https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Improved Ensemble-Based Satellite-Derived XCO_2 and XCH_4 Climate Data Records Using EMMA v5.1

Blanca Fuentes Andrade[1], Maximilian Reuter[1], Michael Buchwitz[1], Heinrich Bovensmann[1], Stefan Noël[1], Oliver Schneising[1], Michael Hilker[1], Robert J. Parker[2,3], Andrew Gerald Barr[4], Yukio Yoshida[5], Yu Someya[5], Hartmut Bösch[1], and John P. Burrows[1]

[1]Institute of Environmental Physics (IUP), University of Bremen, Bremen, Germany
[2]National Centre for Earth Observation, Space Park Leicester, University of Leicester, Leicester, UK
[3]Earth Observation Science, School of Physics and Astronomy, University of Leicester, Leicester, UK
[4]Earth Science Group, SRON Space Research Organisation Netherlands, CA Leiden, the Netherlands
[5]National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan

Correspondence: Blanca Fuentes Andrade (bfuentes@iup.physik.uni-bremen.de)

Abstract.

Satellite observations of the column-averaged dry-air mole fractions of carbon dioxide (XCO_2) and methane (XCH_4) are essential for monitoring their global atmospheric distribution and variability, as well as for related climate applications. To generate long-term, consistent Climate Data Records (CDR), ensemble merging approaches such as the ensemble median algorithm (EMMA) have been used.

Here, we introduce a new version of the EMMA algorithm, v5.1, and the four merged datasets that were created with it, namely merged Level 2 (L2) and corresponding gridded Level 3 (L3) products for both XCO_2 and XCH_4. EMMA v5.1 keeps the core idea from previous EMMA versions of the ensemble median being a good choice for merging multiple satellite data products, while taking a new approach: for each gas and time period, one of the L2 input products is used as a reference, and the ensemble is used to compute an empirical spatio-temporal adjustment for this input dataset. This new design eliminates conceptual limitations of former EMMA versions, making the L2 dataset eligible for hotspots emission estimations in periods in which the reference product has sufficient spatial coverage. Additionally, new and updated input L2 products were included in the ensemble, and the L3 datasets have higher spatial resolution, i.e. monthly 1°x1°, compared to the previous monthly 5°x 5° resolution. The four merged datasets cover the time period from 2003 to 2024.

All four merged datasets were validated against ground-based Total Carbon Column Observing Network (TCCON) retrievals, showing good agreement. The XCO_2 L2 (L3) merged datasets have a spatio-temporal bias of 0.47 ppm (0.48 ppm). The estimated single-sounding (single-cell) random error is 1.51 ppm (1.82 ppm), very close to the average reported uncertainty of 1.66 ppm (1.65 ppm). The estimated linear drift is 0.00 ± 0.12 ppm y^{-1} (0.00 ± 0.11 ppm y^{-1}).

The XCH_4 L2 (L3) merged datasets have a spatio-temporal bias of 7.37 ppb (4.31 ppb). The estimated single-sounding (single-cell) random error is 20.87 ppb (9.01 ppb), very close to the average reported uncertainty of 21.11 ppb (11.05 ppb). The estimated linear drift is 0.42 ± 1.14 ppb y^{-1} (0.11 ± 0.62 ppb y^{-1}).

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

1 Introduction

Carbon dioxide (CO_2) and methane (CH_4) are the most important anthropogenic greenhouse gases, whose cumulative emissions significantly contribute to global warming (e.g., Gulev et al., 2021). Due to their importance for the Earth's radiative balance, the atmospheric concentrations of CO_2 and CH_4 are considered as Essential Climate Variables (ECV) by the Global Climate Observing System (GCOS) (e.g., GCOS-154, 2011; GCOS-200, 2016). Monitoring atmospheric concentrations of these two greenhouse gases both long term and globally is essential to determine their trends over time, as well as their sources and sinks. For the global monitoring of these two ECVs, ground-based measuring networks and satellite observations are used. Ground-based networks provide high-precision measurements at specific locations, having sparse spatial coverage, while satellite measurements offer potentially global coverage.

Column-averaged dry-air mole fractions of CO_2 and CH_4, denoted as XCO_2 and XCH_4, respectively, and hereafter collectively referred to as XGHG, from space-borne instruments were first retrieved using measurements of the SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric ChartographY; Burrows et al. (1995), Bovensmann et al. (1999)) on ESA's ENVISAT satellite (2002-2012). Since 2002, multiple space-borne sensors have become operational (as shown in Fig. 1). These sensors have different lifetimes and spatio-temporal sampling patterns. Additionally, there are several independently developed algorithms to retrieve XCO_2 and XCH_4 from the radiance spectra of reflected solar radiation measured by these sensors, leading to several Level 2 (L2) products. Climate Data Records (CDRs) are long-term, consistent, and high-quality datasets suitable for the evaluation of trends and variability of the ECVs. To produce CDRs with the largest temporal coverage from existing L2 products, they need to be combined into a consistent merged multi-sensor and/or multi-algorithm product.

A multi-sensor CH_4 dataset was developed by Balasus et al. (2023) with the aim to correct systematic biases in TROPOspheric Monitoring Instrument (TROPOMI) L2 data using Greenhouse gases Observing SATellite (GOSAT) XCH_4 observations. The blended TROPOMI+GOSAT product, generated using a machine learning model that predicts TROPOMI-GOSAT differences based on a set of predictor variables from the TROPOMI product, benefits from the high observation density of TROPOMI and the high accuracy of GOSAT, while being bound to the availability of observations from two sensors and the used retrieval algorithms.

The individual L2 input products are typically validated against Total Carbon Column Observing Network (TCCON) retrievals and show good agreement with the TCCON measurements (Wunch et al., 2010, 2011, 2015; Laughner et al., 2024). Further away from the TCCON stations, the deviation among the XCO_2 and XCH_4 values retrieved with the different algorithms tends to be higher (Reuter et al., 2013, 2020). To create a merged product that is robust against such deviations and potentially reduces bias patterns from individual input products, Reuter et al. (2013) developed the ensemble median algorithm (EMMA), a merging algorithm based on the assumption that the ensemble median from the Level 3 (L3) data is a good representative of the true atmospheric state. This approach has been used since 2016 to operationally generate merged XCO_2 and XCH_4 datasets within the Copernicus Climate Change Service (C3S). The corresponding products are available through the Copernicus Climate Data Store (CDS, https://cds.climate.copernicus.eu/). The merged datasets generated with the EMMA

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.



algorithm have been used for comparison with climate models (Lauer et al., 2017; Gier et al., 2020) as well as for estimation of annual mean growth rates (e.g., Buchwitz et al., 2018; Reuter et al., 2020; C3S and WMO, 2025).

The EMMA versions up to v4.6, used to create the merged XCO_2 and XCH_4 multi-sensor and multi-algorithm datasets as described by Reuter et al. (2013, 2020), use the following procedure. First, all L2 products are averaged onto a monthly 10°x 10° latitude-longitude grid. For each grid cell, the EMMA algorithm identifies the L2 product whose gridded value corresponds to the ensemble median of all available input products. The merged L2 product is created by transferring, for each grid cell, all individual L2 soundings corresponding to the product selected as the ensemble median. Finally, the merged L2 product is aggregated to create L3 XCO_2 and XCH_4 products, formatted according to the Observations for Model Intercomparisons Project (Obs4MIPs; e.g., Waliser et al., 2020) conventions, with monthly 5°x 5° resolution.

The L2 data sets generated using EMMA up to its 4.6 version were not suitable or optimized for applications that require dense spatial sampling without jumps, such as obtaining emission information on hot spots. Due to the design of the algorithm, especially in terms of sampling, the L2 product may have discontinuities at the monthly 10°x 10° boundaries, leading to possible cut-offs or inconsistencies at these boundaries. Additionally, regions with dense spatial sampling are not prioritized by the merging algorithm, leading to the potential override of L2 data that would be suitable for emission estimation by sparser data from another sensor.

In EMMA v5.1, we have taken a new approach that also assumes that the ensemble median is a good choice, analogous to previous EMMA versions. However, the new approach, in contrast to former EMMA versions, is to select a reference L2 product per time period, while the ensemble of input L2 products is used for the computation of an empirical spatio-temporal adjustment for the reference product. With this new design, the potential jumps at the monthly 10°x 10° boundaries are eliminated, and the new merged L2 product has the spatial resolution and sampling density of the selected reference products. The merged L2 product is again gridded to create L3 Obs4MIPs products at higher resolution, i.e. monthly 1°x 1°, compared to the previous monthly 5°x 5° resolution. Additional improvements of the EMMA v5.1 products with respect to its predecessors is the inclusion in the ensemble of data from more sensors, as well as updated versions of the input datasets.

The input datasets used for the generation of the merged XGHG L2 and L3 datasets is described in Sect. 2. Section 3 describes the merging algorithm (Sect. 3.1) and the validation method (Sect. 3.2). The resulting XGHG L2 and L3 are described in Sect. 4 and the conclusions are presented in Sect. 5.

2 Datasets

2.1 Satellite data

To generate the merged XCO_2 and XCH_4, we used data from six instruments, as illustrated in Fig. 1. These instruments are SCIAMACHY (SCanning Imaging Absorption spectroMeter of Atmospheric CHartographY; Burrows et al., 1995; Bovensmann et al., 1999) on ENVISAT, TANSO-FTS on GOSAT (Kuze et al., 2009, 2016; Yokota et al., 2009) and TANSO-FTS-2 on GOSAT-2 (Greenhouse gases Observing SATellite 2, Suto et al., 2021; Imasu et al., 2023), which permit retrieval of both XCO_2 and XCH_4; OCO-2 (Orbiting Carbon Observatory-2, Crisp et al., 2017), and OCO-3 (Orbiting Carbon Observatory-3,

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Eldering et al., 2019) on the International Space Station (ISS), which permit retrieval of XCO_2; and TROPOMI (Veefkind et al., 2012) on Sentinel-5 Precursor (S5P), which permits retrieval of XCH_4. All these instruments measure radiance spectra of backscattered solar radiation containing absorption bands of CO_2 and CH_4, which the retrievals interpret to estimate atmospheric XCO_2 and XCH_4 values. There are several independently developed algorithms to retrieve XCO_2 and XCH_4 from the spectra measured by these sensors. An overview of the L2 input products used to generate the merged datasets is given in Table 1 for XCO_2 and 2 for XCH_4.

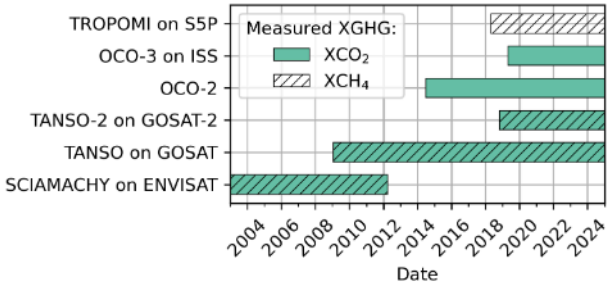


Figure 1. Overview of the satellite sensors used to generate the merged XGHG products, together with the time coverage and XGHG (i.e. XCO_2 and/or XCH_4) retrieval that each sensor permits.

Table 1. Satellite-derived L2 input data products used to generate the merged XCO_2 product. The table lists the satellite instrument, the name and version of the L2 algorithm, the institution and related references.

Satellite instrument	Algorithm and version number	Institution	Reference
SCIAMACHY	BESD v02.01.02	IUP	Reuter et al. (2010, 2011, 2016)
GOSAT	NIES v03.05bc	NIES	Yoshida et al. (2013, 2023); Someya et al. (2023)
GOSAT	RemoTeC v2.3.8	SRON	Butz et al. (2011); Detmers (2017a); Wu (2021)
GOSAT	UoL-FP v7.3	UoL	Cogan et al. (2012); Boesch and Di Noia (2021)
GOSAT	ACOS v9r	NASA	O'Dell et al. (2012); Taylor et al. (2022)
GOSAT	FOCAL v4.0	IUP	Noël et al. (2022); Appendix B[a]
OCO-2	ACOS v11.2	NASA	Crisp et al. (2021); Taylor et al. (2023)
OCO-2	FOCAL v12	IUP	Reuter et al. (2017b, a, 2021)
GOSAT-2	NIES v02.20 & v02.21	NIES	Yoshida et al. (2023); Yoshida and Oshio (2025)
GOSAT-2	RemoTeC v2.1.0	SRON	Barr et al. (2025)
GOSAT-2	FOCAL v4.0	IUP	Noël et al. (2022); Appendix B[a]
OCO-3	ACOS v11r	NASA	Crisp et al. (2021); Taylor et al. (2023)

[a] Appendix B contains the changes of GOSAT and GOSAT-2 FOCAL v4.0 with respect to Noël et al. (2022).

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 2. Satellite-derived L2 input data products used to generate the merged XCH_4 product. The table lists the satellite instrument, the name and version of the L2 algorithm, the institution and related references.

Satellite instrument	Algorithm and version number	Institution	Reference
SCIAMACHY	WFMD v4.0	IUP	Schneising et al. (2011); ESA CCI GHG project team (2018)
GOSAT	FOCAL-FP v4.0	IUP	Noël et al. (2022); Appendix B [a]
GOSAT	FOCAL-PR v4.0	IUP	Noël et al. (2022); Appendix B[a]
GOSAT	NIES v03.05bc	NIES	Yoshida et al. (2013, 2023); Someya et al. (2023)
GOSAT	RemoTeC-FP v2.3.8	SRON	Butz et al. (2011); Detmers (2017a); Wu (2021)
GOSAT	RemoTeC-PR v2.3.9	SRON	Butz et al. (2011); Detmers (2017b)
GOSAT	UoL-FP v7.3	UoL	Cogan et al. (2012); Boesch and Di Noia (2021)
GOSAT	UoL-PR v9.0	UoL	Parker et al. (2011, 2020)
GOSAT-2	FOCAL-FP v4.0	IUP	Noël et al. (2022), Appendix B[a]
GOSAT-2	FOCAL-PR v4.0	IUP	Noël et al. (2022), Appendix B[a]
GOSAT-2	RemoTeC-FP v2.1.0	SRON	Barr et al. (2025)
GOSAT-2	RemoTeC-PR v2.1.1	SRON	Barr et al. (2025)
GOSAT-2	NIES v02.20 & v02.21	NIES	Yoshida et al. (2023); Yoshida and Oshio (2025)
GOSAT-2	UoL_PR v1.0	UoL	Parker et al. (2011, 2025)
TROPOMI	WFMD v2.0	IUP	Schneising et al. (2026)

[a] Appendix B contains the changes of GOSAT and GOSAT-2 FOCAL v4.0 with respect to Noël et al. (2022).

2.2 SLIM data

We use the Simple cLImatological Model (SLIM) for atmospheric CO_2 and CH_4 (Noël et al., 2022) in its 2025 version as common a priori for harmonizing the different L2 input data sets (described in Sect. 3.1.1).

SLIM CO_2 and CH_4 use model-based climatologies to estimate the spatial distribution and annual cycle of CO_2 and CH_4, respectively. SLIM2025 CO_2 uses CO_2 mole fraction information from CAMS global CO_2 atmospheric inversion v22r1 (CAMS, 2020; Chevallier et al., 2005, 2010; Chevallier, 2013). SLIM2025 CH_4 has been derived from CAMS global CH_4 atmospheric inversion v22r2 (CAMS, 2020; Segers, 2022). Additionally, SLIM adjusts the used climatologies for the annual growth using growth rates from NOAA (https://gml.noaa.gov/ccgg/trends/gr.htm, last accessed: 12/08/2025). If the used NOAA annual increases had not yet stabilised at the time of extraction, subsequent revisions reveal a corresponding bias in the SLIM dataset as originally constructed. Such revisions of the annual increases mainly affect CH_4 and, if they occur, are most likely towards the end of the time series. Both SLIM2025 CO_2 and SLIM2025 CH_4 use 20 height layers, defined in such a way that each layer contains the same number of dry-air molecules. SLIM CO_2 and CH_4 empirically extrapolate from a climatology of past averaged modeled CO_2 and CH_4 fields, so new or changing phenomena are not reflected in SLIM CO_2 and CH_4.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

2.3 Validation data

Column-averaged dry-air mole fractions of CO_2 and CH_4 from the Total Carbon Column Observing Network (TCCON, http://www.tccon.caltech.edu/) in its GGG2020 version (Laughner et al., 2024) were used for calibration of the uncertainties of the different L2 input products and the validation of the final merged product (described in Sects. 3.1.2 and 3.2, respectively).

TCCON is a network of ground-based Fourier transform spectrometers that measure solar spectra in the infrared region, and report column-average mole fractions of CO_2, CH_4 and other atmospheric species. The TCCON project started in 2004 and the number of operational sites has increased over time. TCCON is the core network used for validation of the satellite XCO_2 and XCH_4 retrieval (e.g., Wunch et al., 2011, 2017). The TCCON Network provides temporally dense observations, but spatially sparse.

An overview of the used TCCON sites is shown in Table A1.

3 Methods

In this section, the EMMA algorithm in its v5.1, used to generate the XCO_2 and XCH_4 multi-sensor and multi-algorithm datasets, is described (Sect. 3.1), together with the validation procedure (Sect. 3.2).

3.1 Merging algorithm

The merging approach of EMMA v5.1. is based on the same idea as previous EMMA versions (Reuter et al., 2013, 2020) that the ensemble median is a good choice to represent the true atmospheric state, but it has a new design. EMMA v5.1. selects a reference L2 product per time period, and uses the ensemble of input L2 products for an empirical spatio-temporal adjustment of the reference products. The input datasets, along with their respective temporal availability, are shown in Fig. 2 and Fig. 3, where the algorithms used as a reference are also illustrated. For the merged XCO_2 datasets, the reference products are SCIAMACHY BESD from the beginning of the dataset on the 1 January 2003 until the 31 March 2009; GOSAT FOCAL from the 1 April 2009 until the 31 July 2014 and OCO-2 FOCAL from the 1 September 2014 until the end of the dataset, on 31 December 2024. For the merged XCH_4 datasets, the reference products are SCIAMACHY WFMD from the beginning of the dataset on the 1 January 2003 until the 31 March 2009; GOSAT FOCAL PR from the 1 April 2009 until the 30 April 2018; and TROPOMI WFMD from the 1 May 2018 until the end of the dataset, on 31 December 2024. We would expect other choice of high-quality reference products with similar coverage, sampling density and precision to lead to comparable results.

The generation of the merged products with EMMA v5.1 is conceptually similar to previous versions of the EMMA algorithm (Reuter et al., 2013, 2020, 2025a). The steps of the algorithm are illustrated in the flowchart in Fig. 4 and explained in Sects. 3.1.1 to 3.1.7. The first four steps (see Sects. 3.1.1 to 3.1.4) follow the same approach as Reuter et al. (2013, 2020) and therefore are only outlined here for completion. The input datasets are the L2 products (described in Sect. 2.1), TCCON GGG2020 data (see Sect. 2.3) and SLIM2025 (see Sect. 2.2).

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

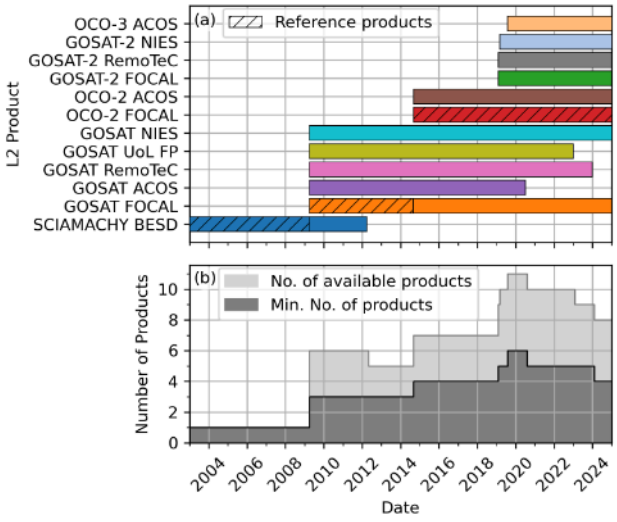


Figure 2. EMMA v5.1 input nominal data availability for XCO_2. Panel (a) shows the nominal temporal availability for the L2 input products The hatched bars illustrate the time periods that each of the selected L2 products is used as the reference product. Panel (b) shows the nominal total number of products (light gray) and the minimum number of algorithms required (dark gray) for the median calculation.

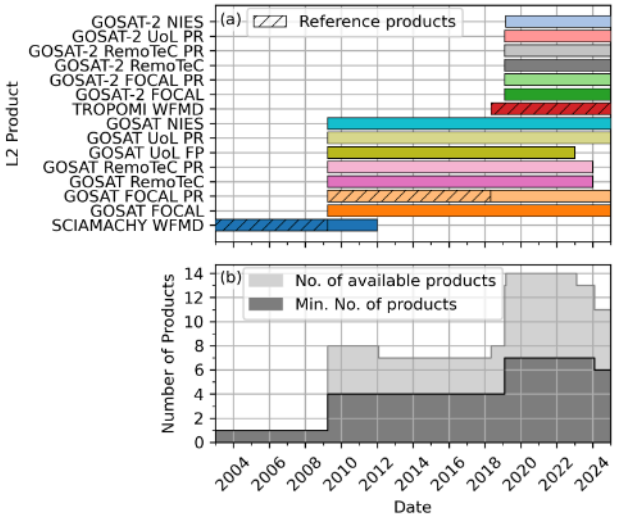


Figure 3. EMMA v5.1. input nominal data availability for XCH_4. Panel (a) shows the nominal temporal availability for the L2 input products The hatched bars illustrate the time periods that each of the selected L2 products is used as the reference product. Panel (b) shows the nominal total number of products (light gray) and the minimum number of algorithms required (dark gray) for the median calculation.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

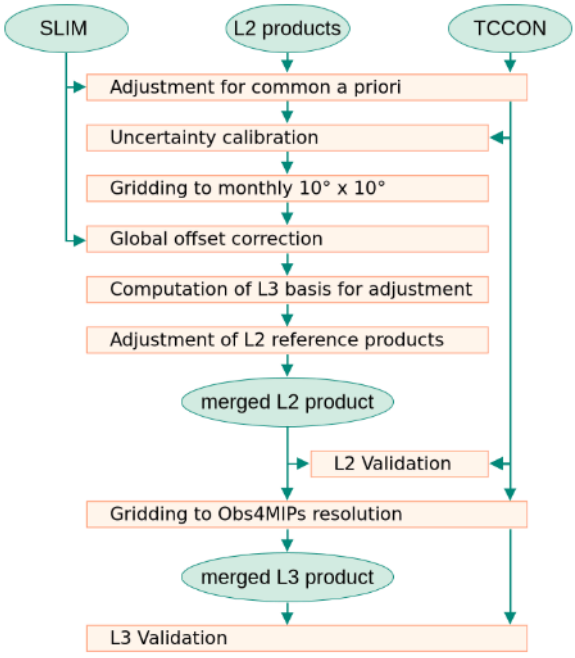


Figure 4. Flowchart showing the main steps of the EMMA v5.1 algorithm used to generate the merged L2 EMMA product from the individual satellite-sensor L2 data products. The input datasets are shown in green ovals; the light-brown boxes indicate each step the datasets go through. The resulting merged L2 product is then gridded to generate the L3 Obs4MIPs format data product.

3.1.1 Adjustment for common a priori

To account for different column a priori profiles, all satellite L2 input data sets, as well as the TCCON data, are adjusted to a common a priori, for which SLIM2025 CO_2 and CH_4 was used. We do this as proposed by Wunch et al. (2011) and as described in Rodgers (2008) or Reuter et al. (2024). The resulting adjustment is typically very small, especially for CO_2, with adjustments on the order of a few tenths of a ppm.

Additionally, we interpolate the column averaging kernels and a priori profiles of the L2 input data sets so that the atmospheric column is split into ten layers defined by an equidistant pressure grid starting at the surface pressure and ending at zero.

3.1.2 Uncertainty calibration

The uncertainty of the individual L2 soundings is calibrated by a scaling factor obtained from their validation against TCCON. This scaling factor is such that the average reported uncertainty is consistent with the standard deviation of satellite minus TCCON ground-based validation data differences. The used scaling factors are listed in the last column of Table 3 for XCO_2 and 4 for XCH_4. The details on how this validation was performed are explained in Sect. 3.2

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

3.1.3 Initial gridding to monthly 10°x 10°

The different L2 input products have different samplings, since they come from different satellites, different algorithms have different filtering approaches, etc. Therefore, the product inter-comparison in the EMMA algorithm is based on aggregated data (i.e. L3), namely monthly averages on a 10°x 10° grid.

For each monthly 10°x 10° grid cell, we compute the standard error of the mean (SEM) as:

$$SEM = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\sigma_i^2}, \quad (1)$$

where σ_i is the scaled (according to Sect. 3.1.2) XCO_2 or XCH_4 uncertainty of the i-th sounding, and n is the total number of soundings in the grid cell. To get statistically robust results, we only use, for each L2 product, those grid boxes with more than five soundings and for which the standard error of the mean (SEM) is estimated to be less than 1 ppm for XCO_2 and 12 ppb for XCH_4.

