This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3593486

Global estimates of daily gapless atmospheric XCH_4 concentrations from satellite and reanalysis data during 2003–2020

Yu Qu, Jing Wei, *Member*, *IEEE,* Hanfa Xing, Xian Shi, Zurui Ao, Xiaoliang Meng

***Abstract*—Methane is a potent greenhouse gas that significantly drives climate change and affects human health and ecosystems by promoting surface ozone formation. However, the spatial continuity of current satellite-derived column-averaged dry-air mole fraction of methane (XCH_4) data remains limited, and the sparse distribution of ground-based observations further hinders our understanding of methane sources and trends. To address this, we developed a novel spatiotemporal Transformer model to fill gaps in satellite retrievals and correct biases, thereby generating a global, daily, gapless XCH_4 dataset at 0.1° × 0.1° (~10 × 10 km near the equator) grid resolution from 2003 to 2020. This model integrates XCH_4 retrievals from the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) and Greenhouse Gases Observing Satellite (GOSAT), together with XCH_4 reanalysis from the Copernicus Atmosphere Monitoring Service (CAMS), meteorological variables, and other auxiliary datasets. Our fused daily XCH_4 dataset yielded a mean coefficient of determination (R^2) value of 0.93, a root-mean-square error (RMSE) of 13.10 ppb, and a bias of 7.45 ppb when independently compared with measurements from the Total Carbon Column Observing Network (TCCON). Furthermore, we performed bias correction on the reconstructed XCH_4 data using the TCCON network and supplementary variables to improve consistency across satellites and enhance accuracy, achieving a mean cross-validation R^2 (CV-R^2) of 0.97, an RMSE of 7.51 ppb, and a bias of 0.51 ppb. We observed a global annual growth rate of 6.09 ppb/yr ($p < 0.001$) from 2003 to 2020, with the average annual trends in developed countries being much lower than those in developing countries, at 5.86 ppb/yr and 6.16 ppb/yr ($p < 0.001$), respectively. Bangladesh exhibited the highest annual average concentration and trend, with average values of 1850.21 ± 41.85 ppb and 8.07 ppb/yr ($p < 0.001$), respectively. This unique global gapless daily XCH_4 dataset enables the precise detection of local emissions and provides valuable support for informing methane emission management policies and enhancing our understanding of global methane sources and their spatiotemporal variations.**

***Index Terms*—XCH_4; Transformer; data fusion; bias correction**

I. INTRODUCTION

Methane (CH_4) is a critical greenhouse gas in the Earth's atmosphere, with a radiative forcing effect approximately 20 times that of carbon dioxide over a century, with a radiative forcing of 0.61 W m^{-2}, contributing approximately 0.5 °C to global warming since pre-industrial times [1-4]. As a potent short-lived climate pollutant, methane plays a major role in climate change and poses direct threats to human health and ecosystems, primarily through its role in the formation of ground-level ozone [5-6]. In the last decade, global methane emissions have increased at an unprecedented rate compared to any other period in the past 30 years. This rise in methane emissions has intensified climate warming, resulting in a variety of detrimental impacts, including the increased frequency of extreme events and a decline in biodiversity [7-8]. The CH_4 concentration in the atmosphere continues to rise and has become one of the important causes of climate change [9]. To address this critical issue and gain a comprehensive understanding of CH_4 dynamics, precise concentration data is fundamental. This data enables the identification of emission sources, characterization of spatiotemporal variability, and attribution of changes to various drivers. Consequently, high-quality and long-term CH_4 datasets are essential for informing effective climate policy, designing robust mitigation strategies, and tackling the challenge of its continuous global rise [10-12].

Currently, ground-based networks and satellite-based remote sensing approaches are commonly used methods for monitoring atmospheric XCH_4 [13]. While ground-based networks like the Total Carbon Column Observing Network (TCCON) can provide high-precision fixed-point measurements, but their sites are sparsely distributed and cannot meet the needs of making refined assessments of the global methane budget [2]. In contrast, satellite-based offers a complementary approach by providing extensive spatial coverage, enabling regional and global XCH_4 monitoring despite lower temporal resolution.

This work was supported by the Scientific Research Innovation Project of Graduate School of South China Normal University (2025KYLX064), the National Natural Science Foundation of China (No.41971406 and No.42271470), National Natural Science Foundation of China (NSFC) (No. 42401556) and Guangdong Basic and Applied Basic Research Foundation (No. 2023A1515110774). (Corresponding author: Hanfa Xing)

Y. Qu and H. Xing are is with the School of Geographical Sciences, South China Normal University (SCNU), Guangzhou 510631, China (e-mail: quyu0813@m.scnu.edu.cn; xinghanfa@m.scnu.edu.cn).

J. Wei is with the College of Environmental Sciences and Engineering, Peking University, Beijing, China & Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA (e-mail: weijing_rs@163.com), and contributed equally to this work.

H. Xing and Z. Ao, is with the BeiDou Research Institute, School of Geographical Sciences, South China Normal University, Foshan 528225, China (e-mails: xinghanfa@m.scnu.edu.cn; aozurui1990@gmail.com).

X. S. is with the School of Surveying, Mapping and Spatial Information, Shandong University of Science and Technology, Qingdao 266590, China (e-mail: Shixian9786@126.com).

X. Meng is with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China (e-mail: xmeng@whu.edu.cn).

Authorized licensed use limited to: University of Maryland College Park. Downloaded on August 02,2025 at 14:41:43 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3593486

[14]. Satellites such as GOSAT, the Tropospheric Monitoring Instrument (TROPOMI), and the newly launched MethaneSAT mission in 2024 have significantly enhanced methane monitoring capabilities, enabling high-precision detection of emissions ranging from regional basins to individual point sources [15]. For example, MethaneSAT and TROPOMI have quantified emissions from the Permian Basin, revealing leakage rates higher than previously estimated [16-17], and identified extreme methane leaks, such as from a natural gas well blowout [18-19]. Satellite observations have also highlighted emission discrepancies in regions like Mexico [20], tracked variations linked to oil price changes during the COVID-19 pandemic [21], and quantified emissions across the US and Canada, identifying contributions from individual basins [22]. However, challenges remain, like retrieval limitations and uncertainties, especially spatial gaps in satellite retrievals, which complicate continuous XCH_4 coverage and limit its applications [22-24].

To generate spatially continuous daily XCH_4 products based on satellite observation data, researchers commonly apply techniques such as geospatial interpolation [25], multisource data fusion, and machine learning (ML) [26-31]. For instance, the National Institute for Environmental Studies employed the ordinary kriging interpolation method to produce GOSAT Level 3 data products, including global monthly mean XCO_2 and XCH_4 data [32]. Cao et al. used the kriging interpolation method to generate monthly XCH_4 data at a 0.1° resolution in the Northern Hemisphere from 2009 to 2021 [13]. Wang et al. (2023) proposed a novel spatiotemporal fusion method to generate long-term daily gapless XCO_2 and XCH_4 products from 2010 to 2020 [33]. Although geospatial interpolation methods are widely used to extend the spatial coverage of studies, accurately quantifying their associated uncertainties remains challenging, especially in regions with sparse observational data [34]. Comparatively, data fusion techniques enable the generation of highly accurate datasets with enhanced spatiotemporal resolution. However, variations in satellite coverage and differences in the spatiotemporal resolutions of individual sensors can introduce inconsistencies and impact the overall reliability of the result [35]. Recently, an increasing number of researchers have utilized ML to establish relationships between predictor variables and satellite-derived XCH_4 data. For example, Li et al. developed an unbiased ML model for estimating atmospheric XCH_4 by integrating satellite, meteorological, and terrain data to map its global distribution in 2021 [4]. To date, studies with high spatiotemporal resolution and long-term, gapless satellite-derived XCH_4 datasets at the global scale are lacking, which limits the insights that can be obtained regarding the dynamics of methane emissions and climate change. Existing ML models for XCH_4 reconstruction face challenges such as data quality issues, limited interpretability, and generalizability [36-38]. While deep learning (DL) models, including those with convolutional layers, have been applied for methane profile retrieval [39], the problem of spatiotemporal autocorrelation in multidimensional features remains unresolved [40-41].

To address these challenges, this study developed a novel DL model, called the four-dimensional spatiotemporal Transformer (4D-STransformer), to generate reliable, high-precision, global, daily, gapless XCH_4 datasets. This model effectively addresses issues related to uneven changes exhibited by geographical and temporal location features and nonlinear relationship modeling, filling gaps in satellite retrieval data. Using SCIAMACHY and GOSAT satellite retrieval XCH_4, CAMS XCH_4 reanalysis data, meteorological variables, and other auxiliary datasets, we generated a global, daily, gapless XCH_4 dataset at 0.1° grid for the period from 2003 to 2020. Bias correction was subsequently applied using TCCON data to enhance the dataset's consistency and reliability. This daily XCH_4 dataset enables detailed analysis of spatiotemporal variations in atmospheric XCH_4, exploration of its long-term trends, and precise detection of methane emission sources, such as wildfires, wetlands, and agricultural activities, across diverse regions and latitudes. Furthermore, explainable artificial intelligence techniques were applied to evaluate the contributions of various features during the data fusion and bias correction processes, providing insights into the drivers of XCH_4 variability.

II. Materials

The study incorporates four main categories of data sources, including multiple satellite XCH_4 retrievals, CAMS XCH_4 reanalysis, meteorological variables from the fifth-generation European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (ERA5), as summarized in Tables S1 and S2. All input datasets are resampled and aggregated to a 0.1° spatial resolution using bilinear interpolation and subsequently used to estimate daily XCH_4 concentrations. In addition, TCCON XCH_4 measurements are primarily utilized to validate the dataset generated in this study.

