https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


# 5 Years of GOSAT-2 Retrievals with RemoTeC: $XCO_2$ and $XCH_4$ Data Products with Quality Filtering by Machine Learning

Andrew Gerald Barr[1], Jochen Landgraf[1], Mari Martinez-Velarte[1], Mihalis Vrekoussis[2,3], Ralf Sussmann[4], Isamu Morino[5], Kimberly Strong[6], Minqiang Zhou[7], Voltaire A. Velazco[8], Hirofumi Ohyama[9], Thorsten Warneke[3], Frank Hase[10], and Tobias Borsdorff[1]

[1]Earth science group, SRON Netherlands Institute for Space Research, Niels Bohrweg 4, 2333 CA Leiden, The Netherlands
[2]Climate and Atmosphere Research Center (CARE-C), The Cyprus Institute, Nicosia, Cyprus
[3]Institute of Environmental Physics, University of Bremen, Bremen, Germany
[4]Karlsruhe Institute of Technology (KIT), IMK-IFU, Garmisch-Partenkirchen, Germany
[5]Satellite Remote Sensing Section and Satellite Observation Center, Earth system Division, National Institute for Environmental Studies (NIES), Onogawa 16-2, Tsukuba, Ibaraki 305-8506, Japan
[6]Department of Physics, University of Toronto MP710A, 60 St. George Street, Toronto, ON, M5S 1A7, Canada
[7]Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China
[8]Deutscher Wetterdienst (DWD), Meteorological Observatory Hohenpeissenberg, 82383 Hohenpeissenberg, Germany
[9]Earth System Division, National Institute for Environmental Studies, Tsukuba, Japan
[10]Institute of Meteorology and Climate Research (IMK-ASF), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

**Correspondence:** Andrew Gerald Barr (a.g.barr@sron.nl)

**Abstract.** Accurately measuring greenhouse gas concentrations to identify regional sources and sinks is essential for effectively monitoring and mitigating their impact on the Earth's changing climate. In this article we present the scientific data products of $XCO_2$ and $XCH_4$, retrieved with RemoTeC, from the Greenhouse Gases Observing Satellite-2 (GOSAT-2), which span a time range of five years. GOSAT-2 has the capability to measure total columns of $CO_2$ and $CH_4$ to the necessary requirements set by the Global Climate Observing System (GCOS), who define said requirements as accuracy < 10 ppb and < 0.5 ppm for $XCH_4$ and $XCO_2$ respectively, and stability of < 3 ppb $yr^{-1}$ and < 0.5 ppm $yr^{-1}$ for $XCH_4$ and $XCO_2$ respectively.

Central to the quality of the $XCO_2$ and $XCH_4$ datasets is the post-retrieval quality flagging step. Previous versions of RemoTeC products have relied on threshold filtering, flagging data using boundary conditions from a list of retrieval parameters. We present a novel quality filtering approach utilising a machine learning technique known as Random Forest Classifier (RFC) models. This method is developed under the European Space Agency's (ESA) Climate Change Initiative+ (CCI+) program and applied to data from GOSAT-2. Data from the Total Carbon Column Observing Network (TCCON) are employed to train the RFC models, where retrievals are categorized as good or bad quality based on the bias between GOSAT-2 and TCCON measurements. TCCON is a global network of Fourier transform spectrometers that measure telluric absorption spectra at infrared wavelengths. It serves as the scientific community's standard for validating satellite-derived $XCO_2$ and $XCH_4$ data. Our results demonstrate that the machine learning-based quality filtering achieves a significant improvement, with data yield increasing by up to 85% and RMSE improving by up to 30%, compared to traditional threshold-based filtering. Furthermore, inter-comparison with the TROPOspheric Monitoring Instrument (TROPOMI) indicates that the quality filtering RFC models generalise well to the full dataset, as the expected behaviour is reproduced on a global scale.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


Low systematic biases are essential for extracting meaningful fluxes from satellite data products. Through TCCON validation we find that all data products are within the breakthrough bias requirements set, with RMSE for $XCH_4$ <15 ppb and $XCO_2$ <2 ppm. We derive station-to-station biases of 4.2 ppb and 0.5 ppm for $XCH_4$ and $XCO_2$ respectively, and linear drift of 0.6 ppb $yr^{-1}$ and 0.2 ppm $yr^{-1}$ for $XCH_4$ and $XCO_2$ respectively.