3.1.4 Global offset correction

We correct for potential offsets between the individual data products by adding a constant global offset for each L2 product. The global offset applied to each product is computed using SLIM as a reference. SLIM reproduces basic features well as it is constructed from state-of-the-art models and the NOAA annual growth rates. However, other models would also work as reference.

The global offset for each product with respect to SLIM is computed by least squares minimization of pairwise differences, i.e. the difference based on all overlapping L3 data is optimized. Only grid boxes with L3 data for multiple products are taken into account. More specifically, we used grid boxes with L3 data for at least as many products as shown by the dark gray areas in panel b of Fig. 2 (for XCO_2) and Fig. 3 (for XCH_4), and with L3 data for more than one product. This minimum number of required products used as a threshold is computed by rounding half the nominal total number of available products at each time (light gray areas of Fig. 2 and 3, panel b) up to the next integer.

The overall effect of the global offset correction is illustrated in the figures shown in Appendix C.

3.1.5 Computation of the L3 basis for the empirical spatio-temporal adjustment

From the previous steps, we obtain a field of a-priori-adjusted uncertainty-calibrated gridded and offset-corrected XGHG data from each of the input L2 products. Let us refer to this set of L3 XGHG fields as X_{pijk}, where the p index indicates the input L2 product and i, j, k each grid cell (latitude, longitude, time).

Following the same conceptual approach as Reuter et al. (2013, 2020), we use X_{pijk} to compute a median field, M_{ijk}, as the median of X along the p dimension. The median is computed only for those L3 grid boxes with valid data (according to the filtering described in Sect. 3.1.3) from at least the minimum number of data products shown in Figs. 2 and 3 (panel b, dark gray areas). For those grid cells with an even number of input products with valid L3 data, we define the median as the mean of the two central values. For the sake of simplicity, from now on we will drop the indices i, j, k that refer to each grid box.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Examples of the gridded input products and the computed median are shown in Appendix D.

Previous EMMA versions ($\leq$ v4.6) selected, for each month and each 10°x 10° grid cell, exactly one product (the product corresponding to the median) of the available individual L2 input products and then transferred all relevant information, e.g. needed for surface flux inverse modelling, from the selected harmonized and uncertainty calibrated L2 data into the corresponding EMMA L2 files. In the new approach taken for EMMA v5.1, one reference product per time period is selected (see hatched bars in panel a of Figs. 2 and 3). The median of the gridded monthly 10°x 10° data is used to compute an empirical spatio-temporal adjustment which is then applied to the L2 data of each of the reference products.

The empirical spatio-temporal adjustment is computed as follows. For each reference product, a basis for the empirical spatio-temporal adjustment, B, is computed as the difference between the L3 XGHG in each grid cell and the corresponding median value, i.e., $B = X_{\text{ref}} - M$, where X_{ref} contains the L3 XGHG field for the corresponding reference products based on the defined periods. The resulting B field, has some gaps due to a) the absence of L2 observations from the reference products that meet the L3 criteria, which leads to an absence of L3 data at some latitudes, longitudes and times, and b) gaps in the median field resulting from the absence of valid L3 data from enough input products. These gaps could lead to unadjusted L2 observations, so they need to be filled to minimize the amount of unadjusted L2 observations. To fill the gaps in the B field, we applied a normalized convolution using a Gaussian kernel whose weights are modified by thresholding.

A common approach for normalized convolution is to normalize by the smoothed validity mask, v_s. This approach gives reliable interpolated values if the underlying field is smooth and the local data density around each gap is sufficiently high. However, due to the spatial variability of the XGHG fields, large regions with gaps can produce extrapolation artefacts that would lead to unrealistic adjustments. To avoid such artefacts, we defined a new mask for the normalized convolution that treats gaps in the B field surrounded by very low data density and previously filled with zeros, as if they contained valid data. The resulting gap-filled B field contains therefore values of zero in regions with low data density, which will lead to zero-valued spatio-temporal adjustments, i.e. they would leave the reference product unadjusted. The details on this modified normalization for the convolution are explained in Appendix E.

The gap filling is applied to each reference product independently. Examples of the B fields before and after gap filling are shown in Figs. 5 and 6, for XCO_2 and XCH_4 respectively, together with the L3 data and the median they were computed from.

From the ensemble of L3 data at monthly 10°x 10° resolution, the following quantities are computed together with the B field. As for the computation of the B field, only grid boxes with data for more than half the total number number of available products for that time period (see dark gray area in Figs. 2 and 3) were considered.

- The data density mask, which is a boolean mask that flags the gaps in B that were treated by the smoothing procedure as if they contained data.

- The Inter-algorithm spread (IAS) field, which informs about potential regional or temporal systematic uncertainties, and is computed from the median absolute deviation (MAD) as:

$$IAS = \frac{\text{MAD}}{\Phi^{-1}(0.75)}, \tag{2}$$

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

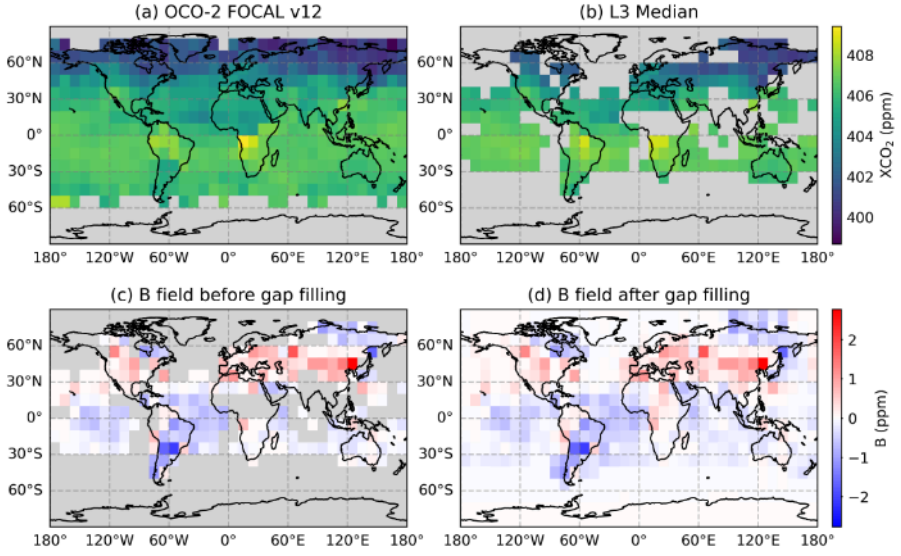


Figure 5. Panel (a) shows the gridded XCO_2 reference product OCO-2 FOCAL at monthly 10°x10° resolution. Panel (b) shows the median, used for the computation of the B field (i.e. the basis for the empirical spatio-temporal adjustment). Panel (c) shows the B field. Panel (d) shows the B field after gap filling. All four panels show global maps for August 2018.

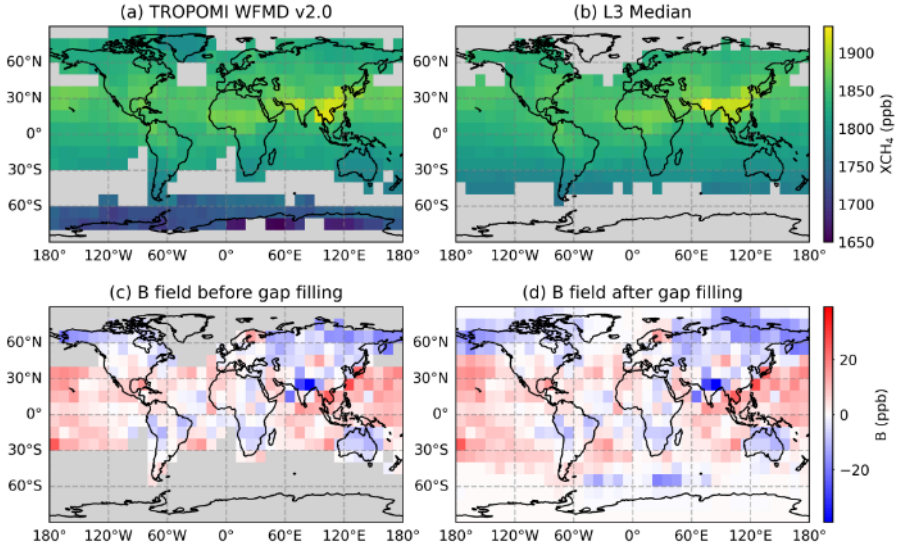


Figure 6. Analogous to Fig. 5, for the gridded XCH_4 reference product TROPOMI WFMD and for the month of September 2019.

where $\Phi^{-1}(0.75)$ is the 75%-percentile of the standard normal probability distribution (Huber and Ronchetti, 2009). The MAD provides a robust estimation of the spread, less influenced by outliers than the standard deviation.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.



– The uncertainty of the B field, σ_B, which was computed, for each grid cell, as the combined uncertainty of the SEM and the IAS:

$$\sigma_B = \sqrt{SEM^2 + IAS^2}. \quad (3)$$

Therefore, analogously to the field B, its uncertainty, σ_B, also has gaps. These gaps were filled applying a normalized convolution with a Gaussian kernel on σ_B^2. Unlike for the gap filling of B, the (unmodified) smoothed validity mask is used for weighting in the gap filling of σ_B. For this reason, the filled σ_B will still have gaps if there is no data in the region of the size of the Gaussian kernel around them.

3.1.6 Adjustment of L2 products and estimation of the corresponding uncertainty

For each L2 sounding of the reference products, the corresponding empirical spatio-temporal adjustment is computed as a trilinear interpolation (in time, latitude and longitude) of the gap-filled B field, and subtracted from the L2 XGHG values.

To obtain the adjustment uncertainty for each XGHG value, the squared gap-filled B-field uncertainty is also interpolated. The outcome of this interpolation contains some gaps (derived from the B field uncertainty; see Sect. 3.1.5), which are filled with the corresponding spatio-temporal bias obtained from the L2 validation and shown in Tables 3 and 4. Additionally, to avoid unrealistically low values of the uncertainty, we set the lower uncertainty bound to be the TCCON accuracy (Wunch et al., 2010), i.e. 0.4 ppm for XCO_2 and 4 ppb for XCH_4. The TCCON data is highly accurate and used as the validation standard for satellite-derived observations, so we do not expect satellite-based estimates to me more accurate.

The L2 merged products contain therefore two components of the XGHG uncertainties: the calibrated 1σ uncertainty of the retrieved values by the corresponding reference product, and the adjustment uncertainty. These two uncertainties should be combined as the root sum of squares to obtain a realistic uncertainty of the adjusted XGHG. It is worth noting that the adjustment uncertainty can be interpreted as an estimate of the accuracy of the XGHG values for the used ensemble, and therefore does not decrease by averaging XGHG soundings.

The data density mask is also tri-linearly interpolated, leading to an estimate of the data density used to compute the empirical spatio-temporal adjustment. The B field gaps surrounded by low or no data density had been filled with zeros (see Sect. 3.1.5), therefore, the empirical spatio-temporal adjustment will be also (close to) zero for L2 soundings in such regions, resulting in (nearly) uncorrected soundings of the reference product. Such soundings are identified by the data density, whose value is 0 if the computed adjustment for a given sounding was zero, and 1 if the computed adjustment was not affected by smoothing with zeros due to missing data in the computed B field. The closer the data density to 0, the fewer grid cells with an ensemble median were used to estimate the corresponding adjustment, so the closer the computed adjustment will be to zero. The data density mask is therefore especially useful in order to ensure that the XGHG values were adjusted by the ensemble.

3.1.7 Gridding to monthly 1°x 1°

To obtain the merged L3 datasets in Obs4MIPs format, the merged L2 EMMA products obtained from the previous step are gridded to spatial 1°x 1° and temporal monthly resolution, in contrast with previous product versions (up to v4.6) with monthly

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

5°x 5° spatial resolution. The gridding is analogous to that in Sect. 3.1.3 and is based on arithmetic unweighted averaging of all soundings falling in a grid box for the adjusted XGHG values, the interpolated data density mask, the column averaging kernel and the a priori profile.

The number of soundings used for averaging, the number of days with L2 data and the standard deviation of the averaged XGHG values is also included in the final L3 product.

For each grid box, we computed the standard error of the mean (SEM) using Eq. 1. In this case, however, due to the high spatial resolution of the L3 product and the fact that many of the L2 observations within one grid cell were taken within a short period of time, we considered the L2 soundings within one day on a 1°x 1° as strongly correlated. Therefore, to compute the SEM, the number of days with XGHG data within a spatial grid cell was used as the effective number of observations within that given grid cell. To reduce noise within the level 3 product and to be consistent with previous versions, we filter out grid boxes whose SEM is larger than 5 ppm for XCO_2 or larger than 60 ppb for XCH_4. Individual measurements have typically lower uncertainties than these thresholds, so this filter only affects regions with low data density and high uncertainty, and is specially relevant in the SCIAMACHY WFMD XCH_4 time period.

Additionally, the total uncertainty is estimated as:

$$\hat{\sigma} = \sqrt{SEM^2 + \sigma_s^2}, \tag{4}$$

where σ_s represents the gridded tri-linearly interpolated adjustment uncertainty, computed as the root mean squared of all soundings falling in a grid box.

The ensemble availability is a boolean mask defined from the data density. The ensemble availability has a value of 1 in the grid cells where the interpolated data density mask is lower than a given threshold, which we chose to be 0.3, and 0 elsewhere (where there is XGHG data). This means that where the ensemble availability is 1, the XGHG values were not (significantly) adjusted by the ensemble, i.e. the adjustment is (close to) zero in these cases.

3.2 Validation

The merged L2 and L3 XGHG EMMA v5.1 products were validated against TCCON GGG2020 observations following the same approach as e.g. Noël et al. (2022); Reuter et al. (2025b). Both the L2 and L3 validation strategies are based on finding collocations for each individual XGHG sounding (L2) or each grid cell (L3) with TCCON data.

The ground-based TCCON retrieved XGHG was adjusted to the same common a priori used for the satellite data, i.e. SLIM2025, as explained in Sect. 3.1.1. For the L3 validation, this adjusted TCCON data was gridded to the same spatial resolution as the L3 merged XGHG product, namely monthly 1°x 1°. Therefore, if several stations lie within the same grid cell, their data is combined.

The collocations for the L2 validation (based on L2 satellite and TCCON GGG2020 XGHG data) are defined, for each L2 sounding, as the pairs satellite L2-TCCON observations taken less than 2 hours apart from each other, over locations separated by less than 250 km and whose altitude difference is less than 250 m. The satellite L2 altitude is obtained from the GLO-90

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

digital elevation model (DEM) (ESA, 2024), interpolated at the longitude and latitude of the ground pixel center. To ensure robust results for the validation, only TCCON stations with 12 or more months with collocations are considered.

Analogously, the collocations for the L3 validation (based on gridded L3 satellite and TCCON GGG2020 XGHG data) are defined, for each merged L3 grid cell, as the pairs satellite L3-TCCON that correspond to the same month, whose cell centers are separated by less than 250 km and whose altitude difference is less than 250 m. The altitude for the L3 validation is as well obtained from the GLO90 DEM, in this case re-gridded to the L3 spatial resolution. To ensure robust results for the validation, only TCCON stations with 30 or more months with collocations are considered.

The validation method is based on a fit of the differences between the satellite and the TCCON XGHG values to a trend model. This trend model consists of a constant to account for an offset, a linear drift with time, and a sinusoidal part to model the seasonal fluctuations. The fit is performed independently for the collocations of each TCCON station, from which per-station statistics are derived. In short, the per-station statistics consist of a regional bias, Δ_{reg}, defined as the mean over time of the fit model; a seasonal bias, Δ_{sea}, defined as as the standard deviation of the seasonal term of the fit. The spatio-temporal bias, Δ_{spt}, is defined as the root sum of squares (RSS) of the regional and seasonal biases. The linear drift, Δ_{dri}, corresponds to the slope of the linear term of the trend model. The single-sounding precision, (L2) or single-grid-cell precision (L3), σ, is computed as the standard deviation of the residual. Finally, the reported uncertainty, σ_{rep}, is computed as the root mean square (RMS) over time of the reported XGHG uncertainty of each sounding (L2) or grid cell (L3).

From these per-station statistics, the overall statistics are derived as follow. From the regional per-station biases, Δ_{reg}, we compute the average station bias (or global offset), $\overline{\Delta_{\text{reg}}}$, as the mean and the station-to-station variability, σ_{reg}, as the standard deviation of Δ_{reg}. The average seasonal bias, $\overline{\Delta_{\text{sea}}}$, is computed as the mean of the per-station seasonal biases, Δ_{sea}. The overall spatio-temporal bias, $\overline{\Delta_{\text{spt}}}$ is the RSS of the station-to-station variability and the average seasonal bias. Both the overall single-sounding (L2) or single-cell (L3) precision, $\overline{\sigma}$, and the overall reported uncertainty, $\overline{\sigma_{\text{rep}}}$, as computed as the RMS from their per-station counterparts. The details of the validation method are explained in Reuter et al. (2025b).

4 Results and discussion

4.1 Merged XCO_2 products

Both a L2 and a L3 XCO_2 product were generated with the EMMA v5.1 algorithm. The L2 XCO_2 dataset generated with EMMA v5.1 is the result of merging the L2 input products listed in Table 1 as described in Sects. 3.1.1 to 3.1.6. By design, the merged L2 datasets inherit their sampling and ground pixel size from the underlying L2 input used as a reference (depicted as hatched bars in Fig. 2 and enumerated in Sect. 3.1). The same applies to the L3 merged dataset, illustrated in Figs. 7 and 8, showing each of these figures three examples of the monthly merged L3 XCO_2 dataset for three different years, corresponding to different reference products.

Figure 7 shows global maps of the monthly L3 XCO_2 dataset for the month of May and for three different years (panels a to c). Despite the different samplings shown in panels a to c, arising from the different reference products, there are also visible

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

similarities among the three maps. The north-south gradient is specially noticeable, with elevated XCO_2 values in the northern hemisphere and lower values in the southern hemisphere.

Figure 8 is analogous to Fig. 7, but showing three plots for the month of July. The difference in the sampling among the three panels is very similar to that in Fig. 7. The spatial pattern is also similar across the three panels a to c, in this case with lower values in the northern hemisphere compared to the southern hemisphere.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

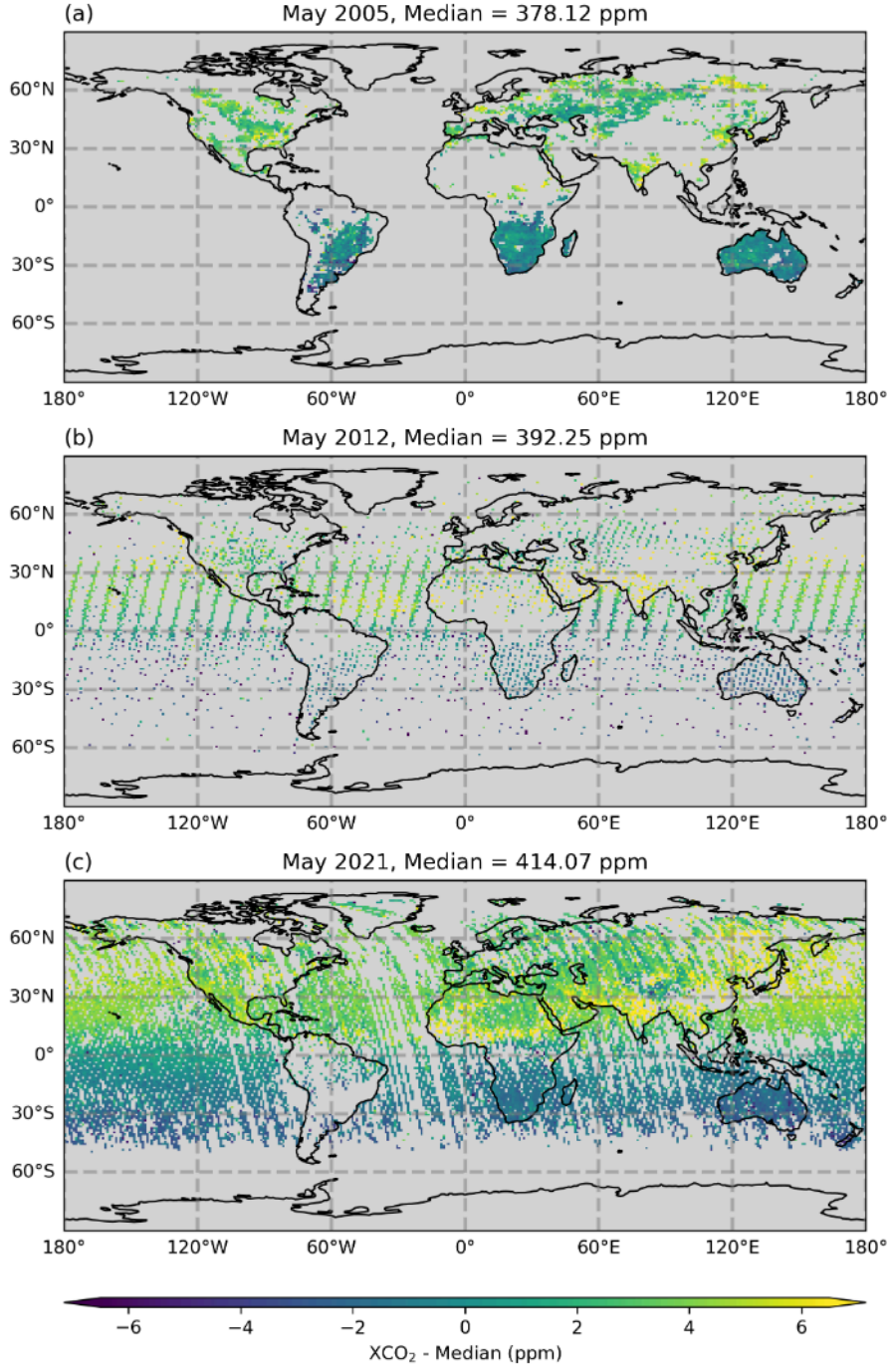


Figure 7. Global maps of the merged L3 XCO_2 dataset for May in three different years corresponding to different reference products. Panel a shows a global map for May 2005, when BESD is the reference algorithm. Panel b depicts May 2012, when GOSAT FOCAL is used as a reference. In panel c, May 2021 is shown, when OCO-2 FOCAL is the reference product. The color map shows the difference between the monthly XCO_2 and the median of the corresponding year, shown in the title of each panel. The illustration of the monthly XCO_2 with respect to the annual median allows us to visualize the XCO_2 regional patterns with the same color scale despite the continuous global annual increase. Gray areas indicate the absence of data.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

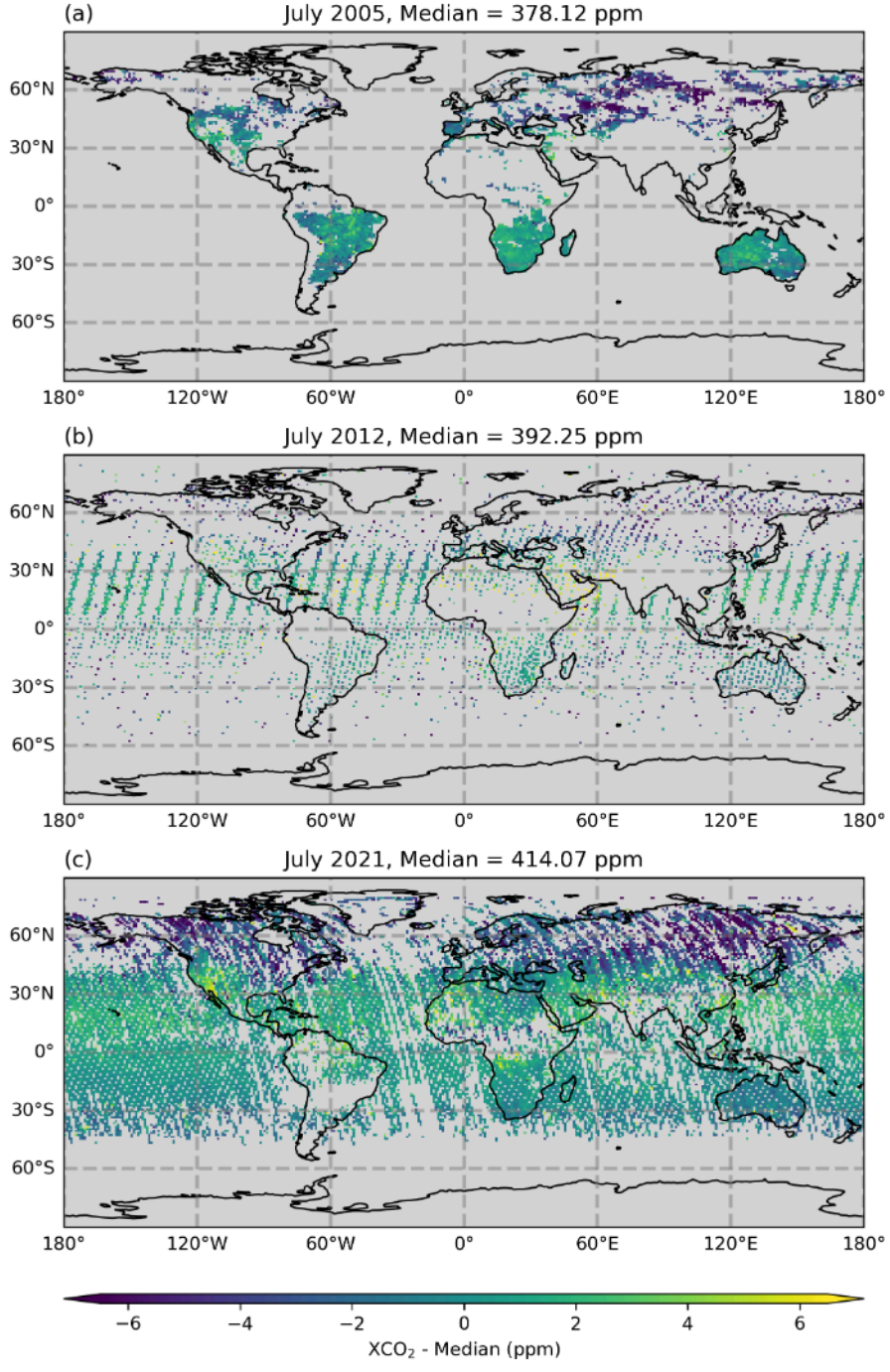


Figure 8. Same as Fig. 7 for the month of July.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

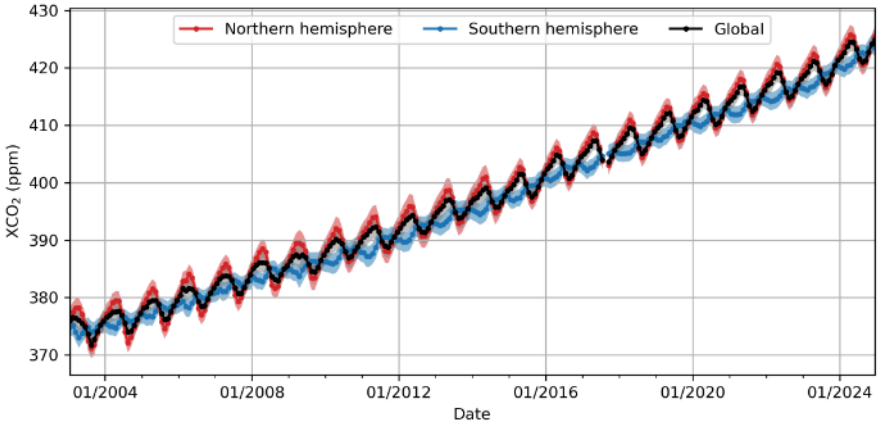


Figure 9. XCO_2 time series for the northern (NH, red) and southern (SH, blue) hemispheres, as well as global (black). Only data for latitudes between 60°S and 60°N and land fraction larger than 0.8 are shown. The shaded areas show the root-mean-squared uncertainty (not to be confused with the uncertainty of the mean) consisting of calibrated measurement noise and the adjustment uncertainty, computed from the inter-algorithm-spread.