A. Satellite XCH_4 Retrievals

The XCH_4 data derived from the SCIAMACHY instrument and the Fourier Transform Spectrometer (FTS) on GOSAT were obtained using optimal estimation. SCIAMACHY launched onboard the Envisat satellite, provided data with spatial and temporal resolutions of 30 × 60 km and 30 days, respectively. The XCH_4 data were derived by inverting spectral information from the SCIAMACHY sensor from 2360-2380 nm [42], using the iterative maximum a posteriori-differential optical absorption spectroscopy (IMAP-DOAS) algorithm [43]. The GOSAT satellite, launched by Japan on January 23, 2009, aims to monitor greenhouse gases. It has a 3-day revisit period and an overpass time of 13:00 local time. The spatial resolution of the data is 10.5 km. In this study, we used GOSAT level-2 XCH_4 product generated via the proxy retrieval algorithm [44]. The XCH_4 retrievals were collected from SCIAMACHY for the period 2003-2009 and from GOSAT for 2010-2020, with quality assurance applied by discarding data flagged as "bad".

B. Reanalysis and Auxiliary Data

The Copernicus Atmosphere Monitoring Service (CAMS) XCH_4 reanalysis data from 2003 to 2020, with a spatial resolution of 0.75° and a temporal resolution of 3 hours, were collected [45]. Six meteorological variables were obtained, including surface pressure (SP), 2-meter temperature (T2M),

Authorized licensed use limited to: University of Maryland College Park. Downloaded on August 02,2025 at 14:41:43 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3593486

10-meter wind components (U10 and V10), and surface net solar radiation (SNSR) from the ERA5-Land hourly reanalysis at 0.1° resolution, along with relative humidity (RH) from the ERA5 reanalysis at 0.25° resolution. Additionally, total column water vapor (TCWV) from the ERA5 reanalysis (0.25°) and surface nitrogen dioxide (NO_2) concentrations from the CAMS reanalysis (0.25°) are also included. Normalized Difference Vegetation Index (NDVI) and Land Surface Temperature (LST), both derived from Moderate Resolution Imaging Spectroradiometer (MODIS) satellite products at a 0.05° resolution, were incorporated due to their strong associations with greenhouse gas concentrations and their substantial influence on methane fluxes between the biosphere and atmosphere [46]. In addition, a Digital Elevation Model (DEM) from the Shuttle Radar Topography Mission (SRTM), with a spatial resolution of 90 meters, was utilized to correct for terrain-induced effects on atmospheric XCH_4 retrievals [47].

C. Ground-based observation

TCCON is a global ground-based network dedicated to measuring the total atmospheric greenhouse gas columns, including methane. It employs a ground-based FTS method to capture near-infrared spectra, which are then used to retrieve atmospheric XCH_4 concentrations. In situ measurements from TCCON are widely used to validate XCH_4 data from satellites such as SCIAMACHY and GOSAT, as well as reanalysis data like CAMS. This study adopts TCCON measurements processed with the GGG2020 version from 30 stations worldwide (Fig. S1).

III. METHODOLOGY

In this study, we developed a Four-Dimensional Spatiotemporal Transformer (4D-STransformer) model to address the primary challenges of gap-filling and bias correction in atmospheric XCH_4 datasets, incorporating a transformer-based encoder that processes spatiotemporal features (Fig. 1).

A. 4D-STransformer framework

Atmospheric methane (CH_4) concentrations exhibit significant spatiotemporal variability, driven by complex interactions with various environmental variables. Consequently, reconstructing satellite-derived XCH_4 through data fusion is a challenging task, particularly when dealing with long sequences and high-dimensional multivariate inputs. Traditional approaches, including statistical regression approaches and conventional machine learning models often struggle to capture complex multivariate relationships, resulting in incomplete spatial or temporal gap-filling and reduced overall accuracy [48-49]. In contrast, the Transformer architecture, which employs a self-attention mechanism, captures long-range dependencies and integrates global contextual information, making it especially suitable for complex multivariate spatiotemporal fusion tasks. Despite its success in fields such as natural language processing and computer vision, the Transformer has been rarely applied in geosciences [50]. In particular, its potential for atmospheric parameter retrieval, such as XCH_4 reconstruction, remains underexplored. Therefore, the application of Transformer-based models presents a promising avenue for improving the accuracy and consistency of satellite-derived XCH_4 datasets through advanced spatiotemporal fusion techniques.

Here, we developed a 4D-STransformer model, an extension of the Transformer architecture that optimizes the integration of spatiotemporal information from independent variables. The spatiotemporal information is represented by Equations (1–3) and (4–6) [51-52]. The spatial positions of sample points (P_S) are represented in three-dimensional Euclidean space using spherical coordinates $[S_1, S_2, and\ S_3]$, while the temporal component (P_T) is captured by three helical trigonometric vectors $[T_1, T_2, and\ T_3]$ to account for periodic variations, such as seasonal and daily cycles of XCH_4. By encoding spatial and temporal attributes in this manner, the model gains the ability to leverage global spatial patterns and periodic temporal dynamics, which are crucial for accurately reconstructing XCH_4 concentrations.

$$S_1 = \sin\left(2\pi \frac{Lon}{360}\right) \quad (1)$$

$$S_2 = \cos(2\pi \frac{Lon}{360}) \sin(2\pi \frac{Lat}{180}) \quad (2)$$

$$S_3 = \cos(2\pi \frac{Lon}{360}) \cos(2\pi \frac{Lat}{180}) \quad (3)$$

$$T_1 = \frac{DOY}{N} \quad (4)$$

$$T_2 = \cos(2\pi \frac{DOY}{N}) \quad (5)$$

$$T_3 = \sin\left(2\pi \frac{DOY}{N}\right) \quad (6)$$

where Lon and Lat represent the longitude and latitude of one point in space, DOY and N represent the day of the year and total days in a year, respectively.

This study employs the 4D-STransformer model to address the key challenge of gap-filling in atmospheric XCH_4 datasets. The model architecture features a Transformer-based encoder that comprises positional encoding, multiple Transformer blocks, a multilayer perceptron, and a linear fully connected output layer to produce XCH_4 predictions. During the training phase, we adopted a meticulous iterative approach to fine-tune hyperparameters, systematically testing diverse combinations to identify the optimal configuration that maximizes model performance. Additionally, we implemented an early stopping strategy to prevent overfitting by halting training once peak performance was achieved. The 4D-STransformer includes four encoder layers, 64-dimensional hidden layers, and four attention heads, with rectified linear unit activation functions applied between layers to enhance learning and enable sparse feature extraction. The model was trained using the mean squared error (MSE) loss function. To generate a global daily full-coverage atmospheric XCH_4 dataset, we employed data fusion and bias correction steps, allowing independent evaluation and optimization of each stage to improve overall accuracy and performance (Fig. 1).

a). XCH_4 data fusion stage

The data fusion phase integrates satellite-derived XCH_4 data from SCIAMACHY and GOSAT satellites with CAMS-XCH_4 reanalysis data, meteorological variables, and other auxiliary

Authorized licensed use limited to: University of Maryland College Park. Downloaded on August 02,2025 at 14:41:43 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3593486

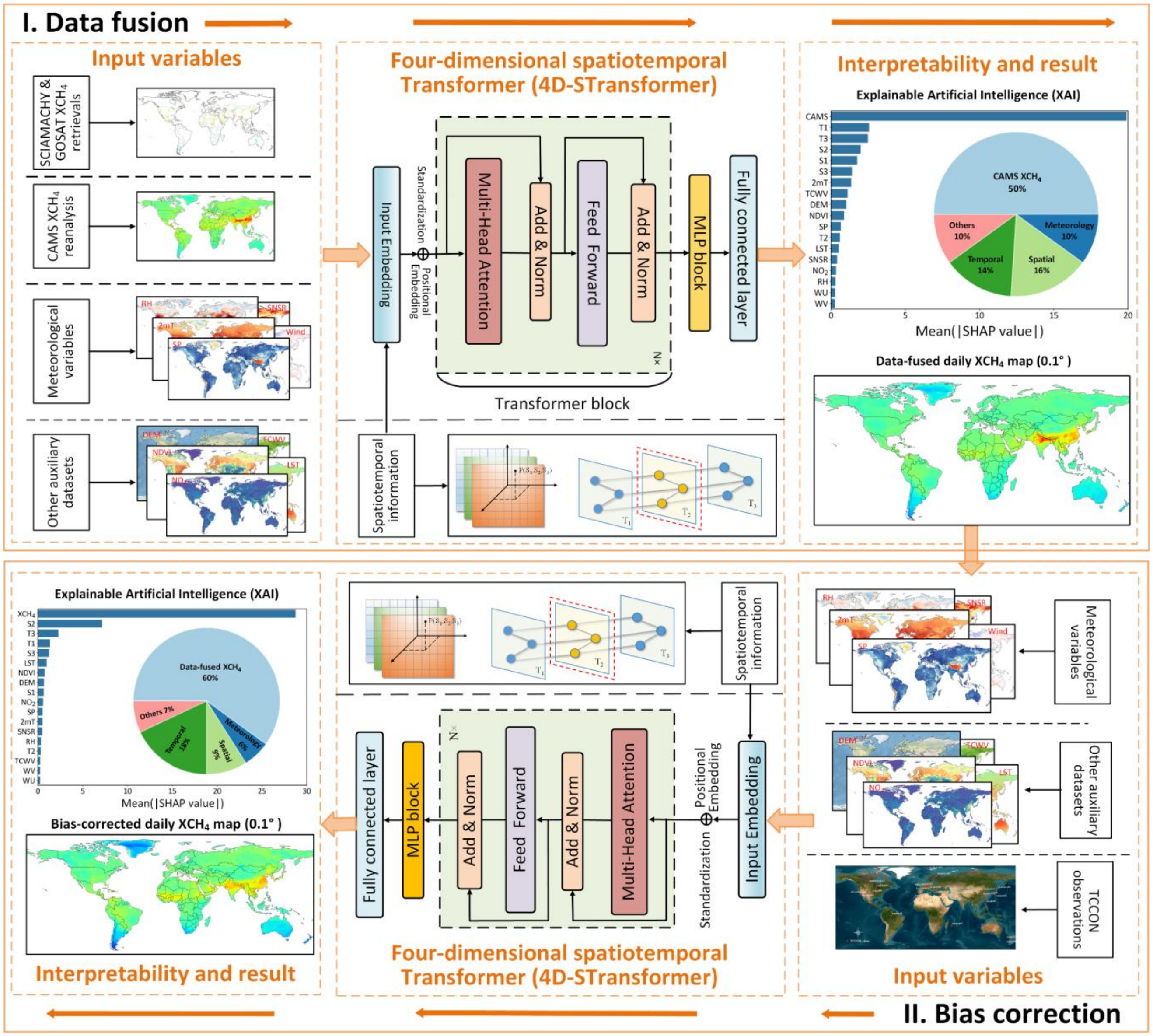

Fig. 1. Workflow for generating global daily gapless XCH_4 concentrations with a 0.1°× 0.1° spatial resolution using the developed four-dimensional spatiotemporal transformer (4D-STransformer) model.