For $XCH_4$, GOSAT-2 and TROPOMI are highly correlated with standard deviations less than 18 ppb and globally averaged biases close to 0 ppb. The inter-satellite bias between GOSAT and GOSAT-2 is significant, with an average global bias of -15 ppb. This is comparable to that seen between GOSAT and TROPOMI, consistent with our findings that GOSAT-2 and TROPOMI are in close agreement.

## 1 Introduction

Anthropogenic emissions of greenhouse gases (GHGs) such as carbon dioxide ($CO_2$) and methane ($CH_4$) over the last century have led to the rapid rise of concentrations of GHGs in the atmosphere (Figure 1.4, IPCC AR6 2021, Tans and Keeling (2020), Cross-Chapter Box 5.2 IPCC AR6 2021). The effect of such changes in atmospheric composition has a clear correlation with the change of climate variables - such as global sea surface temperature anomoly or sea level - with $CO_2$ excess over preindustrial level directly proportional to global mean surface temperature anomaly, relative to 1850-1900 (Figure 1.6, IPCC AR6 2021). Indeed, the emergence of trend in climate variables above the natural year-to-year variability has been firmly established (Banks and Wood, 2002; Giorgi and Bi, 2009; Lyu et al., 2014; Hawkins and Sutton, 2012; IPCC AR5, 2014; Tebaldi and Friedlingstein, 2013), on a global scale as well as regional ones (Mahlstein et al., 2011; Hawkins et al., 2020; Rohde and Hausfather, 2020). The ramifications of a warming climate are serious with significant negative implications affecting the entire globe.

Satellite retrievals of concentrations of $CO_2$ and $CH_4$, or rather column-averaged dry air mole fractions, denoted $XCH_4$ and $XCO_2$, play an essential role in monitoring the changing climate, as these variables can be used alongside inverse modelling of surface fluxes to estimate uptake and emission of GHG surface fluxes (Bergamaschi et al., 2009; Chevallier et al., 2007, 2005; Meirink et al., 2006; Metz et al., 2023). In particular satellite measurements that are sensitive to near-surface variations in GHG concentrations are essential, and tight requirements are necessary to accurately calculate fluxes and so quantify emissions. The Global Carbon Observing System (GCOS) has classed measurements of $CO_2$ and $CH_4$ columns as Essential Climate Variables (ECVs), and defines requirements as being accurate enough to be able to determine sources and sinks on regional scales (GCOS, 2016). To this end, ESA's Climate Change Initiative (CCI) seeks to achieve this with the GHG-CCI+ project in which ECVs of $CO_2$ and $CH_4$ columns are delivered globally.

Particular emphasis is placed on systematic biases in satellite data, such as the change in bias over time, of which the requirements on $XCO_2$ and $XCH_4$ are less than 0.5 ppm $yr^{-1}$ and 3 ppb $yr^{-1}$ respectively (GCOS, 2016). Furthermore, the station-to-station bias of sites, or accuracy, from the The Total Carbon Column Observing Network (TCCON), defined as the standard deviation of all station biases, should be less than 0.5 ppm for $XCO_2$ and less than 10 ppb for $XCH_4$ (GCOS, 2016).

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


The Japan Aerospace Exploration Agency (JAXA) operated satellite GOSAT-2 (Greenhouse Gases Observing Satellite-2) has onboard the TANSO-FTS-2 instrument (Thermal And Near infrared Sensor for carbon Observation-Fourier Transform Spectrometer-2), which has sufficient sensitivity to measure regional sources and sinks, and provides back-scattered level-1B radiance spectra with 10 km circular ground pixels, covering the globe every 6 days in sun-synchronous orbit (Suto et al., 2021; Imasu et al., 2023). TANSO-FTS-2 has an intelligent pointing system, allowing better coverage than its predecessor GOSAT. Also onboard is the dedicated cloud imager TANSO-CAI-2 (Thermal And Near infrared Sensor for carbon Observation-Cloud and Aerosol Imager-2) (Kuze et al., 2009, 2016; Yoshida et al., 2012). GOSAT-2 operates in the near-infrared (NIR), short-wave infrared (SWIR) bands, and the thermal infrared.