The spatial patterns shown in Figs. 7 and 8 illustrate the inter-annual cycle of XCO_2, mainly driven by vegetation, and more pronounced in the northern hemisphere (e.g., Keeling and Graven, 2021; Liu et al., 2024). This seasonal variability is also evident in Fig. 9, which shows the XCO_2 time series computed from the L3 merged dataset for latitudes between 60°S and 60°N and land fraction larger than 0.8. The time series also show the well-known long-term increase in XCO_2 (e.g., Friedlingstein et al., 2025).

During the time in which the SCIAMACHY BESD data is used as the reference product (from the beginning of the dataset until the end of March 2009), there is no other L2 input products, i.e. there is no ensemble to compute an adjustment, so the SCIAMACHY BESD is included in the merged L2 (and L3) using an adjustment of zero almost everywhere. Only the months of February and March contain a partial adjustment by the smoothed B field computed from the ensemble after the 1 April 2009. During the time period in which SCIAMACHY is the reference product, it only measured over land.

In August 2017 there is a gap in both L2 and L3 XCO_2 datasets. This gap arises from the corresponding OCO-2 input product used as a reference, and is originated due to a reboot of the sensor.

The overall validation results for both the L2 and L3 XCO_2 datasets are shown in Table 3, together with the validation results of the individual input L2 products. The table shows the overall statistics described in Sect. 3.2. An inter-comparison among the validation results for the different products is however not straightforward. The criteria for the L2 and L3 validations, though very similar, are inherently different. Among the L2 products, which use the same criteria for the collocations, the different samplings and temporal availability lead to different stations with collocations for the different products. Additionally, the EMMA L2 product itself contains different samplings for different time periods. Therefore, for the TCCON stations with collocations across the different periods, the validation is typically dominated by the period in which the reference product has

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

a higher data volume. For this reason, the EMMA L2 data set was also independently validated for the time periods determined by the different reference products, whose outcome is also shown in Table 3.

The EMMA L2 XCO_2 data set has an overall station bias (global offset) of 0.35 ppm, and a station-to-station variability (regional bias) of 0.40 ppm. These regional bias is smaller than that of all other input L2 products, with the exception of SCIAMACHY BESD. However, the SCIAMACHY validation is derived from only five TCCON stations, in contrast to the more than 18 stations used for all the other products, so the statistics for the SCIAMACHY BESD validation are expected to be less robust. The seasonal bias of the EMMA L2 XCO_2 dataset is 0.26 ppm, and the spatio-temporal bias is 0.47 ppm, the latter being lower than that of all other input products. The linear drift of 0.00 ± 0.12 ppm y^{-1} suggest high stability, similarly to the input products. The single sounding precision is 1.51 ppm, only higher than that of the OCO-2 and OCO-3 products, and close to the mean reported uncertainty. These results show an overall good performance of the EMMA L2 XCO_2 dataset, analogously to all the other input products.

The most straightforward comparison is between the EMMA L2 (OCO-2) and OCO-2 FOCAL, since both products have the same sampling and the same period was used for the validation. EMMA L2 (OCO-2) slightly improved the regional and therefore the spatio-temporal bias, as well as the linear drift, at the expense of a higher global offset. Additionally, the similarity between the validation results of EMMA L2 (OCO-2) and EMMA L2 show the strong influence in the EMMA L2 validation results of the period in which OCO-2 FOCAL is used as a reference.

The validation results of the L3 dataset are similar to those of the L2 validation, with a slightly larger station-to-station bias of 0.42 ppm and a higher value of the single-cell precision, due to possible representation errors. Appendix F1 contains more detailed information on the validation results.

Figures 10 and 11 show the overall agreement between the merged XCO_2 EMMA v5.1 L2 and L3 products, respectively, with TCCON data.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

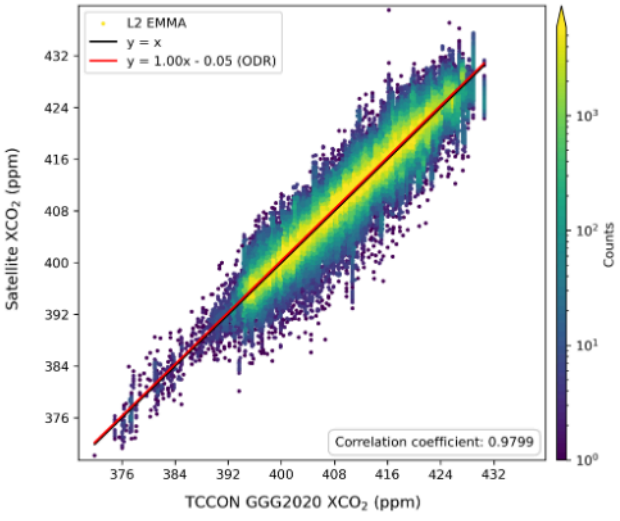


Figure 10. XCO_2 EMMA v5.1 L2 product vs TCCON data, where each dot stands for one collocation. The one-to-one line (black), orthogonal distance regression (red) and Pearson correlation coefficient are also shown.

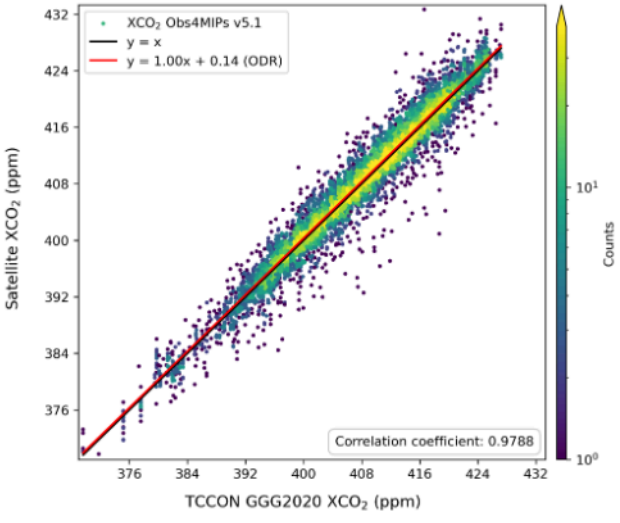


Figure 11. Same as Fig. 10 for XCO_2 EMMA v5.1 Obs4MIPs (L3) product vs TCCON data.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 3. Overview of the validation results for each L2 XCO_2 input product and for the merged L2 and L3 EMMA v5.1 XCO_2 datasets. For each product (first column) the overall validation statistics described in Sect. 3.2 are shown, namely the average station bias, $\overline{\Delta_{reg}}$ together with the station-to-station variability, σ_{reg}, the average seasonal bias, $\overline{\Delta_{sea}}$, the overall spatio-temporal bias, $\overline{\Delta_{spt}}$, the overall single-sounding (L2) and single-cell (L3) precision, $\overline{\sigma}$, and the ratio between the latter and the overall reported uncertainty, $\overline{\sigma_{rep}}$. The latter corresponds to the scaling factor used for the uncertainty calibration of the input L2 products (see Sect. 3.1.2). The scaling factor was only applied to the uncertainties of the input L2 products; for the EMMA products it gives a measure of agreement between the reported and the single-sounding (L2) or single-cell (L3) precision.

Product	$\overline{\Delta_{reg}} \pm \sigma_{reg}$ (ppm)	$\overline{\Delta_{sea}}$ (ppm)	$\overline{\Delta_{spt}}$ (ppm)	$\overline{\Delta_{dri}}$ (ppm y^{-1})	$\overline{\sigma}$ (ppm)	$\overline{\sigma}/\overline{\sigma_{rep}}$
SCIAMACHY BESD	-0.25 ± 0.23	0.62	0.66	0.18 ± 0.18	1.88	0.96
GOSAT FOCAL	0.45 ± 0.82	0.51	0.96	-0.08 ± 0.15	2.71	0.99
GOSAT-2 FOCAL	0.32 ± 0.81	0.53	0.97	-0.12 ± 0.22	2.78	1.09
OCO-2 FOCAL	0.02 ± 0.44	0.25	0.51	0.02 ± 0.11	1.50	0.93
GOSAT ACOS	-0.10 ± 0.49	0.28	0.56	0.07 ± 0.25	1.68	1.35
OCO-2 ACOS	0.11 ± 0.47	0.23	0.53	-0.03 ± 0.12	1.22	2.31
GOSAT RemoTeC	0.39 ± 0.98	0.51	1.11	0.18 ± 0.33	2.38	1.13
GOSAT-2 RemoTeC	-0.06 ± 0.47	0.47	0.67	0.07 ± 0.21	2.02	0.94
GOSAT UoL FP	0.17 ± 0.59	0.61	0.85	0.06 ± 0.14	1.85	0.96
GOSAT NIES	0.68 ± 0.70	0.34	0.78	0.01 ± 0.15	2.01	1.73
GOSAT-2 NIES	3.39 ± 0.83	0.55	1.00	0.06 ± 0.18	2.29	3.79
OCO-3 ACOS	0.12 ± 0.48	0.28	0.55	0.08 ± 0.17	1.42	2.21
EMMA L2	0.35 ± 0.39	0.26	0.47	0.00 ± 0.12	1.51	0.91
EMMA (SCIAMACHY)*	-0.04 ± n.a.	0.84	n.a.	0.18 ± n.a.	1.83	0.91
EMMA L2 (GOSAT)	0.44 ± 0.66	0.50	0.83	0.24 ± 0.44	2.22	0.84
EMMA L2 (OCO-2)	0.34 ± 0.40	0.25	0.48	0.00 ± 0.12	1.50	0.91
EMMA L3	0.29 ± 0.42	0.24	0.48	0.00 ± 0.11	1.82	1.10

*Results computed from only one station, therefore some fields are not applicable (n.a.).

4.2 Merged XCH_4 products

The EMMA v5.1 algorithm was also used to generate both a L2 and a L3 XCH_4 product.

Analogous to the XCO_2 dataset, the EMMA v5.1 L2 XCH_4 dataset is the result of merging the L2 input products listed in Table 2 as described in Sects. 3.1.1 to 3.1.6. The L2 input products used as a reference are marked as hatched bars in Fig. 3 and described in Sect. 3.1. Figures 12 and 13 show examples of the L3 dataset for the months of April and October, respectively, each of them showing three global maps (panels a to c) of the XCH_4 dataset, corresponding to three different years with three different reference products.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Despite the different samplings shown in panels a to c, the spatial pattern in the XCH_4 maps also shows seasonal similarities, arising from the seasonality of its sources (like wetlands or biomass burning) and sinks (mainly reaction with hydroxyl radicals), and driven by meteorological conditions (e.g., Dowd et al., 2023; Kirschke et al., 2013). This seasonal cycle is again visible in the time series shown in Fig. 14, showing, as for XCO_2, the XCH_4 time series computed from the L3 merged dataset for latitudes between 60°S and 60°N and land fraction larger than 0.8. In Fig. 14 also the overall increase of atmospheric CH_4 over time is visible (e.g., Saunois et al., 2025). The time series shows also the higher XCH_4 values in the northern hemisphere, resulting from higher emissions compared to the southern hemisphere.

In Fig. 14 we can observe a large uncertainty until 2009, corresponding to the time period in which SCIAMACHY is the reference (and only) product. The uncertainty is specially high after November 2005, due to the loss of some methane-sensitive spectral detector pixels (Schneising et al., 2012).

As is the case for the merged XCO_2 datasets, during the time in which the SCIAMACHY product, SCIAMACHY WFMD, is used as the reference (from the beginning of the dataset until the end of March 2009), there is no ensemble to compute an empirical spatio-temporal adjustment, so the SCIAMACHY WFMD is included in the merged datasets using an adjustment of zero almost everywhere (with the exception of the result of smoothing the B field computed from the ensemble between the 1 April 2009 and the end of the SCIAMACHY WFMD dataset). As for the XCO_2 datasets, the lower quality of the SCIAMACHY measurements compared to other products is as well reflected in the merged XCH_4 datasets. Additionally, after November 2005, SCIAMACHY measured only over land.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

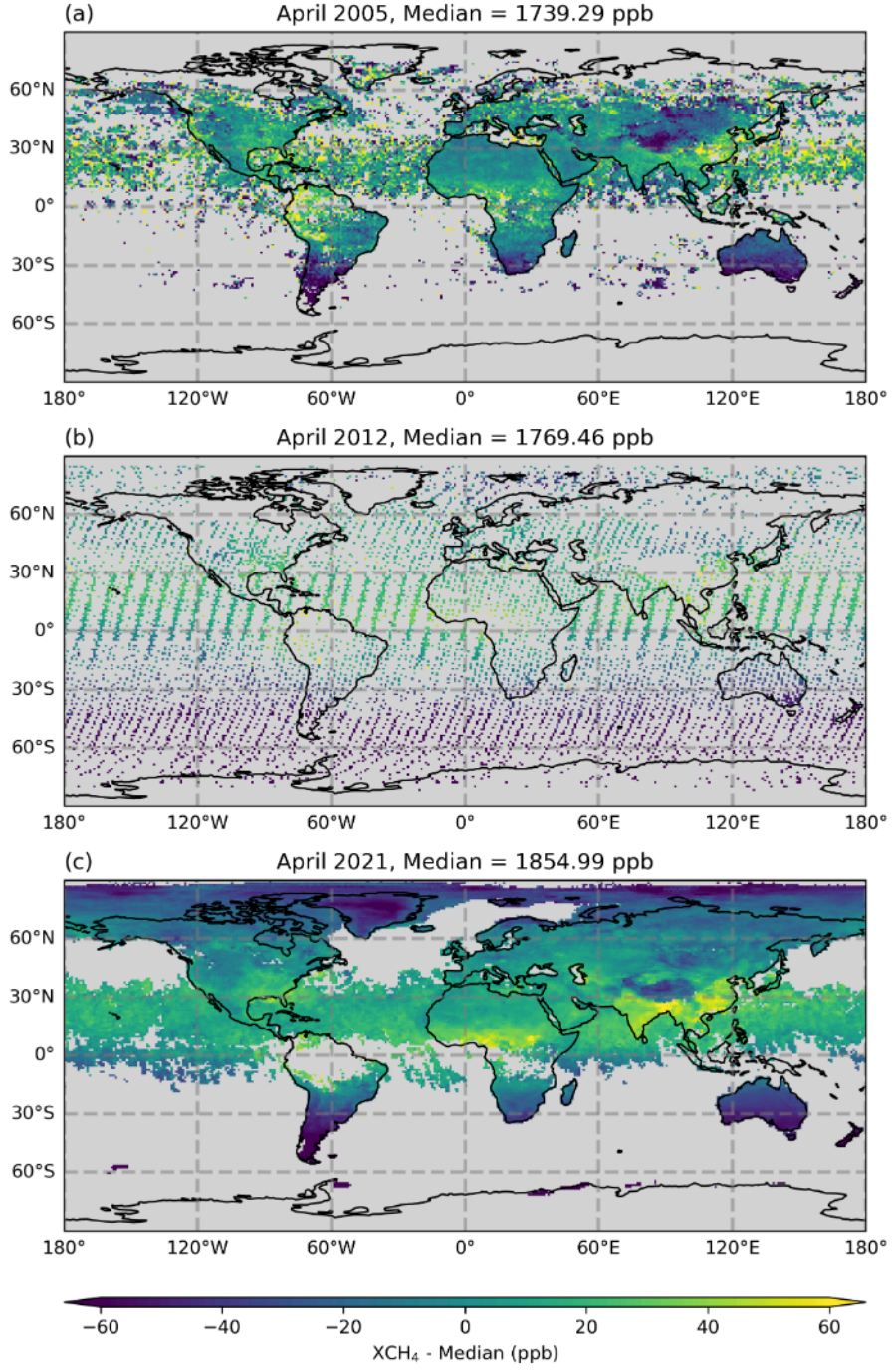


Figure 12. Global maps of the merged L3 XCH_4 dataset for April in three different years, corresponding to different reference products. Panel a shows a global map for April 2005, when SCIAMACHY WFMD is the reference product. Panel b depicts April 2012, when GOSAT FOCAL PR is used as a reference. In panel c, April 2021 is shown, when TROPOMI WFMD is the reference product. The color map shows the difference between the monthly XCH_4 and the median of the corresponding year, represented as A and shown in the title of each panel. The illustration of the monthly XCH_4 with respect to A allows us to visualize the XCH_4 regional patterns with the same color scale despite the continuous global annual increase. Gray areas indicate the absence of data.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

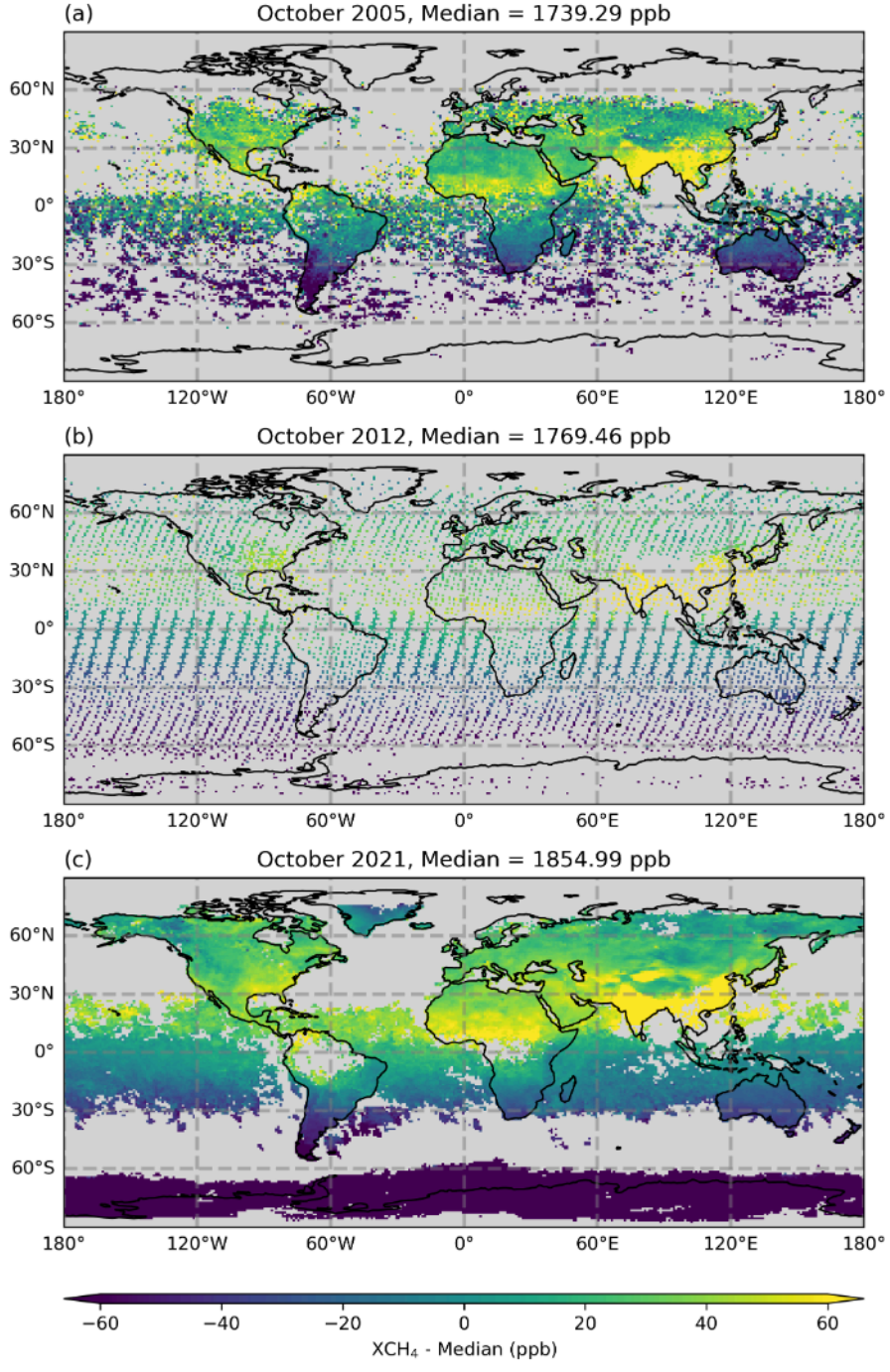


Figure 13. Same as Fig. 12 for the month of October.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

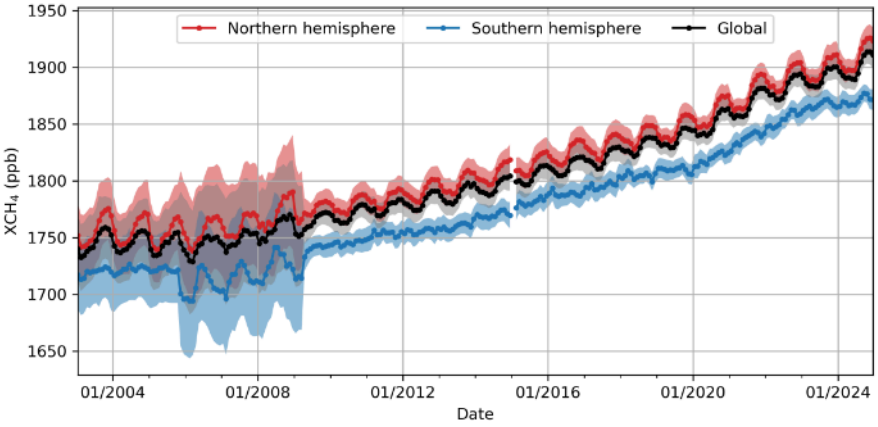


Figure 14. Analogous to Fig. 9 for the XCH_4 dataset.