datasets. The 4D-STansformer model was trained using these combined inputs to generate a gapless global daily XCH_4 dataset. By leveraging the spatial and temporal dependencies of the data, the model was able to produce continuous and accurate XCH_4 estimates across both space and time. Available SCIAMACHY and GOSAT retrievals XCH_4 are regarded as the observations, and the missing values are predicted by regressing the 4D-STransformer model with spatially continuous auxiliary variables, including CAMS-XCH_4, five meteorological variables, including surface pressure (SP), 2-meter temperature (T2M), wind speed (U/V10), surface net solar radiation (SNSR), as well as other related variables, i.e., normalized difference vegetation index (NDVI), total column water vapor (TCWV), Nitrogen dioxide (NO_2), land surface temperature (LST), and digital elevation model (DEM), and spatiotemporal terms (P_S and P_T).

$$Fused_{XCH_4} \sim f_{DL}(CAMS_{XCH_4}, SP, T2M, U10, V10, NDVI, TCWV, NO_2, LST, DEM, P_S, P_T) \quad (7)$$

b). XCH_4 bias correction stage

The bias correction phase addresses the inconsistencies and errors in the XCH_4 data derived from the SCIAMACHY and GOSAT satellites, which arise from differences in retrieval algorithms and sensor characteristics. We applied a bias correction step using the 4D-STransformer model, which was provided with XCH_4 retrievals from the TCCON network as reference observations. The model was also fed with the XCH_4 fusion results from the first phase, along with meteorological variables and additional datasets. By using TCCON data as a reliable reference, the bias correction process effectively adjusted the satellite data, reducing discrepancies between the satellite-derived XCH_4 and ground-based observations:

Authorized licensed use limited to: University of Maryland College Park. Downloaded on August 02,2025 at 14:41:43 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3593486

$$Bias_{XCH_4} \sim f_{DL}(Fused_{XCH_4}, SP, T2M, U10, V10, NDVI, TCWV, NO_2, LST, DEM, P_S, P_T) \quad (8)$$

This two-step approach, encompassing data fusion and bias correction, enhances the interpretability of the model while enabling quantification of improvements, such as reducing average bias and errors across various validation methods.

B. Feature contribution analysis

To explain the contributions of different predictive factors to the output of the 4D-STransformer, the SHapley Additive exPlanations (SHAP) method was applied to the training data to evaluate the importance of each predictive factor based on its SHAP values [53]. We used summary plots to understand the impacts of certain individual predictors on the predicted outcome and to explain the predictors of all the training data.

In the data fusion stage (Fig. 1), CAMS XCH_4 was the primary predictor, with an importance value of 50%. It provides a high-frequency, daily baseline background field that effectively captures XCH_4 variability across diverse regions and time periods, thus enhancing the model's predictive accuracy. Meteorological variables had a combined importance value of 10%, and these variables influence XCH_4 spatiotemporal distributions by impacting methane emissions and uptake: surface pressure and relative humidity affect microbial activity in wetlands and soils, key methane sources; surface net solar radiation drives photochemical reactions influencing methane oxidation rates; and surface pressure and winds modulate atmospheric transport and mixing, affecting XCH_4 dispersion and regional accumulation. Spatiotemporal variables (16% spatial and 14% temporal) significantly influence the 4D-STransformer model's XCH_4 predictions by capturing both seasonal and localized emission patterns. Other predictors collectively contributed an importance value of 10%. For example, land surface temperature positively correlates with XCH_4 as higher temperatures enhance methane emissions from soils and wetlands; NDVI indirectly affects XCH_4 by reducing atmospheric carbon dioxide in summer through photosynthesis, altering greenhouse gas ratios; DEM accounts for topographic influences on atmospheric XCH_4 spatial distributions via local atmospheric mixing and emission patterns; and nitrogen dioxide reflects its indirect link to combustion-related methane emissions from shared sources like fossil fuel activities [54].

In the bias correction stage (Fig. 1), the reconstructed-XCH_4 replaced the CAMS XCH_4 in the first step, with an importance value of 61%. Compared to the data fusion stage, the importance value of meteorological variables decreased to 5%, while spatial and temporal variables reached 19% and 9%, respectively. The remaining predictors had a combined importance value of 7%, reflecting their adjusted roles in refining the model's accuracy during bias correction.

IV. RESULTS

A. Validation and comparison of XCH_4 retrievals

The TCCON measurements taken within ±1 hour of the satellite overpass times (~10:00 and 13:30 local time) were matched with the fused XCH_4 dataset within a 2° radius of each station. The 2° window was chosen to balance the spatial resolution of the satellite observation data and the coverage of ground-based measurements, ensuring a sufficient number of matched data points for robust validation. Validation results were obtained for the satellite observations (Fig. 2a), the CAMS reanalysis data (Fig. 2b), and in situ measurements from TCCON sites. The validation results obtained for the satellite observations showed an R^2 of 0.91, RMSE values of 12.57 ppb, and MAE values of 9.56 ppb, along with a mean bias of 5.86 ppb, respectively. Similarly, the validation results produced for the CAMS data exhibited an R^2 of 0.90, with RMSE and MAE values of 17.24 and 13.82 ppb, respectively, and a mean bias of 11.07 ppb. Our fused data showed the best accuracy against TCCON data, yielding an R^2 of 0.93, an RMSE of 13.10 ppb, an MAE of 10.38 ppb, and a mean bias of 7.45 ppb (Fig. 2c).

Finally, we conducted bias correction on the global daily XCH_4 dataset utilizing TCCON observations as the benchmark, employing three distinct 10-fold cross-validation (CV) methods to assess the performance of the 4D-STransformer model. In the sample-based CV (Fig. 2d), random samples were selected to evaluate the overall performance of the model comprehensively; in the temporal-based CV (Fig. 2e), the dataset was divided by days to test the model's temporal consistency across periods; and in the spatial-based CV (Fig. 2f), the data was segmented by geographical regions to examine the model's spatial generalization effectively. Prior to bias correction, systematic biases were evident, particularly across varying temporal and spatial patterns. After applying bias correction, the 4D-STransformer demonstrated substantial improvements: mean biases for the sample-based, temporal-based, and spatial-based CV were reduced to 0.02, -1.30, and -0.87 ppb, respectively, indicating effective bias mitigation. Additionally, the CV-R^2 values reached 0.97, 0.95, and 0.95; MAE values decreased to 5.43, 7.48, and 7.25 ppb; and RMSE values dropped to 7.51, 9.86, and 9.61 ppb, respectively. These results highlight the robustness and reliability of the 4D-STransformer in correcting biases and improving the accuracy of the global XCH_4 dataset across diverse temporal and spatial domains.

We further evaluated the 4D-STransformer performance in reconstructing the XCH_4 dataset by excluding 2020 TCCON data (Fig. S2). Our bias-corrected XCH_4 dataset demonstrates a notable improvement in accuracy compared to the original, as evidenced by the increase in R^2 from 0.90 to 0.92 against the 2020 TCCON observations. The RMSE decreased from 14.36 ppb to 9.36 ppb (a 35% improvement), and the MAE reduced from 11.52 ppb to 7.08 ppb (a 39% improvement). The bias was also significantly decreased from 9.94 ppb to -1.91 ppb. These results further confirm that our bias correction method remains robust and reliable, even when ground-based measurements from an entire year are excluded from the validation.

Authorized licensed use limited to: University of Maryland College Park. Downloaded on August 02,2025 at 14:41:43 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3593486