GOSAT was the first dedicated GHG observing satellite, and has been used in a wide variety of scientific studies relevant to $CO_2$ and $CH_4$ since 2009, from level 2 (Butz et al., 2011; Schepers et al., 2012; Parker et al., 2020; Taylor et al., 2022) to higher level studies. Crucial for the carbon cycle, fluxes of $CO_2$ have been inferred from total columns on regional scale (Chevallier et al., 2009; Basu et al., 2013; Detmers et al., 2015) as well as global scales (Turner et al., 2015; Jiang et al., 2021; Kou et al., 2023). Also for $CH_4$, global flux estimates and emissions (Maasakkers et al., 2019; Zhang et al., 2021) have been derived from GOSAT measurements, and also compared to the TROPOspheric Monitoring Instrument (TROPOMI) (Liang et al., 2023) and airborne in-situ measurements (Tadić et al., 2012). Science with TANSO-FTS-2 has been limited, and more restricted to total column products (Noël et al., 2022; Yoshida et al., 2023) or instrument studies (Suto et al., 2021; Uno et al., 2021), however a couple of recent studies have looked at retrievals from the Cloud and Aerosol Imager-2 (CAI-2) (Gogoi et al., 2023; Shi et al., 2021).

TCCON provides the most robust measure of the accuracy of total columns of GHGs measured by satellites (Wunch et al., 2010, 2011, 2015), and has for long been widespread used as the conventional validation for $XCH_4$ and $XCO_2$ retrievals (e.g. Dils et al. (2014); Malina et al. (2022)). It is a global network of Fourier transform spectrometers that observe, among others, $XCO_2$ and $XCH_4$ with a root mean square error (RMSE) on mole fractions of 0.15 % and 0.2 % respectively (Toon et al., 2009). These measurements are scaled to aircraft measurements for calibration (Washenfelder et al., 2006; Deutscher et al., 2010; Messerschmidt et al., 2011; Geibel et al., 2012). TCCON measures direct sunlight and can therefore only be performed under clear-sky conditions hampering coverage of the time-series.

## 2 Data products and Input Data

In this article we present the novel level 2 GOSAT-2 scientific data products developed by SRON the Netherlands Institute for Space Research. $XCO_2$ and $XCH_4$ are retrieved deploying the RemoTeC retrieval algorithm, and are processed within the GHG-CCI+ project (Dils et al., 2014; Buchwitz et al., 2015). RemoTeC uses two different retrieval approaches which we discuss further in section 3. From the two configurations available to RemoTeC, three column-averaged dry air mole fractions products are produced. GOSAT-2 has been operational since February 2019. The data products discussed here cover the time range of the first observations until the end of 2023. Data products are available from the ESA Climate Office, under version

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


v2.0.3 in Climate Data Research Package 9 (CDRP9) [1]. More information about the three SRON GOSAT-2 data products can be found in the Algorithm Theoretical Basis Documents (ATBDs) and Product User Guides (PUGs)[2].

The SRON GOSAT-2 data products are generated from calibrated TANSO-FTS-2 L1B data from v210.210 for 2019 until June 2023, made available by the National Institute for Environmental Studies (NIES). For the second half of 2023 we used L1B from v220.220. Instrument line shape (ILS) information is taken from Suto et al. (2021).

A pre-processing step brings meteorological data, surface data and satellite data together before the retrieval is run. Meteorological input data are taken from the ECMWF ERA5 reanalysis product on 137 altitude layers and a $0.75° \times 0.75°$ latitude/longitude grid (Hersbach et al., 2020). Surface information was taken from the extended Shuttle Radar Telemetry Mission (SRTM) digital elevation map. The model used, DEM3, has global coverage at 90 meter spatial resolution[3], extending the original SRTM which is limited to latitudes of 56°S to 60°N. The solar reference spectra used for the retrieval is compiled from the full resolution spectrum of Kurucz (1994).

Absorption cross sections come from the HITRAN 2008 database for spectroscopic parameters (Rothman et al., 2009). Apriori column density profiles for $CO_2$ and $CH_4$ we take from TM5 (Huijnen et al., 2010) and TM4 (Meirink et al., 2006) model simulations respectively. For the $XCH_4$ Proxy product, $XCO_2$ data is used from the CAMS global inversion-optimised greenhouse gas concentrations of Chevallier (2023). These are surface air-sample instantaneous 3 hourly mean columns on $1.9° \times 3.75°$ grids.

## 3 Retrieval

RemoTeC is a retrieval algorithm developed for the retrieval of trace gas column-averaged dry air mole fractions from measured level 1b radiance spectra in the near-infrared (NIR) and shortwave-infrared (SWIR) bands. It has been used extensively for the retrieval of trace gases from GOSAT observations to produce the SRON $XCH_4$ and $XCO_2$ data products (Butz et al., 2009, 2010), as well as the operational products of TROPOMI (Hu et al., 2016, 2018; Lorente et al., 2021) and SCHIAMACHY (Frankenberg et al., 2005, 2011; Dils et al., 2006, 2014) . Below we outline the retrieval approach. The same approach is also used to generate the GOSAT-2 data products.