In January 2015 there is a gap in both L2 and L3 XCH_4 datasets, arising from a GOSAT instrument adjustment that affected all GOSAT XGHG products. This gap resulted in a gap in the merged XCH_4 datasets, since a GOSAT product was used as a reference for this period.

The overall statistics resulting from the validation of the L2 and L3 XCH_4 datasets with the ground-based TCCON retrievals (see Sect. 3.2) are shown in Table 4, together with the validation results of the individual input L2 products. The EMMA L2 (L3) XCH_4 dataset has an overall offset, $\overline{\Delta_{\text{reg}}}$, of -6.09 ppb (-6.39 ppb) and a regional variability, σ_{reg}, of 6.63 ppb (3.37 ppb). The seasonal bias, $\overline{\Delta_{\text{sea}}}$, is 3.21 ppb (2.68 ppb) and the total spatio-temporal bias, $\overline{\Delta_{\text{spt}}}$, is 7.37 ppm (4.31 ppb). The linear drift, $\overline{\Delta_{\text{dri}}}$, was estimated to be 0.42 ± 1.14 ppb y^{-1} (0.11 ± 0.62 ppb y^{-1}). The single-sounding precision, σ_{reg}, is 20.87 ppb (9.01 ppb), very similar to the average reported precision, $\overline{\sigma_{\text{rep}}}$, of 21.11 ppb (11.05 ppb), leading to a ratio between the former and the latter of 0.99 (0.82). More detailed information on the validation results is shown in Appendix F2.

As mentioned in Sect. 4.1, a comparison among the validation results for the different products should be done carefully. In this case, the inclusion of different stations for the validation of the different products leads to higher differences among the results.

The EMMA L2 validation is strongly influenced by the stations of Izaña (iz) and Lauder (lh), as shown in Table F3. The number of collocations found for these two stations is much smaller (only about 0.002% of the total amount of collocations) than for all other sites. For the site of Lauder (lh), there are only collocations during the time in which SCIAMACHY was used as a reference, which causes a significant increase of the single-sounding error. For the site of Izaña (iz) there are collocations during the time in which TROPOMI WFMD is the reference product. However, due to the high altitude of the station compared to its surroundings, there are much less collocations than for the other stations, while having the same weight in the overall validation results.

Figures 15 and 16 show the merged XCH_4 EMMA v5.1 L2 and L3 datasets, respectively, plotted against their corresponding TCCON collocated data.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 4. Overview of the validation results for each L2 XCH_4 input product and for the merged L2 and L3 EMMA v5.1 XCH_4 datasets. The table is analogous to Table 3.

Product	$\overline{\Delta_{reg}} \pm \sigma_{reg}$ (ppb)	$\overline{\Delta_{sea}}$ (ppb)	$\overline{\Delta_{spt}}$ (ppb)	$\overline{\Delta_{dri}}$ (ppb y^{-1})	$\overline{\sigma}$ (ppb)	$\overline{\sigma}/\overline{\sigma_{rep}}$
SCIAMACHY WFMD	-2.35 ± 6.21	14.42	15.70	0.71 ± 9.55	98.33	1.17
GOSAT FOCAL	-7.55 ± 4.33	3.00	5.27	-0.42 ± 1.39	16.90	0.86
GOSAT FOCAL PR	-6.22 ± 4.79	2.01	5.20	0.14 ± 0.71	13.63	0.90
GOSAT-2 FOCAL	-6.05 ± 4.29	2.71	5.08	0.33 ± 1.39	13.98	0.83
GOSAT-2 FOCAL PR	-6.31 ± 3.80	1.73	4.18	0.15 ± 0.93	12.09	0.83
GOSAT RemoTeC	8.89 ± 3.84	1.93	4.30	0.79 ± 1.32	13.44	0.98
GOSAT RemoTeC PR	5.63 ± 3.59	2.63	4.45	-0.75 ± 1.38	14.28	1.07
GOSAT-2 RemoTeC	-0.51 ± 3.00	2.23	3.74	1.69 ± 1.23	12.12	0.91
GOSAT-2 RemoTeC PR	4.43 ± 3.53	1.81	3.96	1.25 ± 0.72	13.88	0.99
GOSAT UoL FP	3.61 ± 3.10	2.48	3.97	-0.15 ± 0.72	12.15	0.85
GOSAT UoL PR	7.40 ± 4.22	2.29	4.80	0.74 ± 1.01	13.54	1.18
GOSAT NIES	-1.79 ± 3.25	1.79	3.71	0.04 ± 0.72	12.04	1.53
GOSAT-2 NIES	-4.04 ± 4.42	3.01	5.35	0.39 ± 1.21	11.32	4.69
GOSAT-2 UoL PR	10.21 ± 4.84	3.71	6.10	2.10 ± 0.55	11.77	0.68
TROPOMI WFMD	-0.01 ± 7.16	2.93	7.73	-0.22 ± 1.16	14.79	1.05
EMMA L2	-6.09 ± 6.63	3.21	7.37	0.42 ± 1.14	20.87	0.99
EMMA L2 (SCIAMACHY)	-19.53 ± 0.30	13.19	13.19	-6.06 ± 2.25	80.15	0.99
EMMA L2 (GOSAT)	-8.22 ± 3.66	1.88	4.12	-0.15 ± 0.93	12.73	0.90
EMMA L2 (TROPOMI)	-4.81 ± 6.34	3.04	7.03	0.57 ± 1.08	14.67	0.91
EMMA L3	-6.39 ± 3.37	2.68	4.31	0.11 ± 0.62	9.01	0.82

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

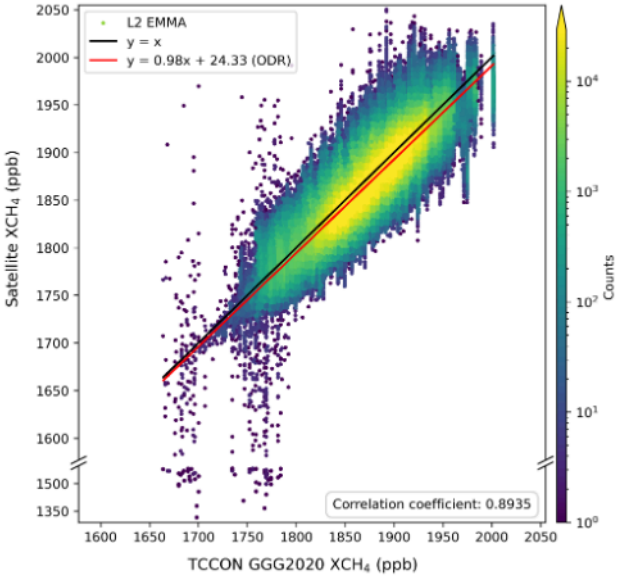


Figure 15. XCH_4 EMMA v5.1 L2 product vs TCCON data, analogous to Fig. 10.

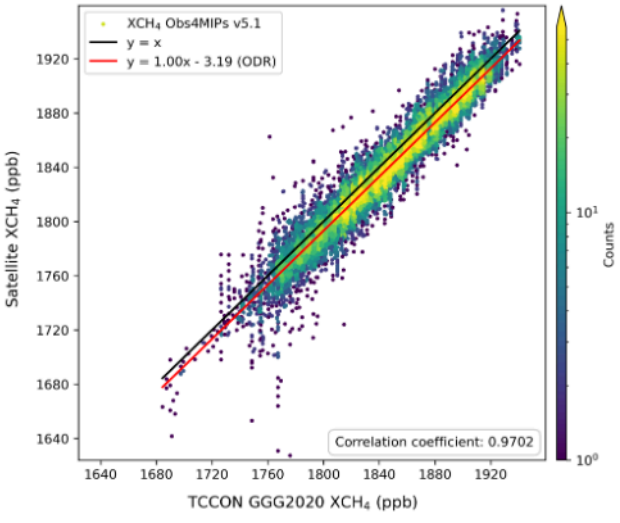


Figure 16. Same as Fig. 10 for XCH_4 EMMA v5.1 Obs4MIPs (L3) product vs TCCON data.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

5 Conclusions

Improved and updated ensemble-based L2 and L3 satellite-derived XCO_2 and XCH_4 data sets were generated using the new EMMA v5.1 algorithm. The four merged datasets cover the time period from 2003 to 2024. The XCO_2 L2 dataset is available at https://doi.org/10.5281/zenodo.18610669 (Fuentes Andrade et al., 2026b). The XCH_4 dataset is available at https://doi.org/10.5281/zenodo.18608583 (Fuentes Andrade et al., 2026a). The L3 datasets will be available at the CDS (https://cds.climate.copernicus.eu/).

The new EMMA v5.1 approach is conceptually similar to earlier EMMA versions. It is based on the same assumption that the ensemble median is a good representative of the true atmospheric state over a region, while adopting a different modus operandi. By selecting a reference L2 product for each time period and applying the empirical spatio-temporal adjustment derived from the ensemble of input products, EMMA v5.1 eliminates potential discontinuities (at monthly 10°x 10° boundaries) inherent to previous EMMA implementations. As a result, the merged L2 data preserve the spatial resolution and sampling of the reference products. Therefore, the chosen reference products lead to an increased sampling density and spatial coverage, mainly in the time period in which OCO-2 FOCAL for XCO_2 and TROPOMI WFMD for XCH_4 were the reference products.

The presented datasets also include data retrieved from a larger number of sensors, as well as new and updated L2 input products. Additionally, the L3 products have higher spatial resolution ($1° \times 1°$) compared to previous versions of these datasets.

From the validation against TCCON data we conclude that the products present long-term stability and no significant biases, constituting therefore long-term and consistent data records of XCO_2 and XCH_4.

Some applications of the XCO_2 and XCH_4 L2 and L3 products presented here are flux inversions and trend analysis. Previous versions of these datasets have been used for comparison with climate models (Laughner et al., 2024; Gier et al., 2020), and estimation of mean annual growth rates (Buchwitz et al., 2018; Reuter et al., 2020; C3S and WMO, 2025). Additionally, the spatial consistency and increased sampling density, which is inherited from the corresponding L2 reference products, makes the new resulting L2 datasets suitable for all the applications of their underlying L2 reference products, such as hotspot emission estimation, which is possible using e.g. TROPOMI WFMD or OCO-2 FOCAL.

Data availability. The EMMA v5.1 merged L2 datasets are available at https://doi.org/10.5281/zenodo.18610669 (Fuentes Andrade et al., 2026b) for XCO_2 and at https://doi.org/10.5281/zenodo.18608583 (Fuentes Andrade et al., 2026a) for XCH_4. The corresponding EMMA v5.1 merged L3 XCO_2 and XCH_4 datasets will be made available at https://cds.climate.copernicus.eu/.

The input L2 products are available at the following addresses. The SCIAMACHY BESD data is available at https://www.iup.uni-bremen.de/~mreuter/besd.php, SCIAMACHY WFMD data at https://www.iup.uni-bremen.de/sciamachy/NIR_NADIR_WFM_DOAS/products/, OCO-2 FOCAL at https://www.iup.uni-bremen.de/~mreuter/focal.php, and TROPOMI WFMD at https://www.iup.uni-bremen.de/carbon_ghg/products/tropomi_wfmd/. The ACOS products for GOSAT, OCO-2 and OCO-3 are available at https://oco2.gesdisc.eosdis.nasa.gov/data/. The GOSAT NIES data can be found at https://data2.gosat.nies.go.jp and GOSAT-2 NIES at https://prdct.gosat-2.nies.go.jp. The GOSAT and GOSAT-2 FOCAL products are available on request from the authors. The GOSAT UoL-FP products can be downloaded from https://cds.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

climate.copernicus.eu/. The GOSAT and GOSAT-2 UoL-PR products, as well as GOSAT RemoTeC data are available at https://catalogue.ceda.ac.uk/. The GOSAT-2 RemoTeC data can be found at https://zenodo.org/records/12180512.

Appendix A: List of TCCON sites

Table A1 shows a list with the TCCON sites used for the validation, sorted from north to south. Their location name, longitude, latitude, altitude and corresponding references are shown along.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table A1. TCCON sites, sorted by descending latitude (i.e. north to south) used for the validation of the merged EMMA v5.1 XGHG products.

Site ID	Location	Longitude (°)	Latitude (°)	Altitude (km)	Reference
eu	Eureka, Canada	-86.44	80.02	0.61	Strong et al. (2022)
ny	Ny-Ålesund, Svalbard, Norway	11.91	78.98	0.02	Buschmann et al. (2022)
so	Sodankylä, Finland	26.64	67.24	0.19	Kivi et al. (2022)
et	East Trout Lake, Canada	-104.47	54.18	0.50	Wunch et al. (2022)
bi	Bialystok, Poland	23.00	53.13	0.19	Petri et al. (2024a)
br	Bremen, Germany	8.85	53.11	0.03	Notholt et al. (2022)
hw	Harwell, UK	-1.32	51.55	0.14	Weidmann et al. (2023)
ka	Karlsruhe, Germany	8.44	49.07	0.11	Hase et al. (2024)
pr	Sorbonne Université, Paris, FR	2.36	48.87	0.06	Té et al. (2022)
or	Orleans, France	2.11	47.99	0.13	Warneke et al. (2024)
gm	Garmisch, Germany	11.04	47.50	0.74	Sussmann et al. (2025)
pa	Park Falls, Wisconsin, USA	-90.36	46.00	0.47	Wennberg et al. (2022c)
rj	Rikubetsu, Hokkaido, Japan	143.89	43.48	0.38	Morino et al. (2022a)
if	Indianapolis, Indiana, USA	-86.00	39.87	0.27	Iraci et al. (2022a)
xh	Xianghe, China	117.00	39.77	0.04	Zhou et al. (2022)
oc	Lamont, Oklahoma, USA	-97.10	36.68	0.32	Wennberg et al. (2025)
tk	Tsukuba, Ibaraki, Japan, 125HR	140.02	36.02	0.03	Morino et al. (2022b)
ni	Nicosia, Cyprus	33.39	35.14	0.18	Petri et al. (2024b)
df	AFRC, Edwards, CA, USA	-117.40	34.72	0.70	Iraci et al. (2022b)
jc	JPL, Pasadena, California, USA	-118.08	34.23	0.39	Wennberg et al. (2022d)
jf	JPL, Pasadena, California, USA	-118.04	34.18	0.39	Wennberg et al. (2022b)
ci	Pasadena, California, USA	-119.02	34.04	0.24	Wennberg et al. (2022a)
js	Saga, Japan	130.14	33.18	0.01	Shiomi et al. (2022)
hf	Hefei, China	117.12	31.89	0.04	Liu et al. (2023)
iz	Izana, Tenerife, Spain	-16.50	28.27	2.37	García et al. (2022)
bu	Burgos, Philippines	120.72	18.51	0.04	Morino et al. (2022c)
ma	Manaus, Brazil	-60.60	-3.21	0.05	Dubey et al. (2022)
db	Darwin, Australia	130.61	-12.48	0.04	Deutscher et al. (2023b)
ra	Reunion Island, France	55.50	-20.90	0.08	De Mazière et al. (2022)
wg	Wollongong, Australia	150.53	-34.47	0.03	Deutscher et al. (2023a)
ll	Lauder, New Zealand	169.21	-45.00	0.37	Sherlock et al. (2022b)
lr	Lauder, New Zealand	169.27	-45.00	0.37	Pollard et al. (2022)
lh	Lauder, New Zealand	169.68	-45.02	0.37	Sherlock et al. (2022a)

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix B: GOSAT and GOSAT-2 FOCAL v4.0 products

The GOSAT and GOSAT-2 FOCAL products v4.0 used in this work are essentially based on the algorithm described in Noël et al. (2022). The main difference is an updated post-processing (filtering and bias correction) which in general results in a larger number of finally valid data, especially over ocean. Furthermore, the temporal coverage of both data sets has been extended until the end of 2024.

The following main changes apply for both GOSAT and GOSAT-2:

- Use of more recent Level 1 product versions as input (v300.300 for GOSAT, v220.221 for GOSAT-2).
- Use of an updated SLIM2024 climatology for a-priori data.
- Use of updated a-priori covariances; especially, use of specific covariances for CH_4 instead of scaled CO_2 values.
- Use of an updated reference data base (also involving SLIM2024) in the post-processing.
- No manually adjusted filtering in post-processing for scattering parameters. All filter parameters and limits are now determined by the variance reduction method (see Noël et al. (2022)).

Specifically for GOSAT-2, the following modifications apply in addition:

- Radiance noise is calculated from the signal-to-noise values given in the Level 1 files.
- The XN_2O retrieval now uses also SLIM2024 a-priori data (instead of a fixed profile). Furthermore, XN_2O is now derived from profile retrieval, not by profile scaling as before.

Appendix C: Global offset correction on the individual gridded input products

Figures C1 and C2 show the influence of the global offset correction.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

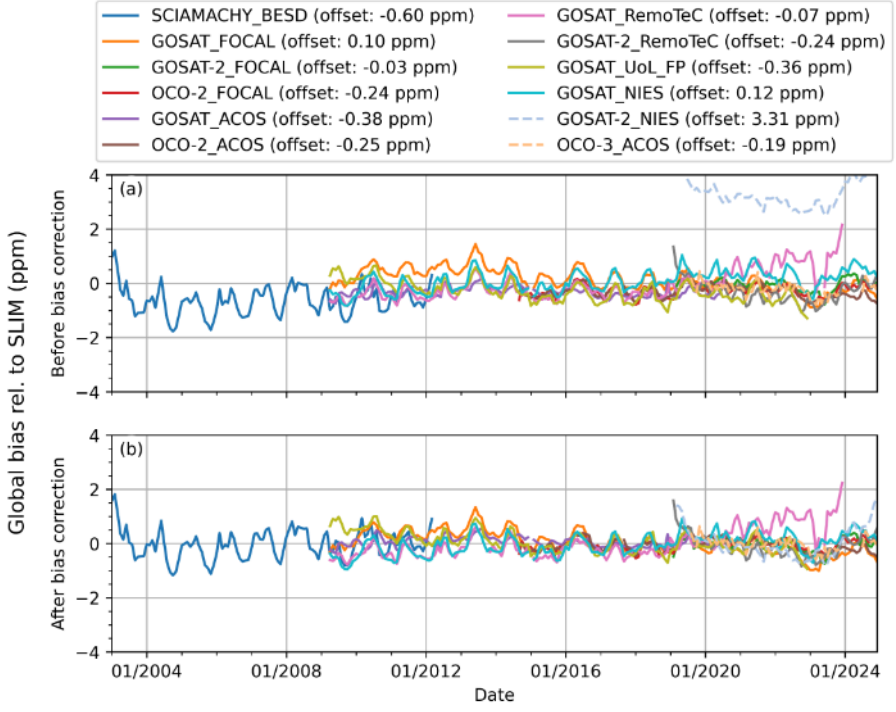


Figure C1. Global monthly average bias for XCO_2 relative to SLIM CO_2 before (panel a) and after (panel b) global offset adjustment. The global offset applied to each L2 product is shown in the legend.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

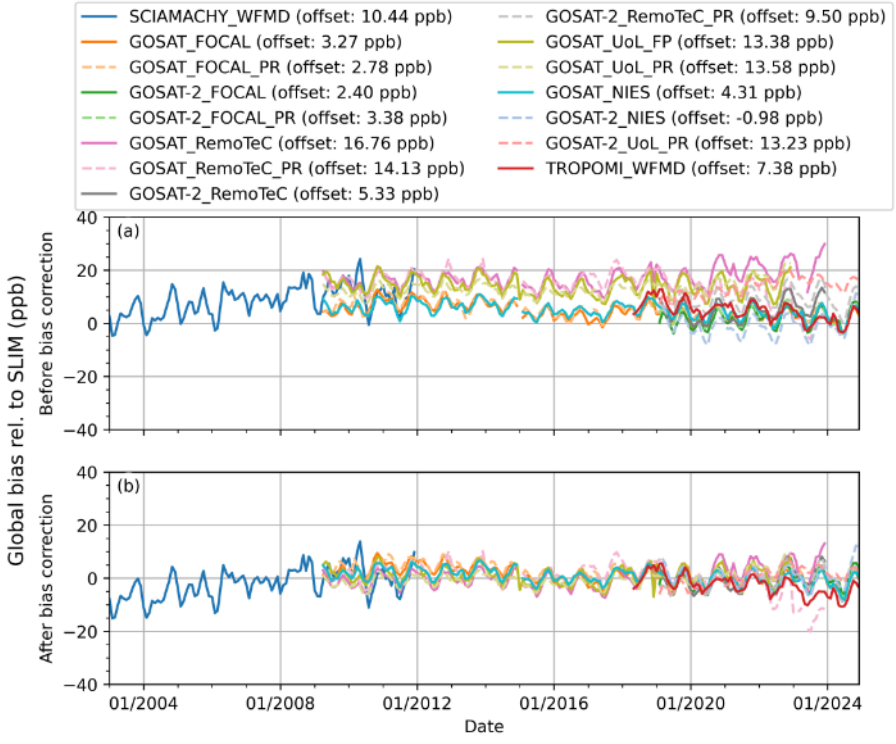


Figure C2. Global monthly average bias for XCH_4 relative to SLIM CH_4 before (panel a) and after (panel b) global offset adjustment. The global offset applied to each L2 product is shown in the legend.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix D: Examples of monthly gridded input products and their median

Figures D1 and D2 show examples of the calculated monthly 10°x 10° XCO_2 and XCH_4 averages, respectively, for the available input products in the corresponding months, together with the corresponding SLIM averages, the median field and the IAS (defined in Sect. 3.1.5). The different XGHG and SLIM averages and the median field show large-scale similarities, such as the north/south gradient, but also variations among the different L3 averages (see panels a to l in Fig. D1 and a to o in D2). The IAS is a measure of these variations among the different L3 averages. Examples of the computed IAS are shown in Fig. D1 (panel p) for XCO_2 and Fig. D2 (panel r) for XCH_4.