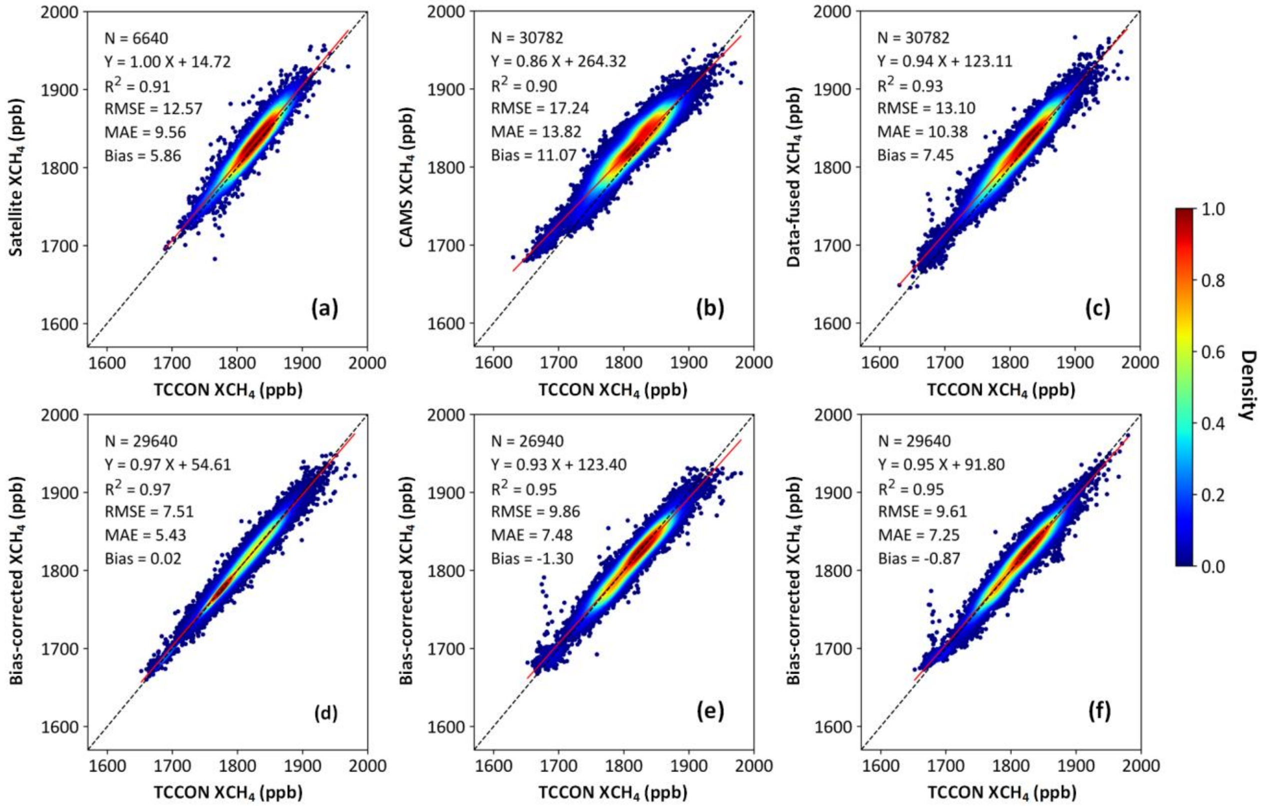

Fig. 2. Density scatterplots of validation results for XCH_4 concentrations derived from (a) satellite data, (b) CAMS reanalysis, (c) our data fusion approach, compared with TCCON ground measurements, as well as cross-validation results of our bias-corrected XCH_4 concentrations using (d) sample-based, (e) temporal-based, and (f) spatial-based 10-fold CV methods. The dashed lines denote the 1:1 line, and the red lines denote the best-fit lines derived from linear regression equations.

B. Daily XCH_4 mapping and emission source identification

We first compared and analyzed the global daily gapless XCH_4 concentration distributions for February 10, 2018, June 20, 2019, and October 15, 2020, using data from SCIAMACHY and GOSAT satellites, CAMS XCH_4 reanalysis data, and our generated results (Fig. 3). The satellite data exhibit sparse daily coverage and incomplete spatial coverage, particularly in remote regions such as high-latitude areas. In contrast, the CAMS XCH_4 provides more continuous global coverage but still faces limitations in accurately capturing localized emissions and trends in specific regions due to model constraints. Our results combine the strengths of satellite observations, CAMS model outputs, and our bias correction methods, offering improved spatial resolution and temporal consistency. While CAMS XCH_4 already assimilates satellite data and performs reanalysis, its spatial resolution of 0.75° is insufficient to capture finer local variations, and satellite data itself suffers from temporal gaps. In contrast, our model enhances the spatial resolution to 0.1° and corrects the systematic biases between SCIAMACHY and GOSAT satellite data and TCCON observations, yielding a more consistent and accurate global XCH_4 dataset. Collectively, our product offers higher spatial resolution and enhanced accuracy compared to existing datasets, making it more suitable for locating methane emission sources and analyzing emission trends, thereby providing significant scientific research value.

Daily atmospheric XCH_4 datasets remain essential for precisely identifying both the timing and the locations of severe pollution events triggered by natural disasters across the globe. Fig. 4 illustrates the spatial distributions of the XCH_4 concentrations across ten distinct regions of interest, each characterized by typical methane emission sources, aiming to assess the effectiveness of our methane retrievals in detecting high global methane emitters. Significant methane anomalies were identified in regions associated with oil and gas extraction as well as coal mining activities, including the Permian Basin in Texas, the United States, and the northern part of Colombia. Higher concentrations were observed at industrial and fossil fuel extraction sites; for instance, elevated methane levels were detected over chemical factories in Russia, power plants in Bangalore, India, coal and copper mines in Shanxi and Jiangxi, China, and coal mining areas in Australia. Additionally, a substantial methane anomaly was observed over an oil facility in Saudi Arabia. These high concentrations in this region are likely attributable to extensive oil extraction, and the related infrastructure, including pipelines and refineries, as methane emissions from oil and gas operations are recognized as major contributors to global greenhouse gas levels [55]. In contrast, natural sources such as wetlands and paddy fields also presented substantial methane emissions. Methane anomalies were prominent over agricultural regions [56-57], including eastern France, which has expansive paddy rice fields, and the wetlands in Central Africa. Wetlands are well-known natural

Authorized licensed use limited to: University of Maryland College Park. Downloaded on August 02,2025 at 14:41:43 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3593486

sources of methane due to anaerobic decomposition of organic matter; however, their emissions are generally more spatially diffuse compared to the more concentrated and localized emissions from industrial activities [58-60]. Overall, the satellite-based XCH_4 observations effectively identified and pinpointed high methane emitters across diverse regions. The successful detection of methane emission sources highlights the capability of our retrievals in monitoring emissions and supporting global methane-related budget management efforts.

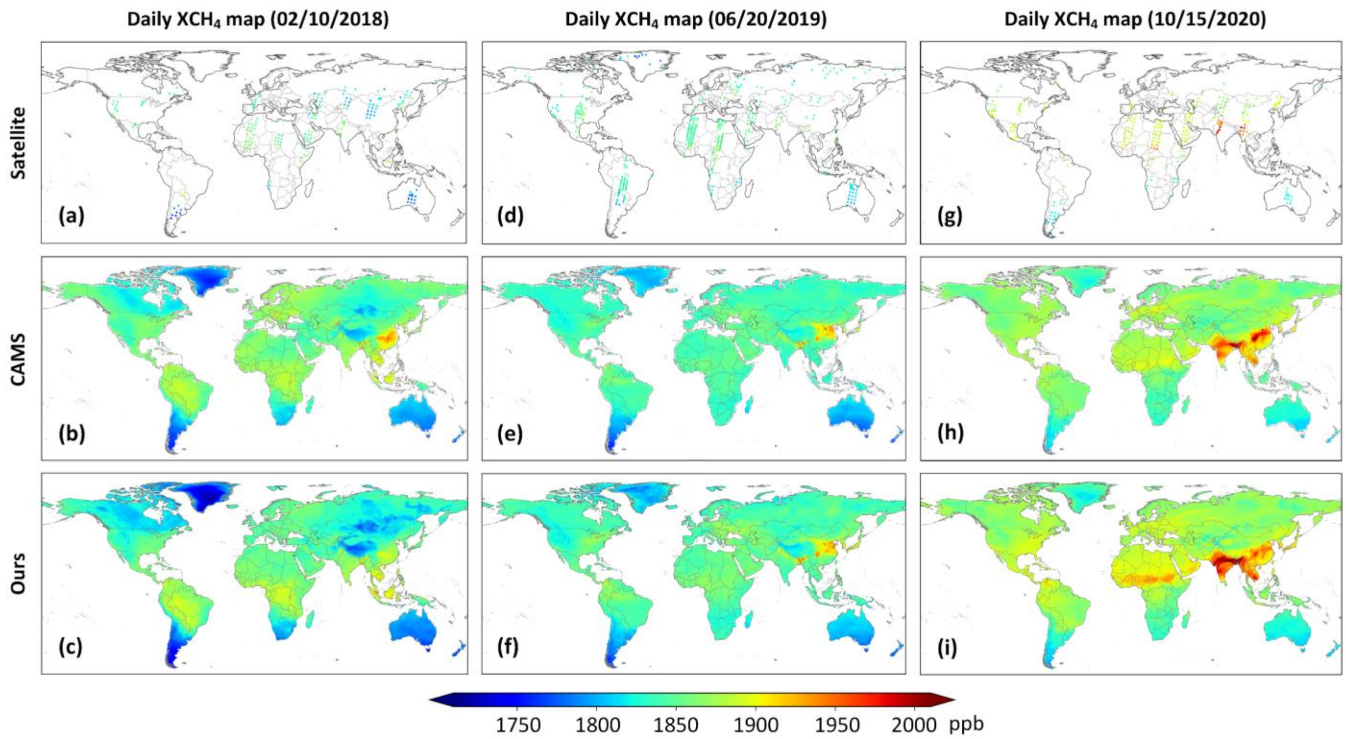

Fig. 3. Comparisons of spatial distribution for global daily XCH_4 concentrations derived from (a, d, g) satellite data, (b, e, h) CAMS reanalysis, (c, f, i) our data fusion approach on several individual days: 02/10/2018, 06/20/2019, and 10/15/2020.