An $XCO_2$ Full Physics product is obtained using the scattering forward model, and an $XCH_4$ Full Physics product is extracted from the same retrieval. Furthermore another $XCH_4$ product (the Proxy product) is obtained with the non-scattering forward model. For the Full Physics approach, light scattering by cirrus and aerosol particles is accounted for in the forward model. For the Proxy retrieval, scattering is neglected and hence atmospheric scattering properties do not need to be calculated (Butz et al., 2009).

[1] https://catalogue.ceda.ac.uk/
[2] Available from https://climate.esa.int/en/projects/ghgs/key-documents/
[3] http://www.viewfinderpanoramas.org/dem3.html

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


### 3.1 Forward Model

Both the scattering and non-scattering forward models have the same general concept in common which we outline here. The atmospheric state vector, $\boldsymbol{x}$, is related to the measurement vector, $\boldsymbol{y}$, through a forward model, F, which in the following equation:

$$\boldsymbol{y} = F(\boldsymbol{x}, \boldsymbol{b}) + \epsilon_y + \epsilon_F \tag{1}$$

where $\epsilon_y$ and $\epsilon_F$ are the error contributions from the measurement noise and forward model respectively, and $\boldsymbol{b}$ is the ancillary vector containing parameters that are not retrieved. In order that the retrieval can be solved iteratively, the forward model must be linearised. For iteration step $n$ the linearised forward model is approximated by:

$$F(\boldsymbol{x}, \boldsymbol{b}) = F(\boldsymbol{x_n}, \boldsymbol{b}) + K(\boldsymbol{x} - \boldsymbol{x_n}) \tag{2}$$

where $\boldsymbol{x_n}$ is the state vector for the $n$-th iteration step and $K$ is the Jacobian matrix at position $\boldsymbol{x_n}$ defined by:

$$K = \frac{\partial F}{\partial \boldsymbol{x}} \tag{3}$$

The inversion method optimises the state vector $\boldsymbol{x}$ with respect to the measurement $\boldsymbol{y}$ after applying the forward model $F$ to $\boldsymbol{x}$. The inversion method is based on the Tikhonov regularization scheme (Philips, 1962; Tikhonov, 1963; Hasekamp and Landgraf, 2005). Regularisation is required because the inverse problem is ill-posed (the measurements $\boldsymbol{y}$ typically contains insufficient information to retrieve all state vector elements independently). The inverse algorithm finds $\boldsymbol{x}$ by minimising the cost function that is the sum of the least-squares cost function and a side constraint weighted by the regularisation parameter $\gamma$ according to

$$\hat{\boldsymbol{x}} = \underset{min}{\boldsymbol{x}} (||S_y^{1/2}(F(\boldsymbol{x}) - \boldsymbol{y})||^2 + \gamma||W(\boldsymbol{x} - \boldsymbol{x_a})||^2) \tag{4}$$

where $S_y$ is the diagonal measurement error covariance matrix, which contains the noise estimate, $\boldsymbol{x_a}$ is an a priori state vector, and $W$ is a diagonal weighting matrix.

### 3.2 Proxy Approach

The Proxy approach is based on a non-scattering retrieval, thus the runtime of processing is around a factor of 4 faster than the Full Physics retrieval. Furthermore, many of the errors in the retrieval, including those due to aerosol, cancel out (Butz et al., 2009; Schepers et al., 2012) following the equation:

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


$$XCH_4 = \frac{[CH_4]}{[CO_2]} XCO_{2,\text{model}} \tag{5}$$

which determines $XCH_4$ from the retrieved total columns $[CH_4]$ and $[CO_2]$. Here, the assumption is that the light path modification by scattering particles such as aerosols is the same for $CH_4$ and $CO_2$ (Schepers et al., 2012). $[CH_4]$ and $[CO_2]$ are total columns retrieved from SWIR-1 at 1.6 $\mu$m, and $XCO_{2,\text{model}}$ is the total column dry air mixing ratio of $CO_2$ from an atmospheric model, on the same grid as GOSAT-2 observations. The main source of uncertainty in this approach is therefore $XCO_{2,\text{model}}$, thus the accuracy of the $XCH_4$ Proxy product is limited by the accuracy of the $XCO_2$ model (Chevallier, 2023).