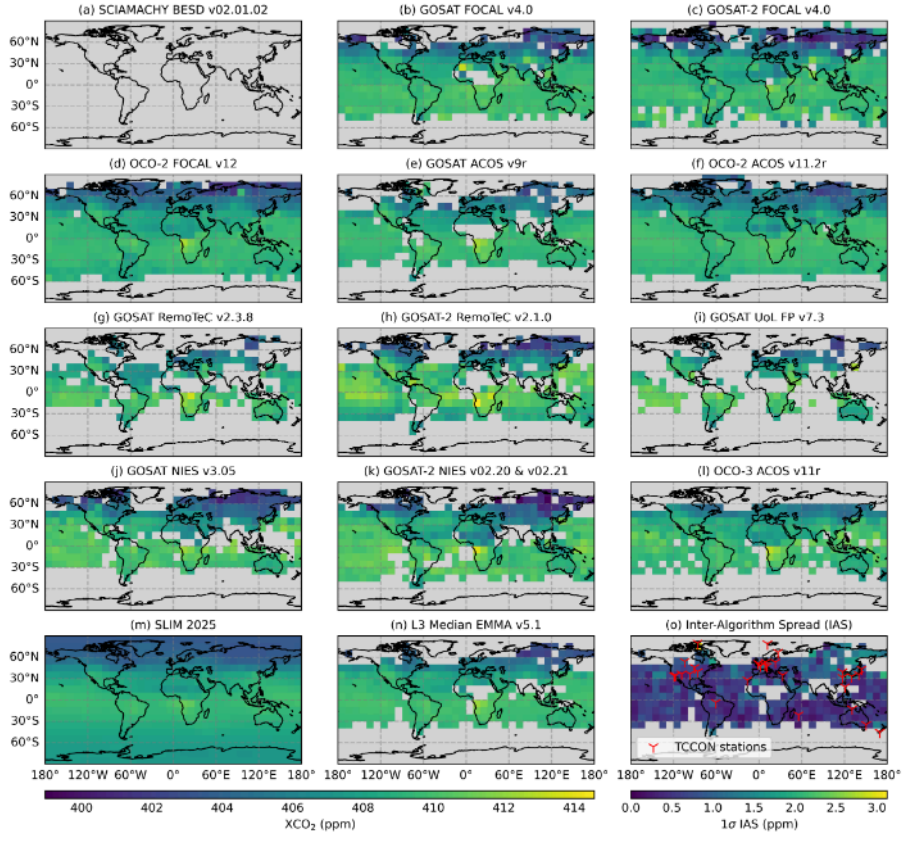


Figure D1. Monthly gridded XCO_2 averages for the L2 input satellite products after adjustment for a common a priori and offset correction (panels a to l), XCO_2 SLIM2025 (panel n), L3 median (panel o) and inter-algorithm spread (defined in Sect. 3.1.5, panel p) for the example of August 2019 for EMMA XCO_2. Gray areas indicate that there is no valid gridded L3 data in the corresponding grid cell. The red markers in panel (p) indicate the locations of TCCON stations.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

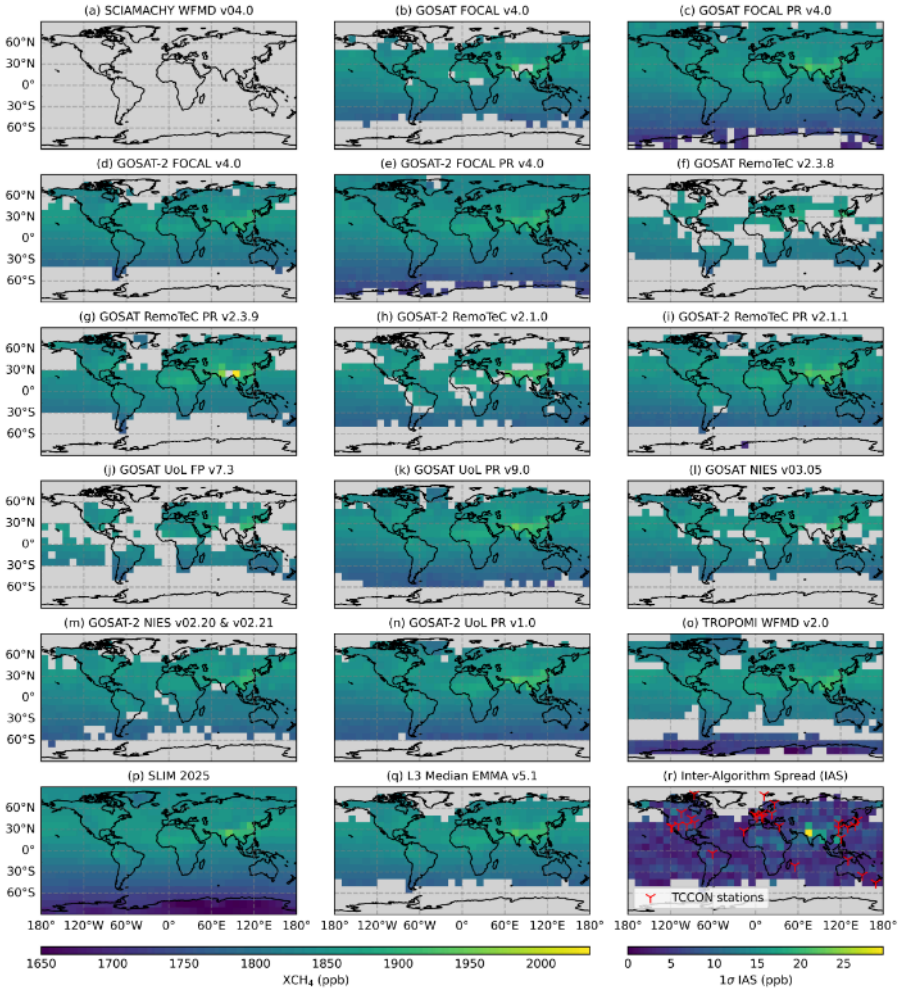


Figure D2. Monthly gridded XCH_4 averages for the L2 input satellite products after adjustment for a common a priori and offset correction (panels a to o), XCH_4 SLIM2025 (panel p), L3 median (panel q) and inter-algorithm spread (defined in Sect. 3.1.5, panel r) for EMMA XCH_4 v5.1 in September 2019. Gray areas indicate that there is no valid gridded L3 data in the corresponding grid cell. The red markers in panel (r) indicate the locations of TCCON stations.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix E: Modified normalized convolution

Let us define the validity mask, v, of the B field as a boolean mask that contains zeros where the B field has gaps and ones elsewhere, i.e.

$$v_i = \begin{cases} 1 & B_i \text{ is valid} \\ 0 & B_i \text{ is gap} \end{cases} \tag{E1}$$

for each element i of B. A synthetic example of a B field with gaps is illustrated in panel (a) of Fig. E1, and its corresponding validity mask in panel (e).

Let us denote with the subscript s the result of applying the smoothing kernel to a given field. Therefore, the smoothed validity mask, v_s, is the result of applying the smoothing kernel to v. A common approach for normalized convolution is to normalize by the smoothed validity mask, v_s. Panel (c) in Fig. E1 shows an example of gap-filling using such a convolution. However, large regions with gaps can lead to extrapolation artefacts, noticeable for example in the large region with low values at the center top of the same panel.

We defined a new mask, denoted u_s, for the normalized convolution. The new mask, u_s results from smoothing a modified validity mask, u. This modified validity mask, u, is derived from modifying v based on the values of v_s, so that u is set to one where the values of v_s are less than or equal to a given threshold, t, (i.e. in regions of low data density) and it keeps the values of v elsewhere:

$$u_i = \begin{cases} 1 & v_{s,i} <= t \\ v_i & v_{s,i} > t \end{cases} \tag{E2}$$

The corresponding examples of u and u_s are shown in panels (g) and (h) of Fig. E1, respectively.

The mask u effectively treats gaps with low data density as valid for the interpolation. These gaps in the B field are filled with zeros, and the result is convolved with a Gaussian kernel and normalized by u_s, as illustrated in panel (d). The resulting gap-filled B field will contain zero values in regions with low data density, which will lead to zero-valued spatio-temporal adjustments, i.e. they would leave the reference product unadjusted.

The Gaussian filters applied in all these steps have a sigma of half the grid-cell size and the convolution kernels are truncated at twice the grid-cell size. The t used as a threshold to define u is 0.001.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

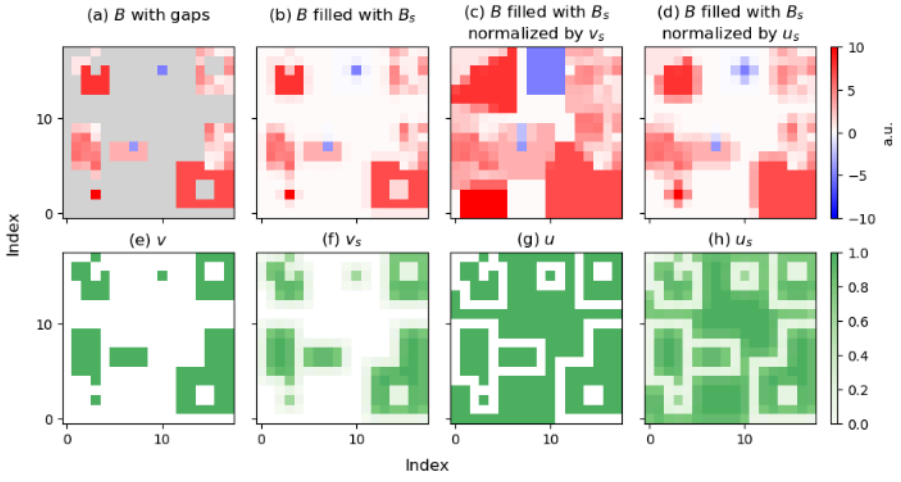


Figure E1. Illustration of the modified normalized convolution using synthetic data. Both the x and y axes in all panels show pixel indices. The top colour bar is expressed in arbitrary units (a.u.). Panel (a) represents the original B field with gaps. Panel (b) illustrates the B field after gap-filling with B_s. Panel (c) and (d) show the B field after gap filling with B_s normalized by v_s and by u_s, respectively. Panels (e) to (h) show the masks used for normalization: (e) illustrates the validity mask, and (f) its smoothed counterpart; (g) illustrates the modified validity mask, and (h) its smoothed version.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix F: Extended validation results

In this section, the validation results of the EMMA v5.1 XGHG (XCO_2 and XCH_4) L2 and L3 datasets are presented in more detail.

F1 Merged XCO_2 products

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table F1. Per station statistics (as explained in Sect. 3.2) obtained for the EMMA v5.1 XCO_2 L2 dataset. The columns indicate, from left to right, the TCCON station id (see Table A1), regional bias (Δ_{reg}), seasonal bias (Δ_{sea}), spatio-temporal bias (Δ_{spt}), linear drift (Δ_{dri}), single-sounding precision (σ), average reported uncertainty (σ_{rep}) and number of collocations (No. coll.). The last column contains the overall statistics, derived from their corresponding per-station values as indicated in Sect. 3.2.

Station	Δ_{reg} (ppm)	Δ_{sea} (ppm)	Δ_{spt} (ppm)	Δ_{dri} (ppm y^{-1})	σ (ppm)	σ_{rep} (ppm)	No. coll.
ny	-0.53	0.35	0.63	-0.17	1.75	1.88	297
so	-0.23	0.14	0.27	-0.08	1.62	1.74	34631
et	0.46	0.22	0.51	-0.08	1.54	1.75	43052
bi	0.25	0.10	0.27	0.14	1.28	1.60	16110
br	0.68	0.33	0.76	-0.07	1.24	1.74	6971
hw	0.08	0.54	0.55	0.28	1.40	1.70	11234
ka	0.50	0.43	0.67	-0.02	1.44	1.64	29200
pr	0.44	0.36	0.57	0.09	1.44	1.65	63688
or	0.64	0.12	0.66	0.07	1.40	1.65	59052
gm	1.18	0.41	1.25	0.10	1.38	1.64	21862
pa	0.07	0.39	0.39	0.02	1.58	1.65	96764
rj	0.48	0.21	0.52	0.02	1.71	1.72	14557
xh	0.58	0.38	0.69	0.06	1.96	1.73	27536
oc	0.50	0.20	0.53	-0.04	1.38	1.59	273239
tk	0.79	0.35	0.86	0.14	1.59	1.74	42218
ni	0.51	0.11	0.52	-0.01	1.44	1.61	136005
df	0.20	0.09	0.22	0.00	1.31	1.56	301007
jf	0.63	0.45	0.77	-0.05	1.84	1.62	10230
ci	0.06	0.25	0.26	0.05	1.66	1.59	282564
js	0.68	0.24	0.72	0.29	1.41	1.69	92087
hf	1.22	0.52	1.33	-0.12	1.80	1.71	26092
iz	0.03	0.23	0.23	-0.04	1.66	1.68	7497
bu	0.19	0.07	0.20	-0.07	1.47	1.64	37062
db	-0.20	0.19	0.27	-0.23	1.50	1.55	130005
ra	0.48	0.10	0.49	-0.18	1.33	1.58	35279
wg	0.17	0.14	0.22	-0.15	1.26	1.60	52083
lr	-0.21	0.19	0.29	-0.09	1.31	1.63	30809
ll	0.29	0.29	0.41	0.11	1.38	1.54	25612
Overall	0.35 ± 0.39	0.26	0.47	0.00 ± 0.12	1.51	1.66	1906743

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table F2. Analogous to Table F1 for the validation results of the EMMA v5.1 XCO_2 L3 dataset. The station IDs that appear aggregated lie within the same grid box are were considered as one station.

Station	Δ_{reg} (ppm)	Δ_{sea} (ppm)	Δ_{spt} (ppm)	Δ_{dri} (ppm y^{-1})	σ (ppm)	σ_{rep}	No. coll.
ny	-0.42	0.36	0.56	-0.11	2.07	1.91	166
so	-0.25	0.07	0.26	0.02	2.26	1.93	730
et	0.58	0.14	0.60	-0.08	2.18	1.64	689
bi	0.24	0.24	0.34	0.18	2.28	1.71	289
ka	0.30	0.30	0.42	-0.00	2.00	1.60	273
pr	0.54	0.39	0.66	-0.06	2.08	1.57	347
gm	0.75	0.36	0.83	0.11	1.57	1.75	173
or	0.76	0.24	0.80	0.03	2.05	1.65	630
pa	0.18	0.45	0.48	-0.03	2.15	1.74	1034
xh	1.15	0.23	1.18	0.03	2.76	1.71	255
oc	0.36	0.14	0.38	0.07	1.67	1.54	1583
tk	0.54	0.35	0.64	0.16	1.80	1.74	261
ni	0.16	0.12	0.20	-0.03	1.17	1.52	571
df	0.72	0.12	0.73	0.02	1.35	1.39	732
jc+ci+jf	-0.66	0.46	0.80	-0.06	1.91	1.57	656
js	0.69	0.14	0.71	0.24	1.67	1.73	587
bu	0.10	0.36	0.38	0.05	1.08	1.58	343
db	0.01	0.06	0.06	-0.16	1.38	1.54	764
ra	0.19	0.22	0.29	-0.19	0.98	1.55	400
wg	0.12	0.05	0.13	-0.09	1.24	1.65	559
lh+lr+ll	0.09	0.25	0.27	-0.10	1.40	1.54	296
Overall	0.29 ± 0.42	0.24	0.48	0.00 ± 0.11	1.82	1.65	11338

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

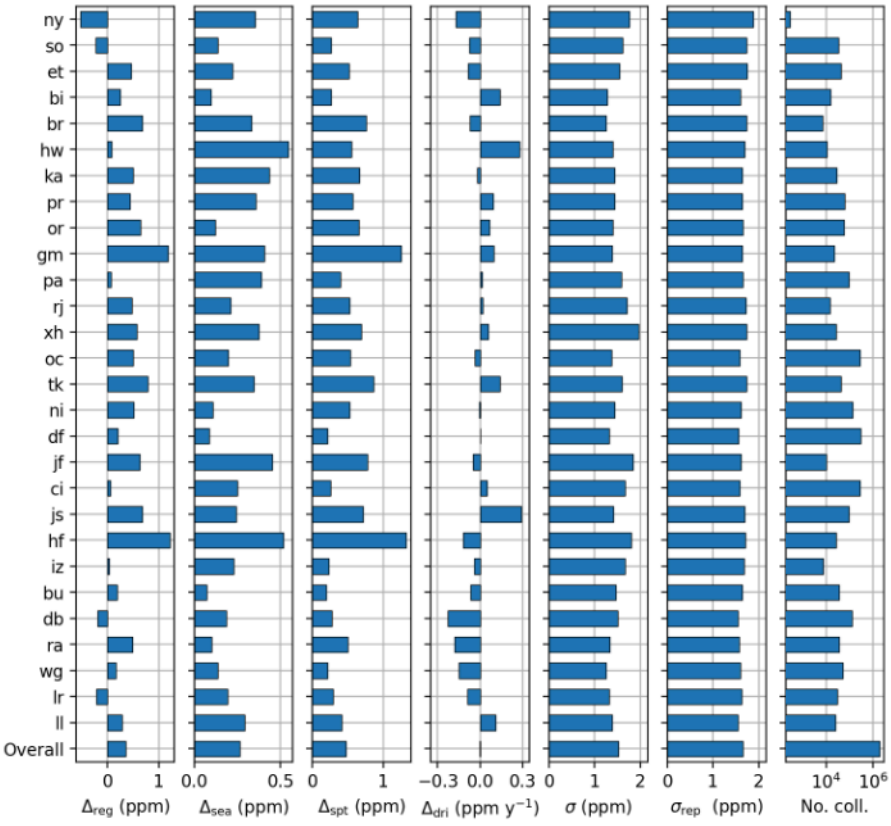


Figure F1. Bar plot of the XCO_2 L2 EMMA validation results shown in Table F1.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

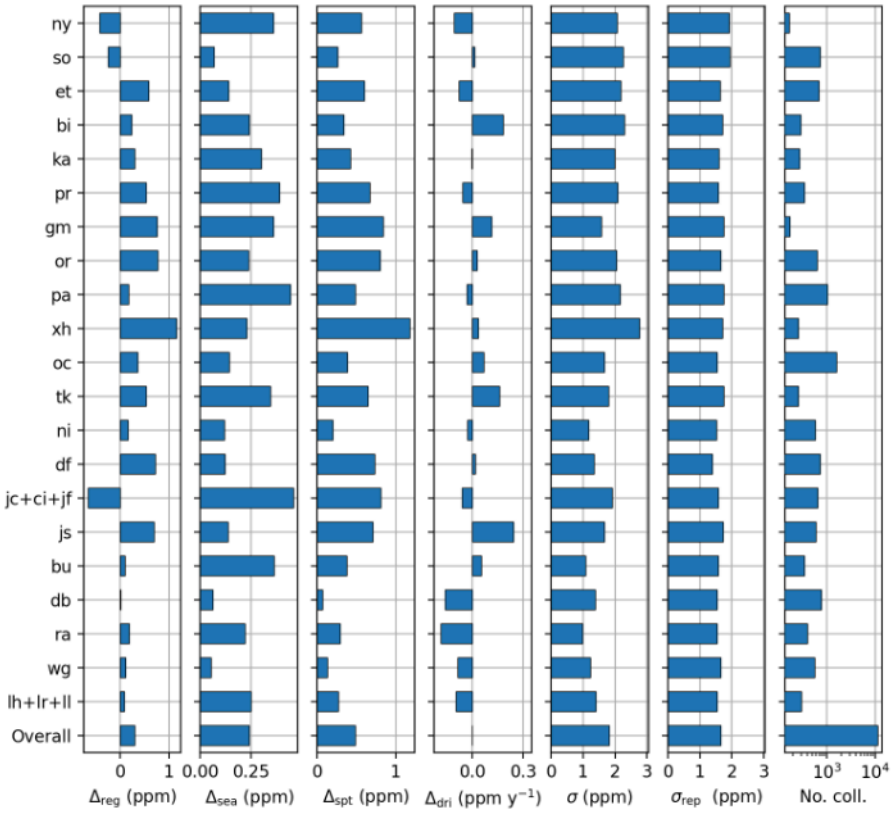


Figure F2. Bar plot of the XCO_2 L3 EMMA validation results shown in Table F2.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

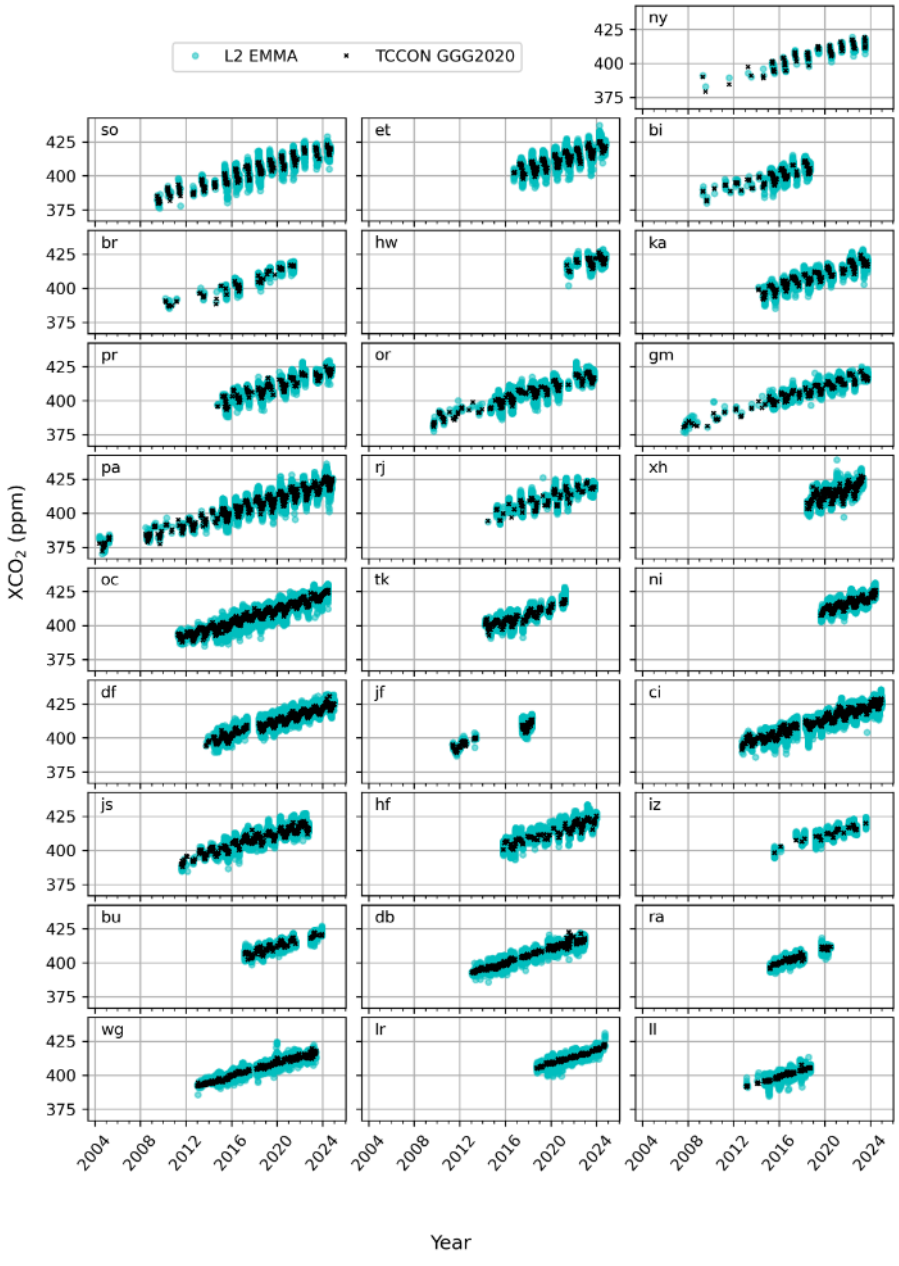


Figure F3. Co-located XCO_2 L2 EMMA and TCCON observations. Each plot shows the collocations for a different station.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

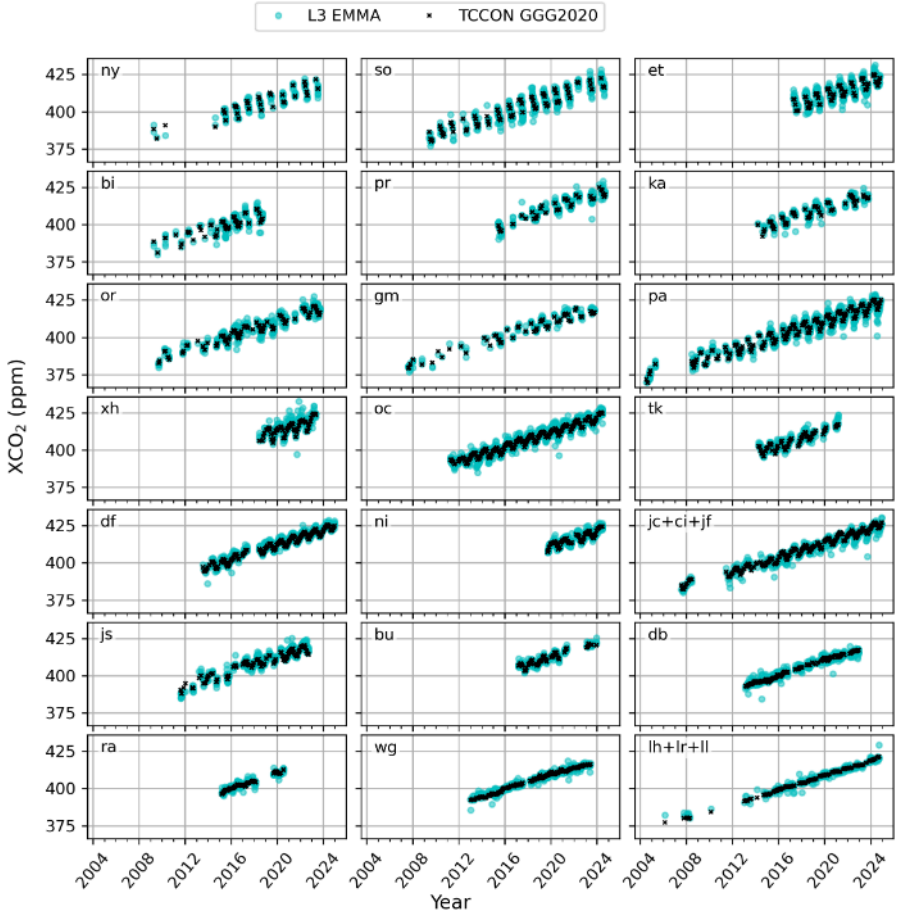


Figure F4. Co-located XCO_2 L3 EMMA and TCCON observations. Each plot shows the collocations for a different station.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table F3. Analogous to F1 for the validation results of the EMMA v5.1 XCH_4 L2 dataset.