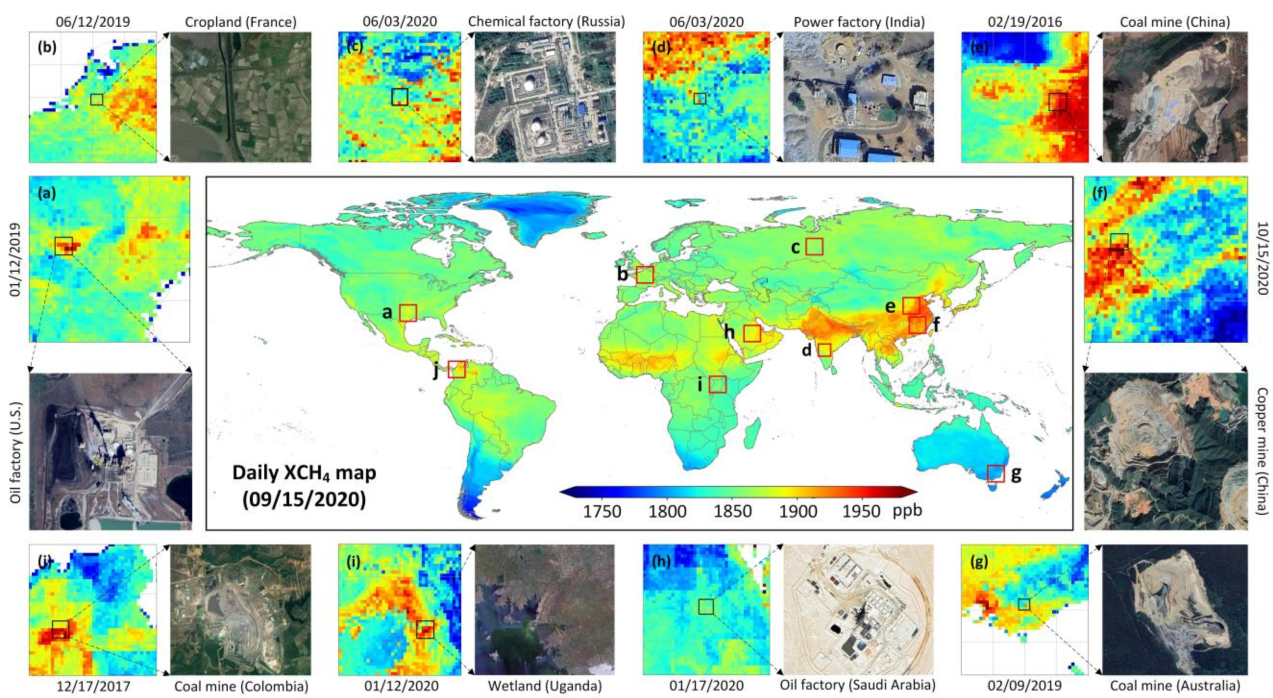

Fig. 4. Identification of major methane emission sources detected from daily gapless XCH_4 data across ten regions of interest on different days around the world.

Authorized licensed use limited to: University of Maryland College Park. Downloaded on August 02,2025 at 14:41:43 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3593486

C. Long-term XCH_4 distributions and trends

Fig. S3 provides the spatial distribution of the global annual mean XCH_4 concentrations for each year from 2003 to 2020. Fig. 5 further compares the global annual mean XCH_4 spatial distributions and temporal trends from 2003 to 2020, derived from satellite retrievals, CAMS reanalysis XCH_4, and our bias-corrected dataset. In the Figure, the satellite results appear continuous because we calculated the annual mean XCH_4 concentrations at the state level for each country. This averaging process smooths out the spatial gaps and striping, facilitating a clearer comparison with our and CAMS data in terms of global distribution and temporal trends. The spatial distribution of our XCH_4 dataset (Fig. 5c) closely aligns with satellite data (Fig. 5a) [33,43] while mirroring global patterns in the CAMS XCH_4 dataset (Fig. 5b). Our XCH_4 dataset with a higher resolution of 0.1°, provides greater spatial detail than CAMS reanalysis XCH_4 (0.75°), enhancing its utility for fine-scale methane emission source analysis.

A consistent global upward trend in XCH_4 is observed, increasing from 1765.61 ± 19.80 ppb in 2003 to 1862.92 ± 23.66 ppb in 2020. The multiyear mean XCH_4 from our dataset is 1797.91 ± 21.33 ppb, closely matching CAMS reanalysis XCH_4 (1800.72 ± 19.57 ppb) and satellite XCH_4 (1800.31 ± 35.11ppb), demonstrating consistency across independent sources. Globally, the atmospheric XCH_4 shows regional variations, with higher concentrations in densely populated and industrial areas such as parts of East and Southeast Asia, Eastern China, northern India, and low latitudes in the United States. These elevated XCH_4 values are often linked to fossil fuel usage, high population density, industrial activities, oil and gas extraction, and human activities [14,54]. In South America, seasonal flooding of the Amazon River basin and its adjacent wetlands creates anoxic conditions that are important sources of high XCH_4 concentrations. The XCH_4 concentrations over Europe are not as intense as those over the Amazon, but the European sources are influenced by population and the exploration and transportation of oil, gas, and coal by the energy sector. As the latitude increases, human activity decreases, and the XCH_4 concentration slightly decreases. On average, developed countries have an XCH_4 concentration of 1801.08 ± 15.48 ppb, while developing countries show a slightly higher average of 1803.22 ± 17.55 ppb. Table S3 shows that Bangladesh has the highest average XCH_4 concentration globally (1850.41 ± 41.85 ppb), followed by India (1829.82 ± 37.43 ppb). The higher XCH_4 concentration in Bangladesh is due to its higher population density, extensive flooding during the summer monsoon, rice cultivation, and the high emission intensity of vertical atmospheric transport [27,61-63]. The high XCH_4 concentrations in India are concentrated around the Kolkata industrial area in the northeastern coastal region, where significant emissions stem from traditional industries such as coal and iron, as well as from extensive rice cultivation in the Indo-Gangetic Plain in northern India [64-65].

To assess the capability of our XCH_4 dataset in capturing interannual methane variations, methane observations from three TCCON sites representing different land covers were selected for comparison. Fig. S4 shows that the long-term variations of XCH_4 at Wollongong (wetland, Australia) from 2013 to 2020, Parkfalls (urban, United States) from 2013 to 2020, and Garmisch (rural, Germany) from 2007 to 2020 were well bias-corrected by our XCH_4 results. Notably, all XCH_4 time series indicated substantial increasing trends of methane over these three regions, with rates of 8.10, 7.56, and 7.69 ppb/yr ($p < 0.001$), respectively. These consistent and alarming upward trends in XCH_4 concentrations underscore the urgent need to implement targeted emission reduction strategies to strengthen greenhouse gas mitigation efforts, particularly given methane's disproportionately high climate impact.

The temporal trends indicate XCH_4 increases of 6.42 ppb/yr from the satellite XCH_4, 6.08 ppb/yr from the CAMS XCH_4, and 6.09 ppb/yr ($p < 0.001$) from the dataset developed in this study (Fig. S5). While the trends differ slightly, our method effectively captures the temporal evolution process of XCH_4, especially in regions with significant methane increases, thus aligning well with satellite data. The XCH_4 trend remained almost unchanged from 2003 to 2006, and strong growth resumed in 2007, becoming stronger starting in 2014. The average annual trend in tropical regions was 6.53 ppb/yr ($p < 0.001$), with the tropics significantly contributing to methane growth since 2007, including during the 2015 El Niño (Fig. S6). The temporal trend maps (Fig. 5d-f) confirm the accuracy of our approach, with our trends closely mirroring the satellite trends in key XCH_4 hotspots, such as South Asia and the Middle East. Additionally, discrepancies in the CAMS XCH_4 data trends, where underestimations are evident, are addressed. Fig. S7 highlights that the largest XCH_4 trends are found in tropical regions, including central Africa, Southeast Asia, and the Amazon Basin [66]. The average annual mean temporal trend in developed countries was 5.86 ppb/yr ($p < 0.001$), whereas the annual mean temporal trend in developing countries was 6.16 ppb/yr ($p < 0.001$). Due to technological progress and industrial structure upgrades, the temporal trends in developed countries were relatively low. Bangladesh had the highest trend (8.07 ppb/yr, $p < 0.001$), followed by the Central African Republic (7.87 ppb/yr, $p < 0.001$) (Table S4 and Fig. S5).

The scatter plots (Fig. 5g-h) demonstrate the validation of our method by comparing the annual mean XCH_4 concentrations and temporal trends from our dataset against those derived from satellite observations and CAMS reanalysis XCH_4 data from the number of country matches. The inclusion of matched countries (N 141 for satellites and 230 for CAMS) ensures that comparisons are made on a consistent and fair basis, avoiding potential biases due to missing data. Our dataset shows a near-perfect correlation with satellite-derived annual mean XCH_4 concentrations (R = 0.98) and exhibits significantly lower errors (MAE = 3.31, RMSE = 4.06 ppb) compared to CAMS (R = 0.86, MAE = 9.89, RMSE = 11.77 ppb). This comparison indicates that our method performs more accurately in capturing XCH_4 concentrations than CAMS XCH_4 data. For temporal trends (Fig. 5i-j), our dataset also demonstrates a strong correlation with satellite data (R = 0.81), accompanied by lower error values (MAE = 0.45, RMSE = 0.55 ppb/yr, $p <$

Authorized licensed use limited to: University of Maryland College Park. Downloaded on August 02,2025 at 14:41:43 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3593486

0.001) compared to CAMS reanalysis XCH_4 data (R = 0.52, MAE = 0.38, RMSE = 0.56 ppb/yr, $p < 0.001$). This consistent outperformance, demonstrating enhanced accuracy in capturing atmospheric XCH_4 concentrations and temporal trends, further substantiates the comprehensive robustness and improved reliability of our developed methodological approach.

