HCHO concentrations: Vertical relationship between surface and column HCHO in China in 2019–2022. Atmospheric Environment 362, 121546. https://doi.org/10.1016/j.atmosenv.2025.121546

Wang, Y., Qi, Q., Liu, Y., 2018. Unsupervised segmentation evaluation using area-weighted variance and Jeffries-Matusita distance for remote sensing images. Remote Sensing 10, 1193.

Wang, Y., Yuan, Q., Li, T., Tan, S., Zhang, L., 2021. Full-coverage spatiotemporal mapping of ambient PM2.5 and PM10 over China from Sentinel-5P and assimilated datasets: Considering the precursors and chemical compositions. Science of The Total Environment 793, 148535. https://doi.org/10.1016/j.scitotenv.2021.148535

Wang, Y., Yuan, Q., Li, T., Yang, Y., Zhou, S., Zhang, L., 2023. 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.

Wang, Y., Yuan, Q., Xiao, R., Li, T., Zhang, L., 2020. Recovery of the Carbon Monoxide Product from S5P-TROPOMI by Fusing Multiple Datasets: A Case Study in Hubei Province, China, in: IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 5529–5532.

Wei, C., Fan, H., Xie, S., Wu, C.-Y., Yuille, A., Feichtenhofer, C., 2022. Masked Feature Prediction for Self-Supervised Visual Pre-Training, in: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, New Orleans, LA, USA, pp. 14648–14658. https://doi.org/10.1109/CVPR52688.2022.01426

https://doi.org/10.5194/essd-2025-817
Preprint. Discussion started: 27 January 2026

Wei, Q., Song, W., Dai, B., Wu, H., Zuo, X., Wang, J., Chen, J., Li, J., Li, S., Chen, Z., 2025. Spatiotemporal estimation of surface NO2 concentrations in the Pearl River Delta region based on TROPOMI data and machine learning. Atmospheric Pollution Research 16, 102353. https://doi.org/10.1016/j.apr.2024.102353

Wu, X., Xiao, Q., Wen, J., You, D., Hueni, A., 2019. Advances in quantitative remote sensing product validation: Overview and current status. Earth-Science Reviews 196, 102875. https://doi.org/10.1016/j.earscirev.2019.102875

Y. Jiang, C. Si, L. Yang, 2024. Improvement Strategies for Mask R-CNN in Satellite Image Analysis, in: 2024 3rd International Conference on Electronics and Information Technology (EIT). Presented at the 2024 3rd International Conference on Electronics and Information Technology (EIT), pp. 739–744. https://doi.org/10.1109/EIT63098.2024.10762551

Yan, C., Fan, X., Fan, J., Wang, N., 2022. Improved U-Net Remote Sensing Classification Algorithm Based on Multi-Feature Fusion Perception. Remote Sensing 14. https://doi.org/10.3390/rs14051118

Yang, Y., Zhou, M., Langerock, B., Sha, M.K., Hermans, C., Wang, T., Ji, D., Vigouroux, C., Kumps, N., Wang, G., 2020. New ground-based Fourier-transform near-infrared solar absorption measurements of XCO_2, XCH_4 and XCO at Xianghe, China.