Station	Δ_{reg} (ppb)	Δ_{sea} (ppb)	Δ_{spt} (ppb)	Δ_{dri} (ppb y^{-1})	σ (ppb)	σ_{rep} (ppb)	No. coll.
eu	8.30	9.06	12.28	4.03	17.08	16.25	68921
ny	-4.71	3.81	6.06	-1.43	20.42	20.20	90264
so	-5.72	3.49	6.70	-0.11	17.87	19.99	310707
et	-0.87	6.18	6.24	-0.80	18.57	17.81	623345
bi	-8.42	3.01	8.95	0.50	13.24	14.12	35725
br	-9.31	0.82	9.35	0.33	12.92	14.51	113924
hw	-3.25	2.79	4.28	0.68	13.66	16.99	110165
ka	-10.58	1.54	10.69	1.28	12.71	15.49	332768
pr	-10.39	1.46	10.49	1.45	12.84	15.14	511688
or	-7.38	0.86	7.43	1.35	12.24	14.86	495324
gm	-1.68	3.00	3.44	0.36	14.68	17.33	128842
pa	-6.14	4.43	7.57	-0.05	16.83	17.12	707961
rj	-1.25	3.05	3.30	1.03	15.07	17.45	21377
xh	1.65	7.17	7.36	0.26	18.12	14.45	679932
oc	-10.03	0.91	10.07	0.08	12.62	13.93	1601270
tk	-6.46	2.95	7.10	1.78	11.66	14.92	23363
ni	-4.73	0.60	4.77	0.96	12.52	15.07	230657
df	-4.04	3.81	5.55	0.39	12.46	13.70	1030927
jf	-14.91	1.17	14.95	0.37	12.76	13.03	3519
ci	-8.84	4.56	9.95	0.54	15.58	13.98	691213
js	-5.19	2.47	5.74	1.24	13.34	15.92	59376
hf	-5.52	1.52	5.73	0.86	15.13	15.55	317486
iz	17.02	1.53	17.09	1.42	12.12	14.38	145
bu	-8.09	2.13	8.36	-0.66	11.68	16.97	61447
db	-11.65	1.72	11.78	-0.90	12.94	14.42	323450
ra	-14.58	4.94	15.40	-2.45	11.23	16.95	12227
wg	-10.36	2.62	10.69	0.68	16.52	17.47	103876
lr	-11.56	3.11	11.97	-0.19	12.47	16.30	90301
ll	-6.92	3.55	7.77	-0.91	11.60	16.71	5327
lh	-16.95	7.97	18.74	0.51	84.11	77.16	192
Overall	-6.09 ± 6.63	3.21	7.37	0.42 ± 1.14	20.87	21.11	8785719

F2 Merged XCH_4 products

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table F4. Analogous to F1 for the validation results of the EMMA v5.1 XCH_4 L3 dataset.

Station	Δ_{reg} (ppb)	Δ_{sea} (ppb)	Δ_{spt} (ppb)	Δ_{dri} (ppb y^{-1})	σ (ppb)	σ_{rep} (ppb)	No. coll.
ny	-4.70	3.52	5.87	-1.30	13.34	14.03	1149
so	-7.13	2.18	7.46	-0.09	11.42	12.50	1844
et	-1.26	3.30	3.53	-0.86	8.52	10.74	1845
bi	-6.44	1.82	6.70	-0.64	10.43	10.77	666
ka	-11.68	2.60	11.97	0.20	7.24	9.01	475
pr	-9.09	2.28	9.37	0.60	7.52	8.48	638
gm	-0.98	3.73	3.86	0.29	8.61	11.08	492
or	-6.14	1.41	6.30	0.12	7.67	9.80	1517
pa	-8.00	4.26	9.06	-0.08	12.34	14.04	2186
rj	-1.23	1.22	1.74	0.86	8.21	10.70	187
xh	0.74	9.12	9.15	0.73	8.15	8.23	448
oc	-8.77	0.53	8.78	-0.19	6.74	7.94	2085
tk	-11.44	4.26	12.21	0.92	8.59	10.66	269
ni	-4.54	0.79	4.61	0.61	6.73	8.56	668
df	-3.88	3.40	5.16	0.66	4.05	8.03	710
jc+ci+jf	-10.29	2.99	10.71	0.55	9.50	12.53	871
js	-6.58	1.71	6.79	1.21	10.42	11.89	833
bu	-5.48	0.93	5.56	-0.45	8.17	12.74	432
db	-7.53	0.93	7.59	0.20	7.74	9.90	828
ra	-8.37	2.83	8.83	-0.68	8.59	11.48	138
wg	-9.25	2.67	9.63	-0.22	11.15	12.05	579
lh+lr+ll	-8.65	2.47	8.99	-0.08	8.02	14.09	386
Overall	-6.39 ± 3.37	2.68	4.31	0.11 ± 0.62	9.01	11.05	19246

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

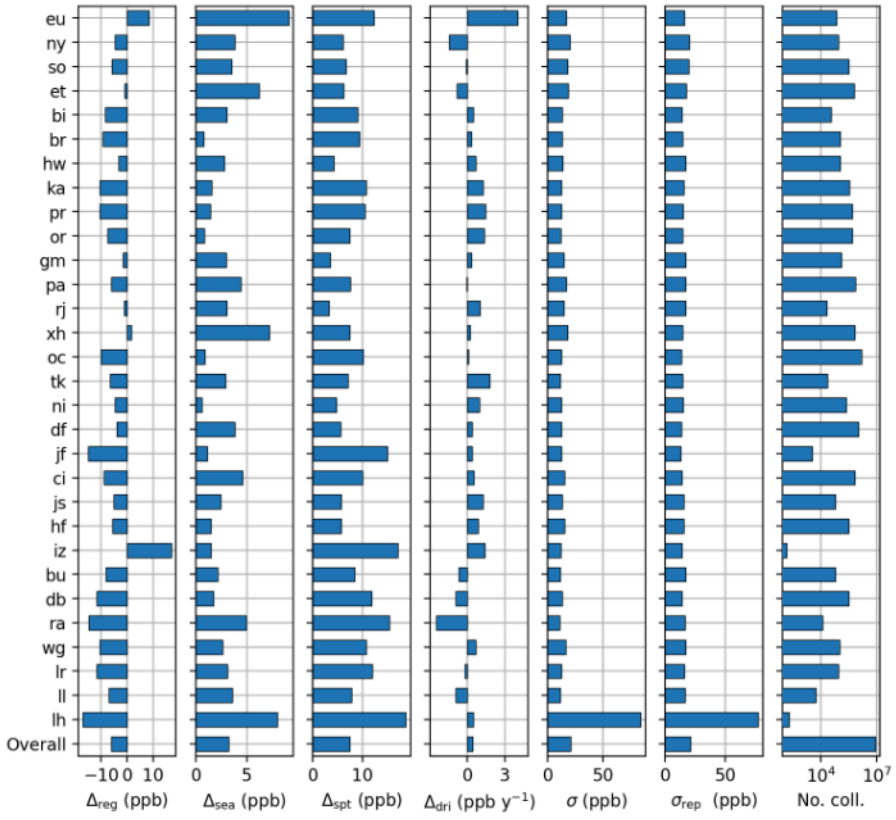


Figure F5. Bar plot of the XCH_4 L2 EMMA validation results shown in Table F3.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

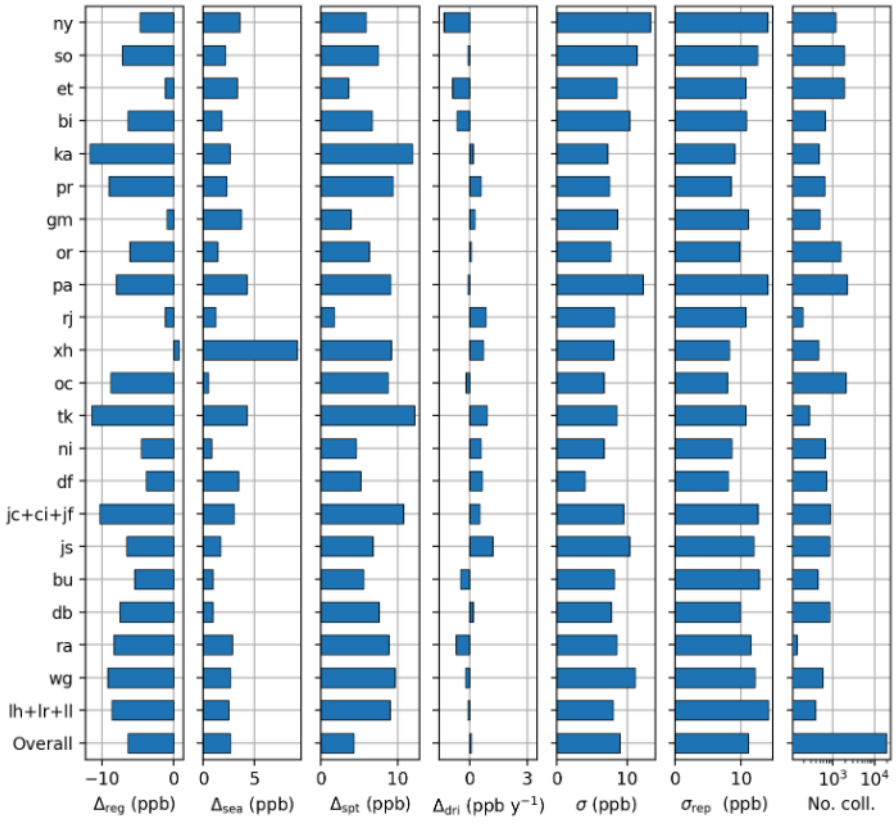


Figure F6. Bar plot of the XCH_4 L3 EMMA validation results shown in Table F4.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

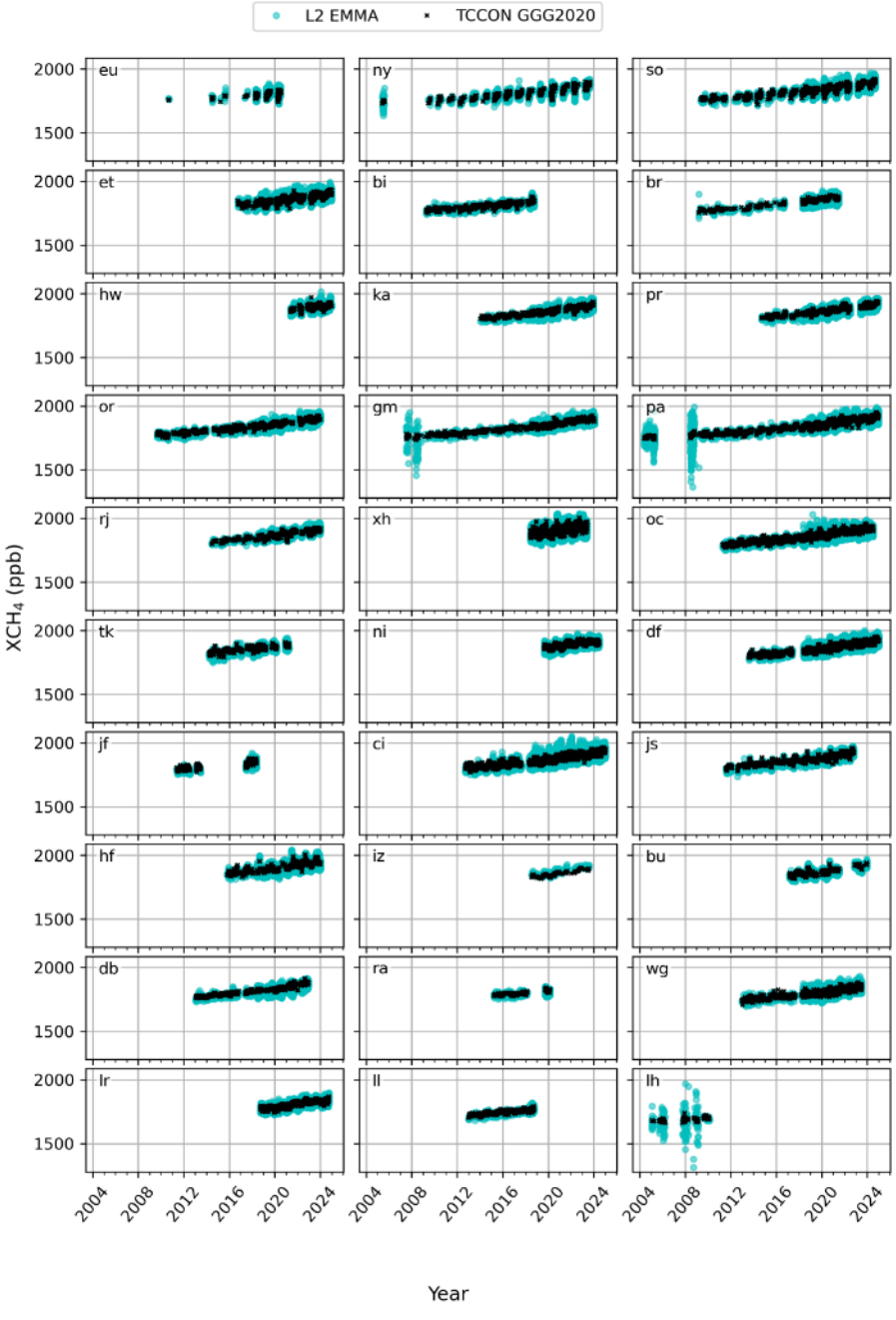


Figure F7. Co-located XCH_4 L2 EMMA and TCCON observations. Each plot shows the collocations for a different station.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

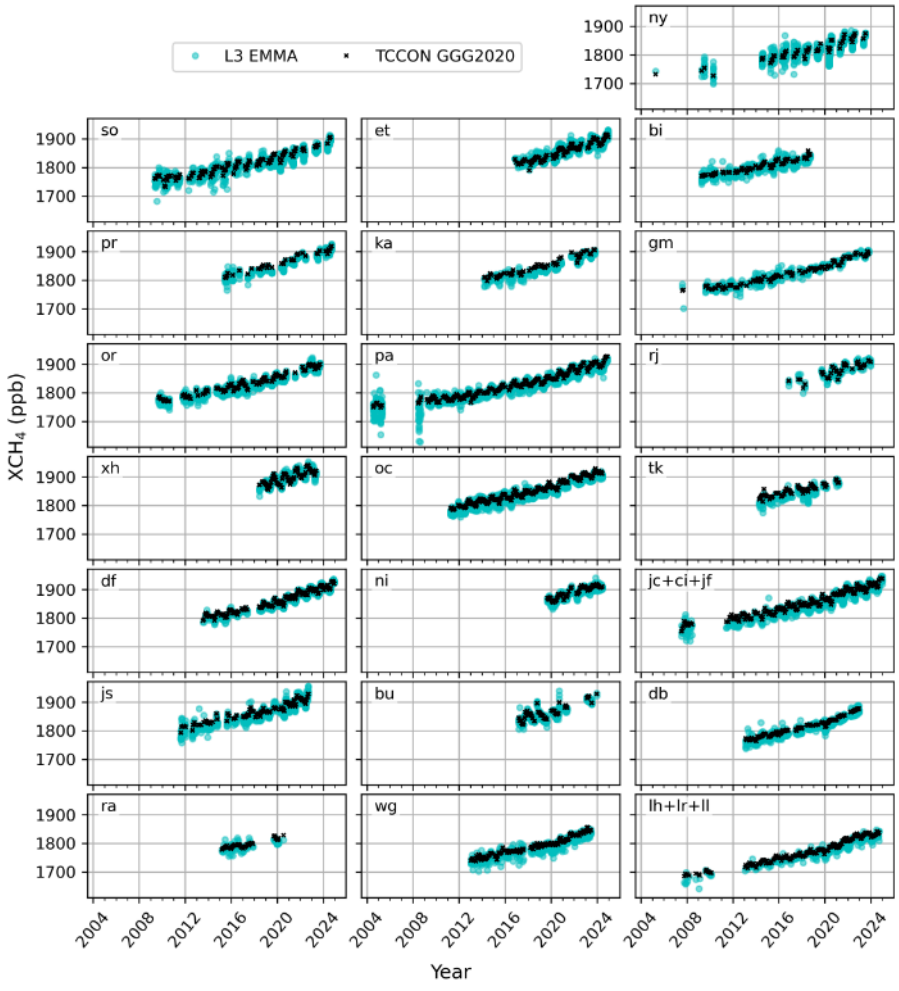


Figure F8. Co-located XCH_4 L3 EMMA and TCCON observations. Each plot shows the collocations for a different station.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix G: Comparison to EMMA version 4.6

This section presents a comparison between the L3 datasets generated with the current EMMA v5.1 and the previous EMMA v4.6, for both XCO_2 (Sect. G1) and XCH_4 (Sect. G2).

G1 Comparison between L3 EMMA v5.1 and v4.6 XCO_2 products

Figures G1 and G2 show examples of global maps of the months of May 2020 and July 2020, respectively, for the L3 EMMA v5.1 product in its original spatial resolution, re-gridded to match the EMMA v4.6 spatial resolution, and the L3 v4.6 product. The plots for both versions of the datasets present similar features, such as the pronounced north-south gradient, with high XCO_2 values in the northern hemisphere at the beginning of the growing season in Fig. G1 and the inversion of this gradient shown in Fig. G2. Overall, both figures show higher coverage of the (re-gridded) v5.1 product compared to its v4.6 counterpart.

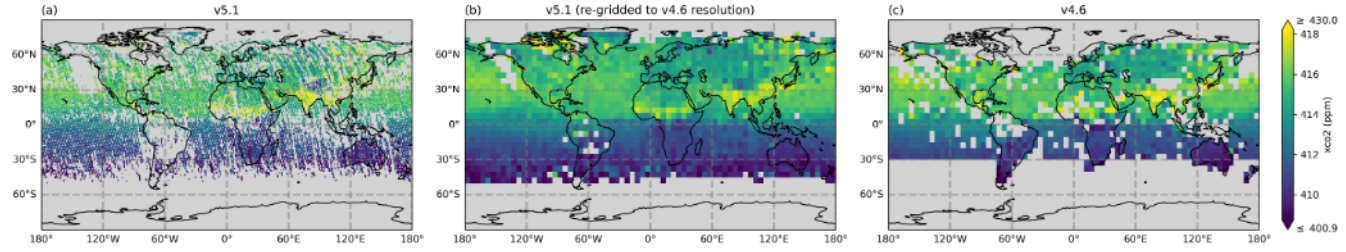


Figure G1. L3 XCO_2 global map for May 2020 showing (a) v5.1 in its original 1°x1° spatial resolution, (b) v5.1 re-gridded to match the v4.6 5°x5° resolution, and (c) v4.6 in its original 5°x5° spatial resolution.

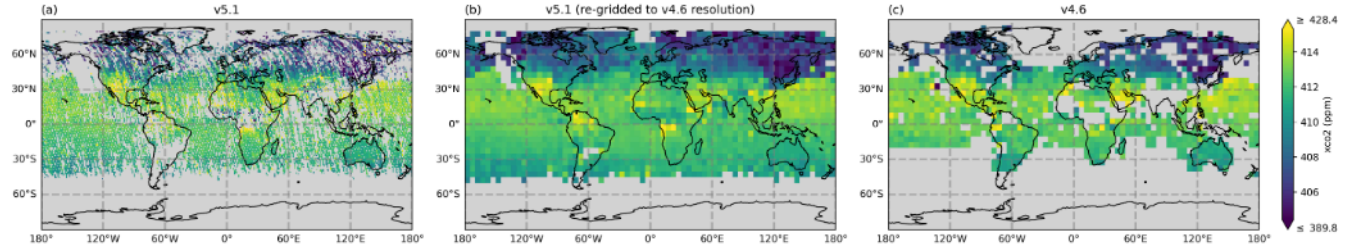


Figure G2. Analogous to G1 for July 2020.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Figure G3 shows the re-gridded L3 EMMA v5.1 product against the v4.6 analogous. In the plot we can observe that the data concentrates around the one-to-one line. Both the slope and the coefficient of determination resulting from the fit are very close to one, showing good agreement between both datasets. Both datasets agree almost perfectly for the lowest XCO_2 values, i.e. below approximately 380 ppm. These values correspond to the period in which SCIAMACHY BESD is the reference and the only available product, so both datasets are almost identical in this period because there is no ensemble to compute the empirical spatio-temporal adjustment from.

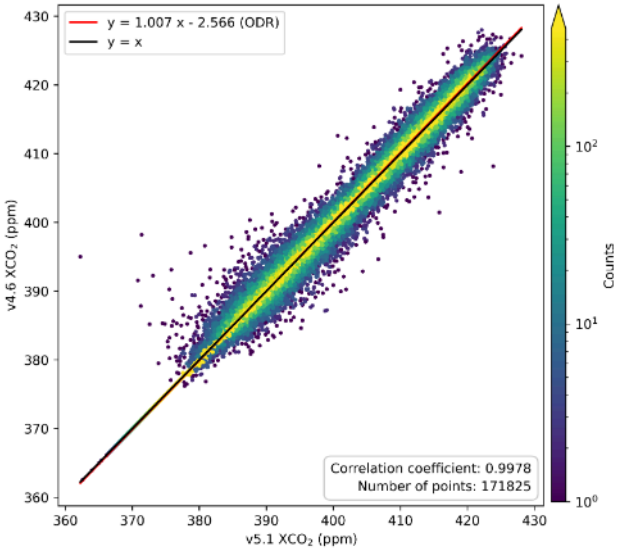


Figure G3. Scatter plot of L3 v4.6 XCO_2 against L3 v5.1 XCO_2 (re-gridded to the v4.6 resolution). Each point represents a grid box (5°x5°) containing data in both datasets. The data was fitted to a straight line, yielding both a slope and a coefficient of determination (r^2) very close to one. Marker colors indicate the number of counts (note the logarithmic scale of the colorbar) in a 2-D histogram, showing that the data concentrates near the one-to-one line (black).

Figure G4 shows the time series for three latitude bands for both the L3 EMMA v5.1 and v4.6 products. For each month and latitude band, the mean XCO_2 values are shown. Only grid boxes with XCO_2 data from both v4.6 and the re-gridded v5.1 products were used for the plot. Both datasets exhibit very similar behaviours in the time series.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

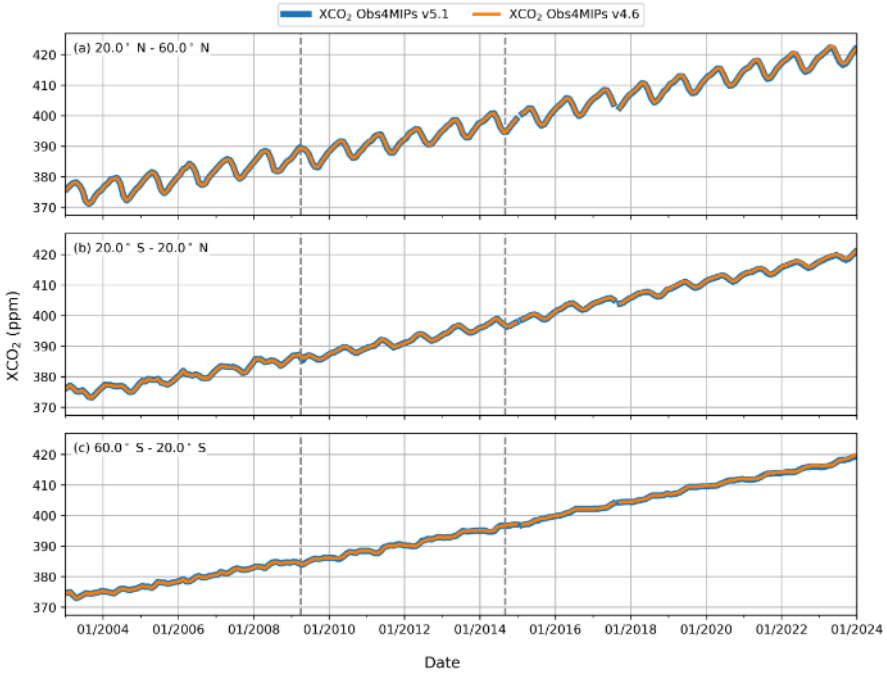


Figure G4. Time series for L3 XCO_2 v5.1 (blue) and superposed L3 XCO_2 v4.6 (orange). The dashed gray vertical lines mark the changes in the reference product, i.e. from the beginning of the dataset until the end of March 2009, SCIAMACHY BESD was used as the reference; from April 2009 until the end of August 2014, GOSAT FOCAL was the reference; and from September 2014 until the end of the dataset period, the reference product was OCO-2 FOCAL. Three latitude bands were used, shown in the corresponding panels: panel (a) shows the latitude band from 20° N - 60° N, panel (b) shows 20° S to 20° N, and panel (c) shows 60° S to 20° S.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

G2 Comparison between L3 EMMA v5.1 and v4.6 XCH_4 products

Figures G5 and G6, which are analogous to Figs. G1 and G2, show examples of global maps of the months of April 2020 and October 2020 for the comparison between the (re-gridded) L3 XCH_4 EMMA v5.1 product and the v4.6 version. Again, the figures illustrate similar features from both dataset versions, such as the spatial latitudinal gradients or regionally enhanced values, as well as an improved spatial coverage of the re-gridded v5.1 dataset with respect to v4.6.

Figure G7 shows the re-gridded XCH_4 L3 EMMA v5.1 against the v4.6 product. Analogous to Fig. G3, most data lie near the one-to-one line, and the linear fit yields both a slope and correlation coefficient very close to one. Also in this case, the lowest XCH_4 values (below approximately 1630 ppb) in both datasets agree almost perfectly due to the lack of ensemble products for the empirical spatio-temporal adjustment during the period in which SCIAMACHY WFMD is the reference.