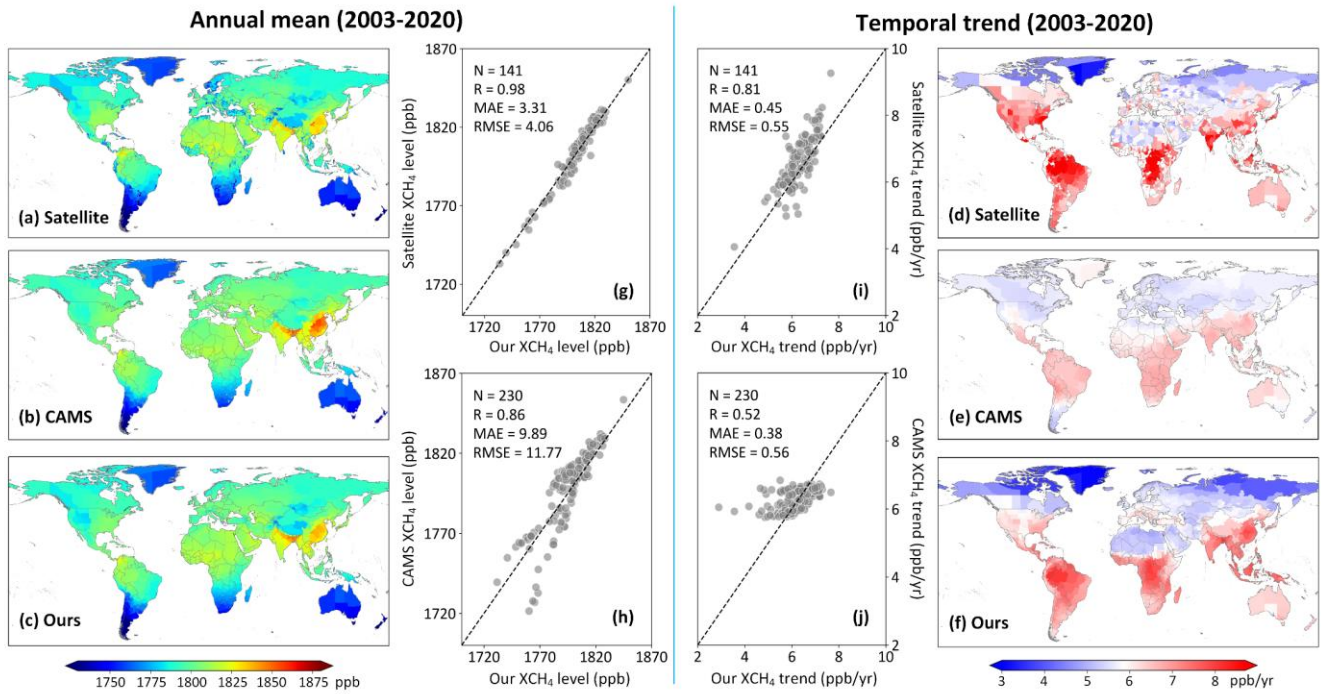

Fig. 5. Comparisons of spatial distribution for global state-level annual mean XCH_4 concentrations derived from (a) satellite data, (b) CAMS reanalysis, and (c) our dataset, along with their temporal trends (d-f) for the period 2003-2020, respectively. Density scatterplots show the comparisons of (g & h) annual mean and (i & j) temporal trends of XCH_4 concentrations between satellite, CAMS, and our datasets at each country during 2003-2020, with the dashed lines representing the 1:1 line.

Atmospheric XCH_4 shows significant seasonal variations (Fig. S8). The highest average concentrations occur during September–November (SON), with elevated levels observed in central South America and southern Africa due to wildfires and biomass burning, as well as in rural eastern China driven by straw burning activities [67]. Tropical wetlands and rice cultivation in Central Africa and the Amazon Basin also contribute to these peaks [68]. In contrast, March–May (MAM) shows the lowest average concentrations, largely due to reduced emissions in high-latitude regions like southern South America and Europe, influenced by clean oceanic air masses [69]. June–August (JJA) displays moderate XCH_4 concentrations, with notable increases over northern India and Southeast Asia linked to intensified rice cultivation and wetland emissions during the monsoon season [68]. December–February (DJF) exhibits slightly higher levels than MAM, particularly in the Northern Hemisphere, driven by reduced atmospheric mixing and increased fossil fuel combustion in winter [69]. Latitudinally, XCH_4 peaks are most pronounced between 10° and 30°N in SON and JJA, reflecting strong tropical wetland emissions, while MAM and DJF display a more uniform distribution across latitudes [70]. Monthly variations (Fig. S9) further confirm a distinct seasonal cycle, with XCH_4 peaking in late JJA to early SON due to enhanced biogenic emissions and reaching a minimum in MAM due to weaker emission activity [71]. These findings emphasize the importance of accounting for seasonal dynamics in atmospheric XCH_4 reconstruction models to improve their accuracy and reliability.

V. DISCUSSIONS

We compared our results with previous studies by applying the same TCCON XCH_4 observations for independent validation (Table I). Our model, with daily temporal and 0.1° spatial resolutions, demonstrated superior accuracy ($R^2 = 0.93$, RMSE = 13.10 ppb), outperforming other reported methods. These include the kriging interpolation method (monthly, 0.1°, R^2 = 0.73, and RMSE = 17.30 ppb) [12], the random forest (RF) model (daily, 0.05°, R^2 = 0.83, and RMSE = 17.16 ppb) [3], the self-supervised approach (daily, 0.25°, R^2 = 0.90, and RMSE = 13.03 ppb) [25], geostatistical analysis (monthly, 1°, R^2 = 0.84) and empirical orthogonal function (EOF) methods (monthly, 0.25°, R^2 = 0.86) [72,73]. After implementing bias correction, our XCH_4 dataset yielded an average CV-R^2 of 0.97 and an RMSE of 7.51 ppb, highlighting the superior accuracy.

Finally, we evaluated our model's performance against widely used AI models through spatiotemporal validation, training the models on data from 2010-2020 and testing on data from the remaining years (Table II). All models demonstrated strong predictive capabilities for global XCH_4 levels, with R^2 to validation data ranging from 0.76 to 0.79. Generally, MAE and

Authorized licensed use limited to: University of Maryland College Park. Downloaded on August 02,2025 at 14:41:43 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3593486

RMSE values were below 12 and 15 ppb, respectively. Among these, tree-based machine learning models showed comparable performance, with R^2 values ranging from 0.76 to 0.79. Deep learning models, leveraging advanced data-mining capabilities and nonlinear problem handling, achieved higher predictive accuracy (e.g., $R^2 = 0.83–0.90$), particularly recent architectures such as Convolutional Neural Networks (CNNs). Transformer, a state-of-the-art deep learning model leveraging self-attention mechanisms, outperformed all other models, as evidenced by the highest correlation and lowest MAE and RMSE values.

TABLE I. COMPARISONS WITH PREVIOUS STUDIES FOCUSING ON GLOBAL XCH_4 ESTIMATION.

Method	Temporal resolution	Spatial resolution	Overall accuracy		Time period	Literature
			R^2	RMSE		
Geostatistical analysis	Monthly	1°	0.84	-	2009-2020	[72]
Empirical orthogonal	Monthly	0.25°	0.86	9.38	2010-2021	[73]
Self-supervised	Daily	0.25°	0.90	13.03	2010-2020	[33]
Kriging interpolation	Daily	0.1°	0.73	17.30	2009-2021	[13]
Random forest	Daily	0.05°	0.83	17.16	2021	[3]
This study	Daily	0.1°	0.93	12.25	2003-2020	

TABLE II. COMPARISON OF PERFORMANCE ACROSS VARIOUS MACHINE LEARNING AND DEEP LEARNING MODELS FOR XCH_4 ESTIMATION.

Category	Core model	R^2	RMSE	MAE
Machine learning	LightGBM	0.76	15.24	11.71
	Extra Trees	0.77	15.77	12.32
	CatBoost	0.79	14.80	11.35
	XGBoost	0.79	14.64	11.23
Deep learning	MLP	0.83	18.63	14.38
	LSTM	0.85	13.86	10.56
	CNN	0.86	10.34	9.48
	Transformer	0.90	11.15	8.35

VI. CONCLUSIONS

This study developed a 4D-STransformer model, a deep learning approach designed to address the complex nonlinear problems involved in decoupling the Earth–atmosphere system and reconstructing a global XCH_4 dataset from multisource datasets. The model generates a global, daily, gapless XCH_4 product at 0.1° spatial resolution by establishing nonlinear relationships between XCH_4 retrievals from the SCIAMACHY and GOSAT satellites, and various auxiliary data, including CAMS XCH_4 reanalysis, meteorological factors, and other relevant variables. Using the XAI-based SHAP method, we revealed the model's internal mechanisms, showing that spatiotemporal information and CAMS reanalysis contribute over 25% to the fusion process. The resulting daily gapless XCH_4 dataset, spanning 2003 to 2020, achieves high accuracy compared to TCCON observations ($R^2 = 0.93$, RMSE = 13.10 ppb). After bias correction, to improve consistency between satellites, cross-validation results showed a CV-R^2 of 0.97 and an RMSE of 7.51 ppb. Compared with previous studies and conventional models, our model achieved superior accuracy. Long-term trend analysis from 2003 to 2020 reveals a global annual increase rate of 6.09 ppb/yr ($p < 0.001$), with tropical regions exhibiting a higher trend of 6.53 ppb/yr ($p < 0.001$). Since 2007, the tropics have been a major driver of methane growth, particularly during the 2015 El Niño event. Our results present a more comprehensive and detailed view of global XCH_4 distribution and variation, particularly in areas that are challenging to monitor with fewer observations.

While this study provides a high-quality global XCH_4 dataset, future research could further enhance overall accuracy and spatial resolution by incorporating advanced deep learning models and additional data sources, such as TROPOMI or MethaneSAT satellite observations. Additionally, regional analyses targeting key methane sources, particularly methane emissions from agriculture or wildfires, would offer deeper insights into localized trends. Integrating XCH_4 dataset with climate models may help refine global climate predictions and inform emission reduction strategies. Continued exploration in these areas will strengthen our understanding of methane's role in climate change.

DATA AVAILABILITY

The generated global daily, gapless XCH_4 datasets are available at https://zenodo.org/records/14184913.

REFERENCES

[1] Etminan M, Myhre G, Highwood E J, et al. Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing[J]. Geophysical Research Letters, 2016, 43(24): 12,614-12,623.

[2] Allen R J, Zhao X, Randles C A, et al. Surface warming and wetting due to methane's long-wave radiative effects muted by short-wave absorption[J]. *Nature Geoscience*, 2023, 16(4): 314-320.

[3] Reuter M, Hilker M, Noël S, et al. A data-driven method to retrieve XCO2 and XCH4 using artificial neural networks in preparation for the European Copernicus CO2 Monitoring Mission CO2M[R]. Copernicus Meetings, 2024.

[4] Li K, Bai K, Jiao P, et al. Developing unbiased estimation of atmospheric methane via machine learning and multi-objective programming based on

Authorized licensed use limited to: University of Maryland College Park. Downloaded on August 02,2025 at 14:41:43 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3593486

TROPOMI and GOSAT data[J]. Remote Sensing of Environment, 2024, 304: 114039.

[5] Masson-Delmotte V, Zhai P, Pirani A, et al. Climate change 2021: the physical science basis[J]. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change, 2021, 2(1): 2391.

[6] Nisbet E G, Fisher R E, Lowry D, et al. Methane mitigation: methods to reduce emissions, on the path to the Paris agreement[J]. *Reviews of Geophysics*, 2020, 58(1): e2019RG000675.

[7] Malley C S, Henze D K, Kuylenstierna J C I, et al. Updated global estimates of respiratory mortality in adults≥ 30 years of age attributable to long-term ozone exposure[J]. Environmental health perspectives, 2017, 125(8): 087021.

[8] Marques A, Martins I S, Kastner T, et al. Increasing impacts of land use on biodiversity and carbon sequestration driven by population and economic growth[J]. Nature ecology & evolution, 2019, 3(4): 628-637.

[9] Te T, Bao C, Bagan H, et al. Mapping seamless monthly XCO2 in East Asia: Utilizing OCO-2 data and machine learning[J]. International Journal of Applied Earth Observation and Geoinformation, 2024, 133: 104117.

[10] Lan, X., Thoning, K., Dlugokencky, E., 2022. Trends in Globally-Averaged CH4, N2O, and SF6 Determined from NOAA Global Monitoring Laboratory Measurements. Version 2022-10.

[11] Bousquet, P., Ciais, P., Miller, J.B., Dlugokencky, E.J., Hauglustaine, D.A., Prigent, C.,Van der Werf, G.R., Peylin, P., Brunke, E.-G., Carouge, C., Langenfelds, R.L. ,Lathi`ere, J., Papa, F., Ramonet, M., Schmidt, M., Steele, L.P., Tyler, S.C., White, J.,2006. Contribution of anthropogenic and natural sources to atmospheric methane variability. Nature 443, 439–443.

[12] IPCC, 2023. Summary for Policymakers. In: Climate Change 2022 - Mitigation of Climate Change. Cambridge University Press, pp. 3–48.

[13] Cao, Xiaoyan, et al. Spatial-temporal variation in XCH4 during 2009–2021 and its driving factors across the land of the Northern Hemisphere. Atmospheric Research 291 (2023): 106811.

[14] Zhang, Lingfeng, Tongwen Li, and Jingan Wu. Deriving gapless CO2 concentrations using a geographically weighted neural network: China, 2014–2020. International Journal of Applied Earth Observation and Geoinformation 114 (2022): 103063.

[15] Jacob D J, Turner A J, Maasakkers J D, et al. Satellite observations of atmospheric methane and their value for quantifying methane emissions[J]. Atmospheric Chemistry and Physics, 2016, 16(22): 14371-14396.

[16] Chan Miller C, Roche S, Wilzewski J S, et al. Methane retrieval from MethaneAIR using the CO 2 proxy approach: a demonstration for the upcoming MethaneSAT mission[J]. Atmospheric Measurement Techniques, 2024, 17(18): 5429-5454.

[17] Zhang Y, Gautam R, Pandey S, et al. Quantifying methane emissions from the largest oil-producing basin in the United States from space[J]. Science advances, 2020, 6(17): eaaz5120.

[18] Jacob D J, Varon D J, Cusworth D H, et al. Quantifying methane emissions from the global scale down to point sources using satellite observations of atmospheric methane[J]. Atmospheric Chemistry and Physics, 2022, 22(14): 9617-9646.

[19] Pandey S, Gautam R, Houweling S, et al. Satellite observations reveal extreme methane leakage from a natural gas well blowout[J]. Proceedings of the National Academy of Sciences, 2019, 116(52): 26376-26381.

[20] Shen L, Zavala-Araiza D, Gautam R, et al. Unravelling a large methane emission discrepancy in Mexico using satellite observations[J]. Remote Sensing of Environment, 2021, 260: 112461.

[21] Lyon D R, Hmiel B, Gautam R, et al. Concurrent variation in oil and gas methane emissions and oil price during the COVID-19 pandemic[J]. Atmospheric Chemistry and Physics, 2021, 21(9): 6605-6626.

[22] Shen L, Gautam R, Omara M, et al. Satellite quantification of oil and natural gas methane emissions in the US and Canada including contributions from individual basins[J]. Atmospheric chemistry and physics, 2022, 22(17): 11203-11215.

[23] Kuze A, Suto H, Nakajima M, et al. 2009. Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring[J]. *Appl. Opt*., 48 (35):6716-6733.

[24] Yoshida, Y.; Ota, Y.; Eguchi, N.; Kikuchi, N.; Nobuta, K.; Tran, H.; Morino, I.; Yokota, T. Retrieval algorithm for CO2 and CH4 column abundances from short-wavelength infrared spectral observations by the Greenhouse gases observing satellite. Atmos. Meas. Tech. 2011, 4, 717–734.

[25] Schepers, D., et al.Methane retrievals from Greenhouse Gases Observing Satellite (GOSAT) shortwave infrared measurements: Performance comparison of proxy and physics retrieval algorithms. Journal of Geophysical Research: Atmospheres 117.D10 (2012).

[26] Bhattacharjee S, Mitra P, Ghosh S K. Spatial interpolation to predict missing attributes in GIS using semantic kriging[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013.

[27] Khaliq M A, Mustafa F, Rehman S U, et al. Spatiotemporal investigation of near-surface CH4 and factors influencing CH4 over South, East, and Southeast Asia[J]. *Science of The Total Environment*, 2024, 922: 171311.

[28] Huang X, Deng Z, Jiang F, et al. Improved consistency of satellite XCO2 retrievals based on machine learning[J]. Geophysical Research Letters, 2024, 51(8): e2023GL107536.

[29] Ai X, Hu C, Yang Y, et al. Quantification of Central and Eastern China's atmospheric CH4 enhancement changes and its contributions based on machine learning approach[J]. Journal of Environmental Sciences, 2024, 138: 236-248.

[30] Mousavi S M, Dinan N M, Ansarifard S, et al. Unveiling the drivers of atmospheric methane variability in Iran: A 20-year exploration using spatiotemporal modeling and machine learning[J]. Environmental Challenges, 2024: 100946.

[31] Wan Y, Chen F, Fan L, et al. Conversion of surface CH4 concentrations from GOSAT satellite observations using XGBoost algorithm[J]. Atmospheric Environment, 2023, 301: 119694.

[32] Watanabe, Hiroshi, et al. "Global mapping of greenhouse gases retrieved from GOSAT Level 2 products by using a kriging method." International Journal of Remote Sensing 36.6 (2015): 1509-1528.

[33] Wang Y, Yuan Q, Li T, et al. Seamless mapping of long-term (2010–2020) daily global XCO 2 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[J]. Earth System Science Data, 2023, 15(8): 3597-3622.

[34] Hua Y, Zhao X, Sun W, et al. Satellite-Based Reconstruction of Atmospheric CO2 Concentration over China Using a Hybrid CNN and Spatiotemporal Kriging Model[J]. Remote Sensing, 2024, 16(13): 2433.

[35] Nguyen H, Katzfuss M, Cressie N, et al. Spatio-temporal data fusion for very large remote sensing datasets[J]. Technometrics, 2014, 56(2): 174-185.

[36] Huang X, Deng Z, Jiang F, et al. Improved consistency of satellite XCO2 retrievals based on machine learning[J]. Geophysical Research Letters, 2024, 51(8): e2023GL107536.

[37] He Q, Ye T, Chen X, et al. Full-coverage mapping high-resolution atmospheric CO2 concentrations in China from 2015 to 2020: Spatiotemporal variations and coupled trends with particulate pollution[J]. Journal of Cleaner Production, 2023, 428: 139290.

[38] Huang C, Hu J, Xue T, et al. High-resolution spatiotemporal modeling for ambient PM2. 5 exposure assessment in China from 2013 to 2019[J]. Environmental Science & Technology, 2021, 55(3): 2152-2162.

[39] Li L, Chen Y, Fan L, et al. A high-precision retrieval method for methane vertical profiles based on dual-band spectral data from the GOSAT satellite[J]. Atmospheric Environment, 2024, 317: 120183.

[40] Tao C, Peng Y, Zhang Q, et al. Diagnosing ozone–NO x–VOC–aerosol sensitivity and uncovering causes of urban–nonurban discrepancies in Shandong, China, using transformer-based estimations[J]. Atmospheric Chemistry and Physics, 2024, 24(7): 4177-4192.

[41] Wei J, Li Z, Chen X, et al. Separating daily 1 km PM2. 5 inorganic chemical composition in China since 2000 via deep learning integrating ground, satellite, and model data[J]. Environmental science & technology, 2023, 57(46): 18282-18295.

[42] Bovensmann H, Burrows J P, Buchwitz M, et al. SCIAMACHY: Mission objectives and measurement modes[J]. Journal of the atmospheric sciences, 1999, 56(2): 127-150.

[43] Frankenberg C, Aben I, Bergamaschi P, et al. Global column-averaged methane mixing ratios from 2003 to 2009 as derived from SCIAMACHY: Trends and variability[J]. Journal of Geophysical Research: Atmospheres, 2011, 116(D4).

[44] Parker, R. J., Webb, A., Boesch, H., Somkuti, P., Barrio Guillo, R., Di Noia, A., Kalaitzi, N., Anand, J. S., Bergamaschi, P., Chevallier, F., Palmer, P. I., Feng, L., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Hase, F., Kivi, R., Morino, I., Notholt, J., Oh, Y.-S., Ohyama, H., Petri, C., Pollard, D. F., Roehl, C., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D.: A decade of GOSAT Proxy satellite CH4 observations, Earth Syst.

Authorized licensed use limited to: University of Maryland College Park. Downloaded on August 02,2025 at 14:41:43 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3593486

Sci. Data, 12, 3383–3412, https://doi.org/10.5194/essd-12-3383-2020, 2020.