Finally, the time series for three latitude bands for both the L3 EMMA v5.1 and v4.6 products are illustrated in Fig. G8, analogous to Fig. G4, showing again similar behaviours in both versions.

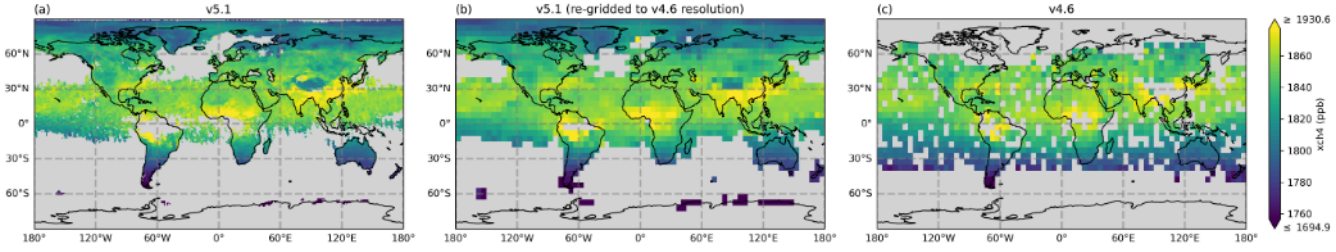


Figure G5. Analogous to G1 for L3 XCH_4 and April 2020.

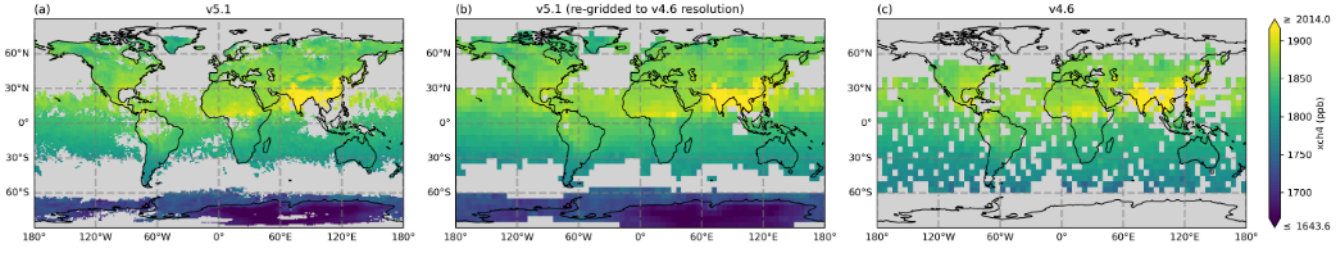


Figure G6. Analogous to G1 for L3 XCH_4 and October 2020.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

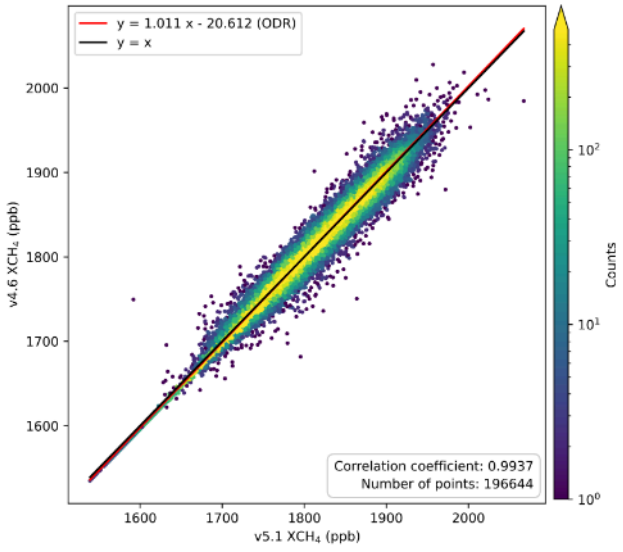


Figure G7. Analogous to G3 for the L3 v4.6 XCH_4 against L3 v5.1 XCH_4 datasets.

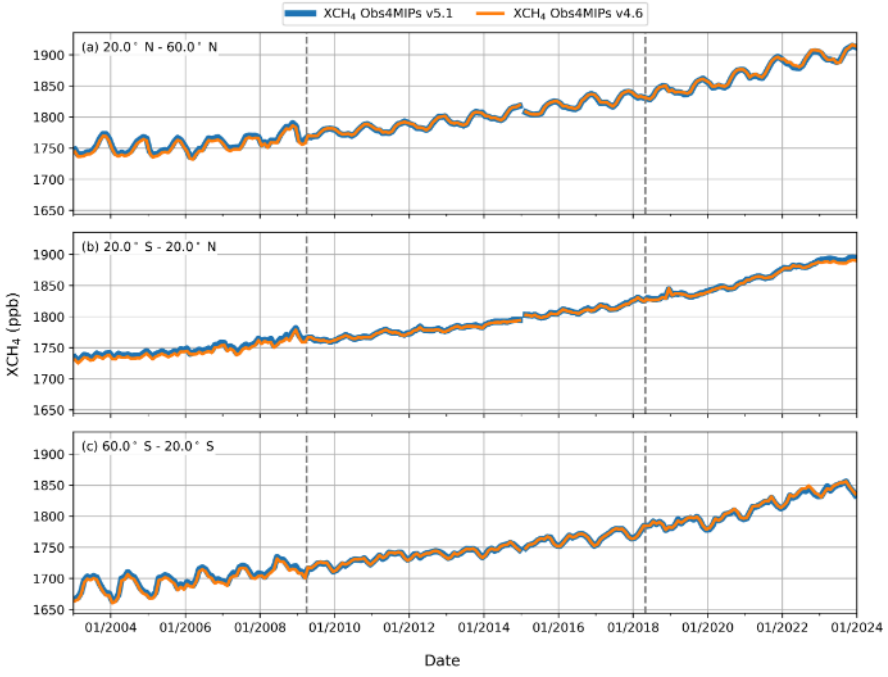


Figure G8. Analogous to G4 for the L3 XCH_4 datasets. The dashed gray vertical lines mark the changes in the reference product, i.e. from the beginning of the dataset until the end of March 2009, SCIAMACHY WFMD was used as the reference; from April 2009 until the end of April 2018, GOSAT FOCAL PR was the reference; and from May 2018 until the end of the dataset period, the reference product was TROPOMI WFMD.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Author contributions. BFA generated the EMMA v5.1 XCO2 and XCH4 L2 and L3 datasets and wrote the manuscript with help of MR and MB. BFA, MR and MB designed the EMMA v5.1 algorithm. SN, OS, MH, RJP,AGB, YY and YS provided input datasets. All authors contributed to significantly improve the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank the following space agencies for making available satellite Level 1 (L1) input data employed for the generation of the L2 data used as a reference by the EMMA algorithm: ESA/DLR for SCIAMACHY L1 data, JAXA for GOSAT Level 1 data through ESA, NASA for the OCO-2 L1 data product and ESA for the GOSAT and TROPOMI L1 products. We thank NIES for providing the operational GOSAT and GOSAT-2 XCO2 and XCH4 Level 2 products; the NASA team for the GOSAT, OCO-2 and OCO-3 ACOS Level 2 XCO2 products; and SRON and the University of Leicester for the GOSAT and GOSAT-2 RemoTeC and UoL products, respectively.

This publication contains modified Copernicus Sentinel data (2018-2024). Sentinel-5 Precursor is an ESA mission implemented on behalf of the European Commission. The TROPOMI payload is a joint development by ESA and the Netherlands Space Office (NSO). The Sentinel-5 Precursor ground-segment development has been funded by ESA and with national contributions from The Netherlands, Germany, and Belgium.

We acknowledge TCCON for making available the XCO_2 and XCH_4 data used in this study. TCCON data were obtained from the TCCON Data Archive, hosted by CaltechDATA. We are grateful to CAMS for their global inversion-optimised greenhouse gas fluxes and concentrations, obtained from the Copernicus Atmosphere Monitoring Service (CAMS) Atmosphere Data Store (ADS), which we used for the SLIM climatologies. We also thank NOAA for making available information on the annual growth rates, which we used for adjusting the annual growth of the SLIM climatologies.

Financial support. The research leading to the presented results has received funding from the European Union though the Copernicus Climate Change Service (C3S), implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF), under the project C3S2_313a with contract number D/565/67345031, from the European Space Agency (ESA) via the projects GHG-CCI+ and SMART-CH4 (ESA contract nos. 4000126450/19/I-NB and 4000142730/23/I-NS), from the German Federal Ministry of Research, Technology and Space (BMFTR) within its project ITMS via grant 01 LK2103A, and from the state and the University of Bremen. The TROPOMI WFMD retrievals presented here were performed on HPC facilities of the IUP, University of Bremen, funded under DFG/FUGG grant nos. INST 144/379-1 and INST 515 144/493-1. RJP was supported by the Natural Environment Research Council through the UK EO Climate Information Service (NE/X019071/1), the UK National Centre for Earth Observation (Grant: NE/W004895/1) and a UKRI Future Leaders Fellowship (Grant: MR/X033139/1).

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

References

Balasus, N., Jacob, D. J., Lorente, A., Maasakkers, J. D., Parker, R. J., Boesch, H., Chen, Z., Kelp, M. M., Nesser, H., and Varon, D. J.: A Blended TROPOMI+GOSAT Satellite Data Product for Atmospheric Methane Using Machine Learning to Correct Retrieval Biases, Atmospheric Measurement Techniques, 16, 3787–3807, https://doi.org/10.5194/amt-16-3787-2023, 2023.

Barr, A. G., Landgraf, J., Martinez-Velarte, M., Vrekoussis, M., Sussmann, R., Morino, I., Strong, K., Zhou, M., Velazco, V. A., Ohyama, H., Warneke, T., Hase, F., and Borsdorff, T.: Five years of GOSAT-2 retrievals with RemoTeC: XCO2 and XCH4 data products with quality filtering by machine learning, Atmospheric Measurement Techniques, 18, 6093–6123, https://doi.org/10.5194/amt-18-6093-2025, 2025.

Boesch, H. and Di Noia, A.: Algorithm Theoretical Basis Document (ATBD) - ANNEX A for products CO2_GOS_OCFP (v7.3), CH4_GOS_OCFP (v7.3) and CH4_GOS_OCPR (v9.0) (CDR6, 2009-2021), accessed 19-01-2026, 2021.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noel, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.: SCIAMACHY: Mission Objectives and Measurement Modes, Journal of the Atmospheric Sciences, 56, 127–150, https://doi.org/10.1175/1520-0469(1999)056%3C0127:SMOAMM%3E2.0.CO;2, 1999.

Buchwitz, M., Reuter, M., Schneising, O., Noël, S., Gier, B., Bovensmann, H., Burrows, J. P., Boesch, H., Anand, J., Parker, R. J., Somkuti, P., Detmers, R. G., Hasekamp, O. P., Aben, I., Butz, A., Kuze, A., Suto, H., Yoshida, Y., Crisp, D., and O'Dell, C.: Computation and analysis of atmospheric carbon dioxide annual mean growth rates from satellite observations during 2003–2016, Atmospheric Chemistry and Physics, 18, 17 355–17 370, https://doi.org/10.5194/acp-18-17355-2018, 2018.

Burrows, J. P., Hoelzle, E., Goede, A. P. H., Visser, H., and Fricke, W.: SCIAMACHY—scanning imaging absorption spectrometer for atmospheric chartography, Acta Astronautica, 35, 445–451, https://doi.org/10.1016/0094-5765(94)00278-T, 1995.

Buschmann, M., Petri, C., Palm, M., Warneke, T., and Notholt, J.: TCCON data from Ny-Ã…lesund, Svalbard (NO), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.NYALESUND01.R0, 2022.

Butz, A., Guerlet, S., Hasekamp, O., Schepers, D., Galli, A., Aben, I., Frankenberg, C., Hartmann, J.-M., Tran, H., Kuze, A., Keppel-Aleks, G., Toon, G., Wunch, D., Wennberg, P., Deutscher, N., Griffith, D., Macatangay, R., Messerschmidt, J., Notholt, J., and Warneke, T.: Toward accurate CO2 and CH4 observations from GOSAT: GOSAT CO2 AND CH4 VALIDATION, Geophysical Research Letters, 38, n/a–n/a, https://doi.org/10.1029/2011GL047888, 2011.

C3S and WMO: European State of the Climate – Report 2024, Copernicus Climate Change Service (C3S) and World Meteorological Organization (WMO), https://library.wmo.int/records/item/69475-state-of-the-climate-in-europe-2024, 2025.

CAMS: CAMS global inversion-optimised greenhouse gas fluxes and concentrations, https://doi.org/10.24381/ed2851d2, copernicus Atmosphere Monitoring Service. Accessed 2026-01-07, 2020.

Chevallier, F.: On the parallelization of atmospheric inversions of CO2 surface fluxes within a variational framework, Geoscientific Model Development, 6, 783–790, https://doi.org/10.5194/gmd-6-783-2013, 2013.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., and Ciais, P.: Inferring CO2 sources and sinks from satellite observations: Method and application to TOVS data, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/10.1029/2005JD006390, 2005.

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morguí, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen,

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

A. T., Wofsy, S., and Worthy, D.: CO2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/10.1029/2010JD013887, 2010.

Cogan, A. J., Boesch, H., Parker, R. J., Feng, L., Palmer, P. I., Blavier, J.-F. L., Deutscher, N. M., Macatangay, R., Notholt, J., Roehl, C., Warneke, T., and Wunch, D.: 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, https://doi.org/10.1029/2012JD018087, 2012.

Crisp, D., Pollock, H. R., Rosenberg, R., Chapsky, L., Lee, R. A. M., Oyafuso, F. A., Frankenberg, C., O'Dell, C. W., Bruegge, C. J., Doran, G. B., Eldering, A., Fisher, B. M., Fu, D., Gunson, M. R., Mandrake, L., Osterman, G. B., Schwandner, F. M., Sun, K., Taylor, T. E., Wennberg, P. O., and Wunch, D.: The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products, Atmospheric Measurement Techniques, 10, 59–81, https://doi.org/10.5194/amt-10-59-2017, 2017.

Crisp, D., O'Dell, C., Eldering, A., Fisher, B., Oyafuso, F., Payne, V., Drouin, B., Toon, G., Laughner, J., Somkuti, P., McGarragh, G., Merrelli, A., Nelson, R., Gunson, M., Frankenberg, C., Osterman, G., Boesch, H., Brown, L., Castano, R., Christi, M., Connor, B., McDuffie, J., Miller, C., Natraj, V., O'Brien, D., Polonski, I., Smyth, M., Thompson, D., and Granat, R.: LEVEL 2 FULL PHYSICS ALGORITHM Theoretical Basis Document, 2021.

De Mazière, M., Sha, M. K., Desmet, F., Hermans, C., Scolas, F., Kumps, N., Zhou, M., Metzger, J.-M., Duflot, V., and Cammas, J.-P.: TCCON data from Réunion Island (RE), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.REUNION01.R0, 2022.

Detmers, R.: Algorithm Theoretical Basis Document (ATBD) – ANNEX B for products CO2_GOS_SRFP & CH4_GOS_SRFP, Copernicus Climate Change Service (C3S) project on satellite-derived Essential Climate Variable (ECV) Greenhouse Gases (CO2 and CH4) data products (project C3S_312a_Lot6), Version 1 (21/08/2017), 2017a.

Detmers, R.: Algorithm Theoretical Basis Document (ATBD) – ANNEX C for product CH4_GOS_SRPR, Copernicus Climate Change Service (C3S) project on satellite-derived Essential Climate Variable (ECV) Greenhouse Gases (CO2 and CH4) data products (project C3S_312a_Lot6), Version 1 (21/08/2017), 2017b.

Deutscher, N. M., Griffith, D. W., Paton-Walsh, C., Jones, N. B., Velazco, V. A., Wilson, S. R., Macatangay, R. C., Kettlewell, G. C., Buchholz, R. R., Riggenbach, M. O., Bukosa, B., John, S. S., Walker, B. T., and Nawaz, H.: TCCON data from Wollongong (AU), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.WOLLONGONG01.R0, 2023a.

Deutscher, N. M., Griffith, D. W., Paton-Walsh, C., Velazco, V. A., Wennberg, P. O., Blavier, J.-F., Washenfelder, R. A., Yavin, Y., Keppel-Aleks, G., Toon, G. C., Jones, N. B., Kettlewell, G. C., Connor, B. J., Macatangay, R. C., Wunch, D., Roehl, C., and Bryant, G. W.: TCCON data from Darwin (AU), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.DARWIN01.R0, 2023b.

Dowd, E., Wilson, C., Chipperfield, M. P., Gloor, E., Manning, A., and Doherty, R.: Decreasing seasonal cycle amplitude of methane in the northern high latitudes being driven by lower-latitude changes in emissions and transport, Atmospheric Chemistry and Physics, 23, 7363–7382, https://doi.org/10.5194/acp-23-7363-2023, 2023.

Dubey, M. K., Henderson, B. G., Allen, N. T., Blavier, J.-F., Roehl, C. M., and Wunch, D.: TCCON data from Manaus (BR), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.MANAUS01.R0, 2022.

Eldering, A., Taylor, T. E., O'Dell, C. W., and Pavlick, R.: The OCO-3 mission: measurement objectives and expected performance based on 1 year of simulated data, Atmospheric Measurement Techniques, 12, 2341–2370, https://doi.org/10.5194/amt-12-2341-2019, 2019.

ESA: GLO-90: Copernicus Global Digital Elevation Model, european Space Agency, 2024.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

ESA CCI GHG project team: ESA Greenhouse Gases Climate Change Initiative (GHG_cci): Column-averaged CH4 from SCIAMACHY generated with the WFMD algorithm (CH4_SCI_WFMD), version 4.0, https://catalogue.ceda.ac.uk/uuid/aa09603e91b44f3cb1573c9dd415e8a8/, 2018.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Landschützer, P., Le Quéré, C., Li, H., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Arneth, A., Arora, V., Bates, N. R., Becker, M., Bellouin, N., Berghoff, C. F., Bittig, H. C., Bopp, L., Cadule, P., Campbell, K., Chamberlain, M. A., Chandra, N., Chevallier, F., Chini, L. P., Colligan, T., Decayeux, J., Djeutchouang, L. M., Dou, X., Duran Rojas, C., Enyo, K., Evans, W., Fay, A. R., Feely, R. A., Ford, D. J., Foster, A., Gasser, T., Gehlen, M., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, O., Harris, I., Hefner, M., Heinke, J., Hurtt, G. C., Iida, Y., Ilyina, T., Jacobson, A. R., Jain, A. K., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Kato, E., Keeling, R. F., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Lan, X., Lauvset, S. K., Lefèvre, N., Liu, Z., Liu, J., Ma, L., Maksyutov, S., Marland, G., Mayot, N., McGuire, P. C., Metzl, N., Monacci, N. M., Morgan, E. J., Nakaoka, S.-I., Neill, C., Niwa, Y., Nützel, T., Olivier, L., Ono, T., Palmer, P. I., Pierrot, D., Qin, Z., Resplandy, L., Roobaert, A., Rosan, T. M., Rödenbeck, C., Schwinger, J., Smallman, T. L., Smith, S. M., Sospedra-Alfonso, R., Steinhoff, T., Sun, Q., Sutton, A. J., Séférian, R., Takao, S., Tatebe, H., Tian, H., Tilbrook, B., Torres, O., Tourigny, E., Tsujino, H., Tubiello, F., van der Werf, G., Wanninkhof, R., Wang, X., Yang, D., Yang, X., Yu, Z., Yuan, W., Yue, X., Zaehle, S., Zeng, N., and Zeng, J.: Global Carbon Budget 2024, Earth System Science Data, 17, 965–1039, https://doi.org/10.5194/essd-17-965-2025, 2025.

Fuentes Andrade, B., Reuter, M., and Buchwitz, M.: Merged L2 methane data from 2003 to 2024 derived from satellite observations, https://doi.org/10.5281/zenodo.18608583, 2026a.

Fuentes Andrade, B., Reuter, M., and Buchwitz, M.: Merged L2 carbon dioxide data from 2003 to 2024 derived from satellite observations, https://doi.org/10.5281/zenodo.18610669, 2026b.

García, O. E., Schneider, M., Herkommer, B., Gross, J., Hase, F., Blumenstock, T., and SepÃºlveda, E.: TCCON data from Izana (ES), Release GGG2020.R1, https://doi.org/10.14291/TCCON.GGG2020.IZANA01.R1, 2022.

GCOS-154, G. C. O. S. G.: Systematic Observation Requirements for Satellite-based Data Products for Climate, https://library.wmo.int/idurl/4/48411, 2011.

GCOS-200, G. C. O. S. G.: The Global Observing System for Climate: Implementation Needs, https://library.wmo.int/idurl/4/55469, 2016.

Gier, B. K., Buchwitz, M., Reuter, M., Cox, P. M., Friedlingstein, P., and Eyring, V.: Spatially resolved evaluation of Earth system models with satellite column-averaged CO2, Biogeosciences, 17, 6115–6144, https://doi.org/10.5194/bg-17-6115-2020, 2020.

Gulev, S. K., Thorne, P. W., Ahn, J., Dentener, F. J., Domingues, C. M., Gerland, S., Gong, D., Kaufman, D. S., Nnamchi, H. C., Quaas, J., et al.: Changing state of the climate system, chap. 2, pp. 287–422, Cambridge University Press, 2021.

Hase, F., Herkommer, B., GroÃŸ, J., Blumenstock, T., Kiel, M., and Dohe, S.: TCCON data from Karlsruhe (DE), Release GGG2020.R2, https://doi.org/10.14291/TCCON.GGG2020.KARLSRUHE01.R2, 2024.

Huber, P. J. and Ronchetti, E. M.: Robust statistics, John Wiley & Sons, Inc., 2009.

Imasu, R., Matsunaga, T., Nakajima, M., Yoshida, Y., Shiomi, K., Morino, I., Saitoh, N., Niwa, Y., Someya, Y., Oishi, Y., Hashimoto, M., Noda, H., Hikosaka, K., Uchino, O., Maksyutov, S., Takagi, H., Ishida, H., Nakajima, M. Y., Nakajima, T., and Shi, C.: Greenhouse Gases Observing SATellite 2 (GOSAT-2): Mission Overview, Progress in Earth and Planetary Science, 10, 33, https://doi.org/10.1186/s40645-023-00562-2, 2023.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Iraci, L. T., Podolske, J. R., Hillyard, P. W., Roehl, C., Wennberg, P. O., Blavier, J.-F., Landeros, J., Allen, N., Wunch, D., Zavaleta, J., Quigley, E., Osterman, G. B., Barrow, E., and Barney, J.: TCCON data from Indianapolis (US), Release GGG2020.R1, https://doi.org/10.14291/TCCON.GGG2020.INDIANAPOLIS01.R1, 2022a.

Iraci, L. T., Podolske, J. R., Roehl, C., Wennberg, P. O., Blavier, J.-F., Allen, N., Wunch, D., and Osterman, G. B.: TCCON data from Edwards (US), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.EDWARDS01.R0, 2022b.

Keeling, R. F. and Graven, H. D.: Insights from time series of atmospheric carbon dioxide and related tracers, Annual Review of Environment and Resources, 46, 85–110, 2021.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J. G., Dlugokencky, E. J., Bergamaschi, P., Bergmann, D., Blake, D. R., Bruhwiler, L., Cameron-Smith, P., Castaldi, S., Chevallier, F., Feng, L., Fraser, A., Heimann, M., Hodson, E. L., Houweling, S., Josse, B., Fraser, P. J., Krummel, P. B., Lamarque, J.-F., Langenfelds, R. L., Le Quéré, C., Naik, V., O'Doherty, S., Palmer, P. I., Pison, I., Plummer, D., Poulter, B., Prinn, R. G., Rigby, M., Ringeval, B., Santini, M., Schmidt, M., Shindell, D. T., Simpson, I. J., Spahni, R., Steele, L. P., Strode, S. A., Sudo, K., Szopa, S., van der Werf, G. R., Voulgarakis, A., van Weele, M., Weiss, R. F., Williams, J. E., and Zeng, G.: Three decades of global methane sources and sinks, Nature Geoscience, 6, 813–823, https://doi.org/10.1038/ngeo1955, 2013.

Kivi, R., Heikkinen, P., and KyrÃ¶, E.: TCCON data from SodankylÃ¤ (FI), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.SODANKYLA01.R0, 2022.

Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T.: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, Applied Optics, 48, 6716–6733, https://doi.org/10.1364/AO.48.006716, 2009.

Kuze, A., Suto, H., Shiomi, K., Kawakami, S., Tanaka, M., Ueda, Y., Deguchi, A., Yoshida, J., Yamamoto, Y., Kataoka, F., Taylor, T. E., and Buijs, H. L.: Update on GOSAT TANSO-FTS performance, operations, and data products after more than 6 years in space, Atmospheric Measurement Techniques, 9, 2445–2461, https://doi.org/10.5194/amt-9-2445-2016, 2016.