[45] Agustí-Panareda, A., Barré, J., Massart, S., Inness, A., Aben, I.,Ades, M., Baier, B. C., Balsamo, G., Borsdorff, T., Bousserez, N., Boussetta, S., Buchwitz, M., Cantarello, L., Crevoisier, C., Engelen, R., Eskes, H., Flemming, J., Garrigues, S., Hasekamp, O., Huijnen, V., Jones, L., Kipling, Z., Langerock, B., Mc- Norton, J., Meilhac, N., Noël, S., Parrington, M., Peuch, V.-H., Ramonet, M., Razinger, M., Reuter, M., Ribas, R., Suttie,M., Sweeney, C., Tarniewicz, J., and Wu, L.: Technical note:The CAMS greenhouse gas reanalysis from 2003 to 2020, At-mos. Chem. Phys., 23, 3829–3859,2023.

[46] Huang N, Gu L, Black T A, et al. Remote sensing‐based estimation of annual soil respiration at two contrasting forest sites[J]. Journal of Geophysical Research: Biogeosciences, 2015, 120(11): 2306-2325.

[47] Jarvis, A., Reuter, H., Nelson, A., & Guevara, E. (2008). Hole-filled seamless SRTM data V4 [Dataset]. International Centre for Tropical Agriculture (CIAT). Retrieved from http://srtm.csi.cgiar.org.

[48] Tian W, Zhang L, Yu T, et al. Using multisource data and time series features to construct a global terrestrial CO 2 coverage by deep learning[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024.

[49] Mustafa F, Xu M. Improving Model-Based Carbon Dioxide Datasets Using Deep Learning and Satellite Observations[J]. IEEE Transactions on Geoscience and Remote Sensing, 2025.

[50] Wei J, Wang Z, Li Z, et al. Global aerosol retrieval over land from Landsat imagery integrating Transformer and Google Earth Engine[J]. Remote Sensing of Environment, 2024, 315: 114404.

[51] Wei J, Li Z, Chen X, et al. Separating daily 1 km PM2. 5 inorganic chemical composition in China since 2000 via deep learning integrating ground, satellite, and model data[J]. Environmental science & technology, 2023, 57(46): 18282-18295.

[52] Cheng F, Li Z, Yang Z, et al. First retrieval of 24-hourly 1-km-resolution gapless surface ozone (O3) from space in China using artificial intelligence: Diurnal variations and implications for air quality and phytotoxicity[J]. Remote Sensing of Environment, 2025, 316: 114482.

[53] Lundberg S M, Lee S I. A unified approach to interpreting model predictions[J]. Advances in neural information processing systems, 2017, 30.

[54] Van Beek C L, Meerburg B G, Schils R L M, et al. Feeding the world's increasing population while limiting climate change impacts: linking N2O and CH4 emissions from agriculture to population growth[J]. environmental science & policy, 2010, 13(2): 89-96.

[55] Wang J, Azam W. Natural resource scarcity, fossil fuel energy consumption, and total greenhouse gas emissions in top emitting countries[J]. Geoscience Frontiers, 2024, 15(2): 101757.

[56] Hayashida S, Ono A, Yoshizaki S, et al. Methane concentrations over Monsoon Asia as observed by SCIAMACHY: Signals of methane emission from rice cultivation[J]. Remote sensing of environment, 2013, 139: 246-256.

[57] Zeng Z C, Byrne B, Gong F Y, et al. Correlation between paddy rice growth and satellite-observed methane column abundance does not imply causation[J]. Nature communications, 2021, 12(1):

[58] Qu Z, Jacob D J, Zhang Y, et al. Attribution of the 2020 surge in atmospheric methane by inverse analysis of GOSAT observations[J]. Environmental Research Letters, 2022, 17(9): 094003.

[59] Peng S, Lin X, Thompson R L, et al. Wetland emission and atmospheric sink changes explain methane growth in 2020[J]. Nature, 2022, 612(7940): 477-482.

[60] Cui S, Liu P, Guo H, et al. Wetland hydrological dynamics and methane emissions[J]. Communications Earth & Environment, 2024, 5(1): 470.

[61] Peters C N, Bennartz R, Hornberger G M. Satellite‐derived methane emissions from inundation in Bangladesh[J]. Journal of Geophysical Research: Biogeosciences, 2017, 122(5): 1137-1155.

[62] Belikov D A, Patra P K, Terao Y, et al. Assessment of the impact of observations at Nainital (India) and Comilla (Bangladesh) on the CH4 flux inversion[J]. Progress in Earth and Planetary Science, 2024, 11(1): 36.

[63] Qu, Z.; Jacob, D.J.; Bloom, A.A.; Worden, J.R.; Parker, R.J.; Boesch, H. Inverse modeling of 2010–2022 satellite observations shows that inundation of the wet tropics drove the 2020–2022 methane surge. Proc. Natl. Acad. Sci. USA 2024, 121, e2402730121.

[64] Kavitha M, Nair P R. Region-dependent seasonal pattern of methane over Indian region as observed by SCIAMACHY[J]. Atmospheric environment, 2016, 131: 316-325.

[65] Tian W, Zhang L, Yu T, et al. Generating Daily High-Resolution Regional XCO 2 by Deep Neural Network and Multisource Data[J]. Atmosphere, 2024, 15(8).

[66] Zhu S, Feng L, Liu Y, et al. Decadal Methane Emission Trend Inferred from Proxy GOSAT XCH_4 Retrievals: Impacts of Transport Model Spatial Resolution[J]. Advances in Atmospheric Sciences, 2022, 39(8): 1343-1359.

[67] Hong X, Zhang C, Tian Y, et al. Quantification and evaluation of atmospheric emissions from crop residue burning constrained by satellite observations in China during 2016–2020[J]. Science of the Total Environment, 2023, 865: 161237.

[68] Lu X, Jacob D J, Zhang Y, et al. Global methane budget and trend, 2010–2017: complementarity of inverse analyses using in situ (GLOBALVIEWplus CH 4 ObsPack) and satellite (GOSAT) observations[J]. Atmospheric Chemistry and Physics, 2021, 21(6): 4637-4657.

[69] Bloom A A, Bowman K W, Lee M, et al. A global wetland methane emissions and uncertainty dataset for atmospheric chemical transport models (WetCHARTs version 1.0) [J]. Geoscientific Model Development, 2017, 10(6): 2141-2156.

[70] Saunois M, Stavert A R, Poulter B, et al. The global methane budget 2000–2017[J]. Earth System Science Data Discussions, 2019, 2019: 1-136.

[71] Parker R J, Webb A, Boesch H, et al. A decade of GOSAT Proxy satellite CH 4 observations[J]. Earth System Science Data, 2020, 12(4): 3383-3412.

[72] Li L, Lei L, Song H, et al. Spatiotemporal geostatistical analysis and global mapping of ch4 columns from gosat observations[J]. *Remote Sensing*, 2022, 14(3): 654.

[73] Lopez F P A, Zhou G, Jing G, et al. XCO_2 and XCH_4 reconstruction using GOSAT satellite data based on EOF-algorithm[J]. Remote Sensing, 2022, 14(11): 2622.

Yu Qu. received his Master's degree in Geomatics Engineering from Shandong University of Science and Technology, Qingdao, China. He is currently pursuing a Ph.D. degree in the School of Geography at South China Normal University (SCNU), Guangzhou, China. His research focuses on satellite data processing, multimodal data fusion, and methane emission inversion.

Jing Wei. (Member, IEEE) received his Ph.D. degree from Beijing Normal University, Beijing, China, in 2020 through a joint doctoral program with the University of Maryland, College Park, MD, USA. He is currently an Assistant Research Scientist at the University of Maryland, and will join Peking University in Beijing, China, as an Assistant Professor. He has authored over 200 articles, and his research focuses on aerosols, particulate matter, chemical composition, and trace gases, leveraging satellite remote sensing, big data, and artificial intelligence. Dr. Wei serves as a Topical Associate Editor for *IEEE Transactions on Geoscience and Remote Sensing*, an Associate Editor for the *Journal of Geophysical Research: Atmospheres*, and an Editor for *Earth System Science Data*.

Authorized licensed use limited to: University of Maryland College Park. Downloaded on August 02,2025 at 14:41:43 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3593486

Hanfa Xing. The author received his Ph.D. degree in Cartography from Central South University, China, in 2012. He is currently working with the School of Geography, Beidou Research Institute, South China Normal University (SCNU), Guangzhou, China. Research interests include remote sensing, land cover, urban landscape, land use analysis and Geographic spatiotemporal big data. Developed a theoretical framework for the collection and expression of geographic big data for domain applications, solving the problem of multisource heterogeneous spatiotemporal big data application in the field of land cover.

Xian Shi received the Master's degree in Geomatics Engineering from Shandong University of Science and Technology, Qingdao, China. She is currently pursuing a Ph.D. degree in Geodesy and Geomatics at Shandong University of Science and Technology. Her research interests include synthetic aperture radar (SAR) interferometry and its applications in forest parameter inversion.

Zurui Ao received the M.S. and Ph.D. degrees in cartography and geographical information systems from Capital Normal University, Beijing, China, in 2014 and 2018, respectively. He was a Post-Doctoral Researcher at Sun Yat-sen University, Guangzhou, China, from 2018 to 2021. He is currently an Associate Researcher with the School of Geography, Beidou Research Institute, South China Normal University, Guangzhou. His research interests include light detection and ranging (LiDAR), deep learning in remote sensing images, and spatiotemporal fusion.

Xiaoliang Meng. The author director of Wuhan University Professor of School of remote sensing and information engineering, Luojia Distinguished Professor. he has presided over a number of national natural science foundation, national key research and development plans, national major projects and other vertical and horizontal scientific research projects; he has won the International Society for Photogrammetry and Remote Sensing (ISPRS) Best Young Author Award; his research direction is intelligent space perception network.

Authorized licensed use limited to: University of Maryland College Park. Downloaded on August 02,2025 at 14:41:43 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.