Lauer, A., Eyring, V., Righi, M., Buchwitz, M., Defourny, P., Evaldsson, M., Friedlingstein, P., de Jeu, R., de Leeuw, G., Loew, A., Merchant, C. J., Müller, B., Popp, T., Reuter, M., Sandven, S., Senftleben, D., Stengel, M., Van Roozendael, M., Wenzel, S., and Willén, U.: Benchmarking CMIP5 models with a subset of ESA CCI Phase 2 data using the ESMValTool, Remote Sensing of Environment, 203, 9–39, https://doi.org/10.1016/j.rse.2017.01.007, 2017.

Laughner, J. L., Toon, G. C., Mendonca, J., Petri, C., Roche, S., Wunch, D., Blavier, J.-F., Griffith, D. W. T., Heikkinen, P., Keeling, R. F., Kiel, M., Kivi, R., Roehl, C. M., Stephens, B. B., Baier, B. C., Chen, H., Choi, Y., Deutscher, N. M., DiGangi, J. P., Gross, J., Herkommer, B., Jeseck, P., Laemmel, T., Lan, X., McGee, E., McKain, K., Miller, J., Morino, I., Notholt, J., Ohyama, H., Pollard, D. F., Rettinger, M., Riris, H., Rousogenous, C., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Wofsy, S. C., Zhou, M., and Wennberg, P. O.: The Total Carbon Column Observing Network's GGG2020 data version, Earth System Science Data, 16, 2197–2260, https://doi.org/10.5194/essd-16-2197-2024, 2024.

Liu, C., Wang, W., Sun, Y., and Shan, C.: TCCON data from Hefei (PRC), Release GGG2020.R1, https://doi.org/10.14291/TCCON.GGG2020.HEFEI01.R1, 2023.

Liu, Z., Rogers, B. M., Keppel-Aleks, G., Helbig, M., Ballantyne, A. P., Kimball, J. S., Chatterjee, A., Foster, A., Kaushik, A., Virkkala, A.-M., Burrell, A. L., Schwalm, C., Sweeney, C., Schuur, E. A. G., Dean, J., Watts, J. D., Kim, J. E., Wang, J. A., Hu, L., Welp, L., Berner, L. T., Mauritz, M., Mack, M., Parazoo, N. C., Madani, N., Keeling, R., Commane, R., Goetz, S., Piao, S., Natali, S. M., Wang, W., Buermann, W., Walker, X., Lin, X., Wang, X., Jin, Y., Yu, K., and Zhang, Y.: Seasonal CO2 amplitude in northern high latitudes, Nature Reviews Earth & Environment, 5, 802–817, https://doi.org/10.1038/s43017-024-00600-7, 2024.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Morino, I., Ohyama, H., Hori, A., and Ikegami, H.: TCCON data from Rikubetsu (JP), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.RIKUBETSU01.R0, 2022a.

Morino, I., Ohyama, H., Hori, A., and Ikegami, H.: TCCON data from Tsukuba (JP), 125HR, Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.TSUKUBA02.R0, 2022b.

Morino, I., Velazco, V. A., Hori, A., Uchino, O., and Griffith, D. W.: TCCON data from Burgos, Ilocos Norte (PH), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.BURGOS01.R0, 2022c.

Notholt, J., Petri, C., Warneke, T., and Buschmann, M.: TCCON data from Bremen (DE), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.BREMEN01.R0, 2022.

Noël, S., Reuter, M., Buchwitz, M., Borchardt, J., Hilker, M., Schneising, O., Bovensmann, H., Burrows, J. P., Di Noia, A., Parker, R. J., Suto, H., Yoshida, Y., Buschmann, M., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Hase, F., Kivi, R., Liu, C., Morino, I., Notholt, J., Oh, Y.-S., Ohyama, H., Petri, C., Pollard, D. F., Rettinger, M., Roehl, C., Rousogenous, C., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Vrekoussis, M., and Warneke, T.: Retrieval of greenhouse gases from GOSAT and GOSAT-2 using the FOCAL algorithm, Atmospheric Measurement Techniques, 15, 3401–3437, https://doi.org/10.5194/amt-15-3401-2022, 2022.

O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO_2 retrieval algorithm – Part 1: Description and validation against synthetic observations, Atmospheric Measurement Techniques, 5, 99–121, https://doi.org/10.5194/amt-5-99-2012, 2012.

Parker, R., Boesch, H., Cogan, A., Fraser, A., Feng, L., Palmer, P. I., Messerschmidt, J., Deutscher, N., Griffith, D. W. T., Notholt, J., Wennberg, P. O., and Wunch, D.: Methane Observations from the Greenhouse Gases Observing SATellite: Comparison to Ground-Based TCCON Data and Model Calculations, Geophysical Research Letters, 38, https://doi.org/10.1029/2011GL047871, 2011.

Parker, R., Bharathan, L., Orr, D., and Cartwright, M. P.: EOCIS: University of Leicester GOSAT-2 Proxy XCH4, v1.0, https://doi.org/10.5285/E7E4499F7F32493DB96635026B95DED8, 2025.

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 CH_4 Observations, Earth System Science Data, 12, 3383–3412, https://doi.org/10.5194/essd-12-3383-2020, 2020.

Petri, C., Deutscher, N. M., Notholt, J., Warneke, T., and Buschmann, M.: TCCON data from Bialystok (PL), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.BIALYSTOK01.R0, 2024a.

Petri, C., Vrekoussis, M., Rousogenous, C., Warneke, T., Sciare, J., and Notholt, J.: TCCON data from Nicosia (CY), Release GGG2020.R1, https://doi.org/10.14291/TCCON.GGG2020.NICOSIA01.R1, 2024b.

Pollard, D. F., Robinson, J., and Shiona, H.: TCCON data from Lauder (NZ), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.LAUDER03.R0, 2022.

Reuter, M., Buchwitz, M., Schneising, O., Heymann, J., Bovensmann, H., and Burrows, J.: A method for improved SCIAMACHY CO2 retrieval in the presence of optically thin clouds, Atmospheric Measurement Techniques, 3, 209–232, 2010.

Reuter, M., Bovensmann, H., Buchwitz, M., Burrows, J., Connor, B., Deutscher, N. M., Griffith, D., Heymann, J., Keppel-Aleks, G., Messerschmidt, J., et al.: Retrieval of atmospheric CO2 with enhanced accuracy and precision from SCIAMACHY: Validation with FTS measurements and comparison with model results, Journal of Geophysical Research: Atmospheres, 116, 2011.

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Reuter, M., Bösch, H., Bovensmann, H., Bril, A., Buchwitz, M., Butz, A., Burrows, J. P., O'Dell, C. W., Guerlet, S., Hasekamp, O., Heymann, J., Kikuchi, N., Oshchepkov, S., Parker, R., Pfeifer, S., Schneising, O., Yokota, T., and Yoshida, Y.: A joint effort to deliver satellite retrieved atmospheric CO2 concentrations for surface flux inversions: the ensemble median algorithm EMMA, Atmospheric Chemistry and Physics, 13, 1771–1780, https://doi.org/10.5194/acp-13-1771-2013, 2013.

Reuter, M., Bovensmann, H., Buchwitz, M., Burrows, J., Heymann, J., and Schneising, O.: Algorithm Theoretical Basis Document Version 5 (ATBDv5) - The Bremen Optimal Estimation DOAS (BESD) algorithm for the retrieval of XCO2 for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), 2016.

Reuter, M., Buchwitz, M., Schneising, O., Noël, S., Bovensmann, H., and Burrows, J. P.: A Fast Atmospheric Trace Gas Retrieval for Hyperspectral Instruments Approximating Multiple Scattering—Part 2: Application to XCO2 Retrievals from OCO-2, Remote Sensing, 9, 1102, https://doi.org/10.3390/rs9111102, 2017a.

Reuter, M., Buchwitz, M., Schneising, O., Noël, S., Rozanov, V., Bovensmann, H., and Burrows, J. P.: A Fast Atmospheric Trace Gas Retrieval for Hyperspectral Instruments Approximating Multiple Scattering—Part 1: Radiative Transfer and a Potential OCO-2 XCO2 Retrieval Setup, Remote Sensing, 9, 1159, https://doi.org/10.3390/rs9111159, 2017b.

Reuter, M., Buchwitz, M., Schneising, O., Noël, S., Bovensmann, H., Burrows, J. P., Boesch, H., Di Noia, A., Anand, J., Parker, R. J., Somkuti, P., Wu, L., Hasekamp, O. P., Aben, I., Kuze, A., Suto, H., Shiomi, K., Yoshida, Y., Morino, I., Crisp, D., O'Dell, C. W., Notholt, J., Petri, C., Warneke, T., Velazco, V. A., Deutscher, N. M., Griffith, D. W. T., Kivi, R., Pollard, D. F., Hase, F., Sussmann, R., Té, Y. V., Strong, K., Roche, S., Sha, M. K., De Mazière, M., Feist, D. G., Iraci, L. T., Roehl, C. M., Retscher, C., and Schepers, D.: Ensemble-based satellite-derived carbon dioxide and methane column-averaged dry-air mole fraction data sets (2003–2018) for carbon and climate applications, Atmospheric Measurement Techniques, 13, 789–819, https://doi.org/10.5194/amt-13-789-2020, 2020.

Reuter, M., Hilker, M., Noël, S., Buchwitz, M., Schneising, O., Bovensmann, H., and Burrows, J. P.: Algorithm Theoretical Basis Document Version 3 (ATBDv3) – Retrieval of XCO2 from the OCO-2 Satellite using the Fast Atmospheric Trace Gas Retrieval (FOCAL) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), https://www.iup.uni-bremen.de/carbon_ghg/docs/GHG-CCIplus/CRDP7/ATBDv3_GHG-CCI_CO2_OC2_FOCA_v10.pdf, accessed: 19-01-2026, 2021.

Reuter, M., Hilker, M., and Buchwitz, M.: Product User Guide Version 5 (PUGv5) for the FOCAL XCO2 OCO-2 Data Product CO2_OC2_FOCA (v11), https://climate.esa.int/media/documents/PUG_CRDP9_v2_GHG-CCI_CO2_OC2_FOCA_v11.pdf, accessed: 19/03/2026, 2024.

Reuter, M., Fuentes Andrade, B., and Buchwitz, M.: C3S Greenhouse Gas (GHG: CO2 & CH4) v4.6: Algorithm Theoretical Basis Document (ATBD), https://confluence.ecmwf.int/x/AS-OIQ, accessed: 12/01/2026, 2025a.

Reuter, M., Fuentes Andrade, B., and Buchwitz, M.: C3S Greenhouse Gas (GHG: CO2 & CH4) v4.6: Product Quality Assessment Report (PQAR), https://confluence.ecmwf.int/x/LC-OIQ, accessed: 12/01/2026, 2025b.

Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, Series on atmospheric, oceanic and planetary physics, World Scientific, Singapore, repr edn., ISBN 978-981-02-2740-1, 2008.

Saunois, M., Martinez, A., Poulter, B., Zhang, Z., Raymond, P. A., Regnier, P., Canadell, J. G., Jackson, R. B., Patra, P. K., Bousquet, P., Ciais, P., Dlugokencky, E. J., Lan, X., Allen, G. H., Bastviken, D., Beerling, D. J., Belikov, D. A., Blake, D. R., Castaldi, S., Crippa, M., Deemer, B. R., Dennison, F., Etiope, G., Gedney, N., Höglund-Isaksson, L., Holgerson, M. A., Hopcroft, P. O., Hugelius, G., Ito, A., Jain, A. K., Janardanan, R., Johnson, M. S., Kleinen, T., Krummel, P. B., Lauerwald, R., Li, T., Liu, X., McDonald, K. C., Melton, J. R., Mühle, J., Müller, J., Murguia-Flores, F., Niwa, Y., Noce, S., Pan, S., Parker, R. J., Peng, C., Ramonet, M., Riley, W. J., Rocher-Ros, G., Rosentreter, J. A., Sasakawa, M., Segers, A., Smith, S. J., Stanley, E. H., Thanwerdas, J., Tian, H., Tsuruta, A., Tubiello, F. N., Weber,

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

T. S., van der Werf, G. R., Worthy, D. E. J., Xi, Y., Yoshida, Y., Zhang, W., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q.: Global Methane Budget 2000–2020, Earth System Science Data, 17, 1873–1958, https://doi.org/10.5194/essd-17-1873-2025, 2025.

Schneising, O., Buchwitz, M., Reuter, M., Heymann, J., Bovensmann, H., and Burrows, J. P.: Long-term analysis of carbon dioxide and methane column-averaged mole fractions retrieved from SCIAMACHY, Atmospheric Chemistry and Physics, 11, 2863–2880, https://doi.org/10.5194/acp-11-2863-2011, 2011.

Schneising, O., Bergamaschi, P., Bovensmann, H., Buchwitz, M., Burrows, J. P., Deutscher, N. M., Griffith, D. W. T., Heymann, J., Macatangay, R., Messerschmidt, J., Notholt, J., Rettinger, M., Reuter, M., Sussmann, R., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D.: Atmospheric greenhouse gases retrieved from SCIAMACHY: comparison to ground-based FTS measurements and model results, Atmospheric Chemistry and Physics, 12, 1527–1540, https://doi.org/10.5194/acp-12-1527-2012, 2012.

Schneising, O., Bovensmann, H., Buchwitz, M., Buschmann, M., Deutscher, N. M., Griffith, D. W. T., Hachmeister, J., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Ohyama, H., Petri, C., Reuter, M., Robinson, J., Roehl, C., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Vrekoussis, M., Wang, W., Warneke, T., Weidmann, D., Wunch, D., Zhou, M., and Bösch, H.: TROPOMI/WFMD v2.0: Improved Retrievals of XCH_4 and XCO with XGBoost-based Quality Filtering, Atmospheric Measurement Techniques, 19, 2407–2435, https://doi.org/10.5194/amt-19-2407-2026, 2026.

Segers, A.: Description of the CH4 Inversion Production Chain, https://atmosphere.copernicus.eu/sites/default/files/2022-10/CAMS255_2021SC1_D55.5.2.1-2021CH4_202206_production_chain_CH4_v1.pdf, accessed 2026-01-07, 2022.

Sherlock, V., Connor, B., Robinson, J., Shiona, H., Smale, D., and Pollard, D. F.: TCCON data from Lauder (NZ), 120HR, Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.LAUDER01.R0, 2022a.

Sherlock, V., Connor, B., Robinson, J., Shiona, H., Smale, D., and Pollard, D. F.: TCCON data from Lauder (NZ), 125HR, Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.LAUDER02.R0, 2022b.

Shiomi, K., Kawakami, S., Ohyama, H., Arai, K., Okumura, H., Ikegami, H., and Usami, M.: TCCON data from Saga (JP), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.SAGA01.R0, 2022.

Someya, Y., Yoshida, Y., Ohyama, H., Nomura, S., Kamei, A., Morino, I., Mukai, H., Matsunaga, T., Laughner, J. L., Velazco, V. A., Herkommer, B., Té, Y., Sha, M. K., Kivi, R., Zhou, M., Oh, Y. S., Deutscher, N. M., and Griffith, D. W. T.: Update on the GOSAT TANSO–FTS SWIR Level 2 Retrieval Algorithm, Atmospheric Measurement Techniques, 16, 1477–1501, https://doi.org/10.5194/amt-16-1477-2023, 2023.

Strong, K., Roche, S., Franklin, J. E., Mendonca, J., Lutsch, E., Weaver, D., Fogal, P. F., Drummond, J. R., Batchelor, R., Lindenmaier, R., and McGee, E.: TCCON data from Eureka (CA), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.EUREKA01.R0, 2022.

Sussmann, R., Rettinger, M., and Mostafavi Pak, N.: TCCON data from Garmisch (DE), Release GGG2020.R1, https://doi.org/10.14291/TCCON.GGG2020.GARMISCH01.R1, 2025.

Suto, H., Kataoka, F., Kikuchi, N., Knuteson, R. O., Butz, A., Haun, M., Buijs, H., Shiomi, K., Imai, H., and Kuze, A.: Thermal and near-infrared sensor for carbon observation Fourier transform spectrometer-2 (TANSO-FTS-2) on the Greenhouse gases Observing SATellite-2 (GOSAT-2) during its first year in orbit, Atmospheric Measurement Techniques, 14, 2013–2039, https://doi.org/10.5194/amt-14-2013-2021, 2021.

Taylor, T. E., O'Dell, C. W., Crisp, D., Kuze, A., Lindqvist, H., Wennberg, P. O., Chatterjee, A., Gunson, M., Eldering, A., Fisher, B., Kiel, M., Nelson, R. R., Merrelli, A., Osterman, G., Chevallier, F., Palmer, P. I., Feng, L., Deutscher, N. M., Dubey, M. K., Feist, D. G., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Liu, C., De Mazière, M., Morino, I., Notholt, J., Oh, Y.-S., Ohyama, H., Pollard,

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

D. F., Rettinger, M., Schneider, M., Roehl, C. M., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Vrekoussis, M., Warneke, T., and Wunch, D.: An 11-year record of XCO2 estimates derived from GOSAT measurements using the NASA ACOS version 9 retrieval algorithm, Earth System Science Data, 14, 325–360, https://doi.org/10.5194/essd-14-325-2022, 2022.

Taylor, T. E., O'Dell, C. W., Baker, D., Bruegge, C., Chang, A., Chapsky, L., Chatterjee, A., Cheng, C., Chevallier, F., Crisp, D., Dang, L., Drouin, B., Eldering, A., Feng, L., Fisher, B., Fu, D., Gunson, M., Haemmerle, V., Keller, G. R., Kiel, M., Kuai, L., Kurosu, T., Lambert, A., Laughner, J., Lee, R., Liu, J., Mandrake, L., Marchetti, Y., McGarragh, G., Merrelli, A., Nelson, R. R., Osterman, G., Oyafuso, F., Palmer, P. I., Payne, V. H., Rosenberg, R., Somkuti, P., Spiers, G., To, C., Weir, B., Wennberg, P. O., Yu, S., and Zong, J.: Evaluating the consistency between OCO-2 and OCO-3 XCO2 estimates derived from the NASA ACOS version 10 retrieval algorithm, Atmospheric Measurement Techniques, 16, 3173–3209, https://doi.org/10.5194/amt-16-3173-2023, 2023.

Té, Y., Jeseck, P., and Janssen, C.: TCCON data from Paris (FR), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.PARIS01.R0, 2022.

Veefkind, J., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H., de Haan, J., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sensing of Environment, 120, 70–83, https://doi.org/10.1016/j.rse.2011.09.027, 2012.

Waliser, D., Gleckler, P. J., Ferraro, R., Taylor, K. E., Ames, S., Biard, J., Bosilovich, M. G., Brown, O., Chepfer, H., Cinquini, L., Durack, P. J., Eyring, V., Mathieu, P.-P., Lee, T., Pinnock, S., Potter, G. L., Rixen, M., Saunders, R., Schulz, J., Thépaut, J.-N., and Tuma, M.: Observations for Model Intercomparison Project (Obs4MIPs): Status for CMIP6, Geoscientific Model Development, 13, 2945–2958, https://doi.org/10.5194/gmd-13-2945-2020, 2020.

Warneke, T., Petri, C., Notholt, J., and Buschmann, M.: TCCON data from OrlÃ©ans (FR), Release GGG2020.R1, https://doi.org/10.14291/TCCON.GGG2020.ORLEANS01.R1, 2024.

Weidmann, D., Brownsword, R., and Doniki, S.: TCCON data from Harwell, Oxfordshire (UK), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.HARWELL01.R0, 2023.

Wennberg, P. O., Roehl, C., Wunch, D., Blavier, J.-F., Toon, G. C., Allen, N. T., Treffers, R., and Laughner, J.: TCCON data from Caltech (US), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.PASADENA01.R0, 2022a.

Wennberg, P. O., Roehl, C. M., Blavier, J.-F., Wunch, D., and Allen, N. T.: TCCON data from Jet Propulsion Laboratory (US), 2011, Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.JPL02.R0, 2022b.

Wennberg, P. O., Roehl, C. M., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R., Keppel-Aleks, G., and Allen, N. T.: TCCON data from Park Falls (US), Release GGG2020.R1, https://doi.org/10.14291/TCCON.GGG2020.PARKFALLS01.R1, 2022c.

Wennberg, P. O., Wunch, D., Yavin, Y., Toon, G. C., Blavier, J.-F., Allen, N. T., Keppel-Aleks, G., and Roehl, C.: TCCON data from Jet Propulsion Laboratory (US), 2007, Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.JPL01.R0, 2022d.

Wennberg, P. O., Wunch, D., Roehl, C. M., Blavier, J.-F., Toon, G. C., and Allen, N. T.: TCCON data from Lamont (US), Release GGG2020.R1, https://doi.org/10.14291/TCCON.GGG2020.LAMONT01.R1, 2025.

Wu, L.: Algorithm Theoretical Basis Document (ATBD) - ANNEX B for products CO2_GOS_SRFP, CH4_GOS_SRFP (v2.3.8, 2009-mid2020), 2021.

Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jiménez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T.,

https://doi.org/10.5194/egusphere-2026-3216
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.



Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M. A.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmospheric Measurement Techniques, 3, 1351–1362, https://doi.org/10.5194/amt-3-1351-2010, 2010.

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The Total Carbon Column Observing Network, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 369, 2087–2112, https://doi.org/10.1098/rsta.2010.0240, 2011.

Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu, C., Feist, D. G., and Wennberg, P. O.: Documentation for the 2014 TCCON Data Release, https://data.caltech.edu/records/kv84h-9ta06, 2015.

Wunch, D., Wennberg, P. O., Osterman, G., Fisher, B., Naylor, B., Roehl, C. M., O'Dell, C., Mandrake, L., Viatte, C., Kiel, M., Griffith, D. W. T., Deutscher, N. M., Velazco, V. A., Notholt, J., Warneke, T., Petri, C., De Maziere, M., Sha, M. K., Sussmann, R., Rettinger, M., Pollard, D., Robinson, J., Morino, I., Uchino, O., Hase, F., Blumenstock, T., Feist, D. G., Arnold, S. G., Strong, K., Mendonca, J., Kivi, R., Heikkinen, P., Iraci, L., Podolske, J., Hillyard, P. W., Kawakami, S., Dubey, M. K., Parker, H. A., Sepulveda, E., García, O. E., Te, Y., Jeseck, P., Gunson, M. R., Crisp, D., and Eldering, A.: Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) X CO2 measurements with TCCON, Atmospheric Measurement Techniques, 10, 2209–2238, https://doi.org/10.5194/amt-10-2209-2017, 2017.

Wunch, D., Mendonca, J., Colebatch, O., Allen, N. T., Blavier, J.-F., Kunz, K., Roche, S., Hedelius, J., Neufeld, G., Springett, S., Worthy, D., Kessler, R., and Strong, K.: TCCON data from East Trout Lake, SK (CA), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.EASTTROUTLAKE01.R0, 2022.

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global Concentrations of CO2 and CH4 Retrieved from GOSAT: First Preliminary Results, Sola, 5, 160–163, https://doi.org/10.2151/sola.2009-041, 2009.

Yoshida, Y. and Oshio, H.: GOSAT-2 TANSO-FTS-2 SWIR L2 Retrieval Algorithm Theoretical Basis Document, https://prdct.gosat-2.nies.go.jp/documents/pdf/ATBD_FTS-2_L2_SWL2_en_03.pdf, accessed: 19-01-2026, 2025.

Yoshida, Y., Kikuchi, N., Morino, I., Uchino, O., Oshchepkov, S., Bril, A., Saeki, T., Schutgens, N., Toon, G. C., Wunch, D., Roehl, C. M., Wennberg, P. O., Griffith, D. W. T., Deutscher, N. M., Warneke, T., Notholt, J., Robinson, J., Sherlock, V., Connor, B., Rettinger, M., Sussmann, R., Ahonen, P., Heikkinen, P., Kyrö, E., Mendonca, J., Strong, K., Hase, F., Dohe, S., and Yokota, T.: Improvement of the retrieval algorithm for GOSAT SWIR XCO2 and XCH4 and their validation using TCCON data, Atmospheric Measurement Techniques, 6, 1533–1547, https://doi.org/10.5194/amt-6-1533-2013, 2013.

Yoshida, Y., Someya, Y., Ohyama, H., Morino, I., Matsunaga, T., Deutscher, N. M., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Notholt, J., Pollard, D. F., Té, Y., Velazco, V. A., and Wunch, D.: Quality Evaluation of the Column-Averaged Dry Air Mole Fractions of Carbon Dioxide and Methane Observed by GOSAT and GOSAT-2, Sola, 19, 173–184, https://doi.org/10.2151/sola.2023-023, 2023.

Zhou, M., Wang, P., Kumps, N., Hermans, C., and Nan, W.: TCCON data from Xianghe, China, Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.XIANGHE01.R0, 2022.