The state vector of the Proxy retrieval contains $CO_2$ and $CH_4$ sub-columns in 12 vertical layers, $H_2O$ total column, Lambertian surface albedo, first order spectral dependence of surface albedo, an intensity offset and first order spectral shifts of Earth and Sun radiancies.

### 3.3 Full Physics Approach

The Full Physics retrieval uses a three-window approach retrieving information from the NIR, SWIR-1 and SWIR-2 bands. The treatment of aerosol in the Full Physics approach leads to more accurate retrieved total columns of trace gases, however the radiative transfer calculations are computationally expensive. The state vector of the Full Physics retrieval is the same as for the Proxy retrieval with additional parameters related to aerosol properties.

Aerosols are characterised by three parameters which relate to the aerosol column and size distribution of particles. The height distribution is approximated as a Gaussian function of centre height, $z_{aer}$ and width $\omega_0$:

$$h(z_k) = N \cdot \exp\left[ - \frac{4ln(z_k - z_{aer})^2}{(2\omega_0)^2} \right] \tag{6}$$

Here, $N$ is the total amount of particles and $z_k$ is the layer height. The size distribution is parameterised by a power law function following:

$$n(r) = \begin{cases} A & \text{for } r \leq r_1 \\ A(r/r_1)^{-\alpha} & \text{for} r_1 < r \leq r_2 \\ 0 & \text{for } r > r_2 \end{cases} \tag{7}$$

where $r_1$ = 0.1 $\mu$m, $r_2$ = 10 $\mu$m and the constant $A$ is determined from normalisation of the size distribution. $N$, $\alpha$ and $z_{aer}$ are included in the Full Physics state vector.

### 3.4 Bias Correction

A bias correction is applied post-retrieval to $XCO_2$ and $XCH_4$ using TCCON as a truth. The bias correction of RemoTeC is a simple empirical relation between $XCH_4$ and the retrieved albedo at 1600 nm, defined by:

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


$$X_{corr} = X_{ret}(a + b\alpha) \quad (8)$$

where $X_{ret}$ and $X_{corr}$ are the bias corrected and retrieved concentrations respectively, $\alpha$ is the retrieved albedo at 1600 nm and $a$ and $b$ are determined such that the difference with TCCON is minimised.

## 4 Quality Filtering

A key step in the retrieval process is the post-processing quality flagging. Data from GOSAT are flagged using a selection of retrieval parameters, such as signal-to-noise ratio or chi-squared, and any data that do not lie within a specified range of these parameters are flagged as bad quality. This method offers a binary quality flag and is described further in section 4.1.

Given the rapidly growing capabilities of machine learning techniques, algorithms such as random forest classifiers (RFCs) provide a much more promising way of filtering satellite global data products. We have applied such a flagging technique to the GOSAT-2 data (see section 4.2). The quality flag of GOSAT-2 takes the form of a quality assurance (QA) value that ranges between 0 and 1, with 0 corresponding to the best quality data. Therefore, users should quality filter their data by taking QA values less than, or equal to, the desired value.

### 4.1 Threshold Criteria Approach

Extensive investigations have been conducted to identify effective retrieval parameters, or combinations of parameters, that are correlated with the quality of $XCH_4$ or $XCO_2$ and that can be used to flag bad data, while at the same time maximising the amount of good quality retrievals (Butz et al., 2010; Schepers et al., 2012). Such a set of criteria have been established also for GOSAT-2 and are listed in Table 1.

Criteria 6 to 8 are excluded for the Proxy product, since these are not in the state vector of the retrieval. The blended albedo in criterion 9 is defined as $A_{\mathrm{bld}} = 2.4 \cdot A(0.76\mu\mathrm{m}) - 1.13 \cdot A(2.0\mu\mathrm{m})$ with the retrieved albedos $A(0.76\mu\mathrm{m})$ and$A(2.0\mu\mathrm{m})$ at the indicated wavelengths (Wunch et al., 2011). Guerlet et al. (2013) investigated the use of the cirrus radiance for data filtering, which is defined as the mean radiance in the spectral range 5154.8–5157.8 $\mathrm{cm}^{\breve{}1}$ ($1.9388 - 1.9399\mu$m). The use of the column ratios for data filtering was first proposed by Taylor et al. (2016) based on the difference in the non-scattering retrieved column from a weak and strong absorption band. For this purpose, in criteria 11 and 12, we use the $CO_2$ and $H_2O$ ratios inferred from the 1.6 $\mu$m and 2.0 $\mu$m spectral range. Finally, $O_2$ ratio is the retrieved $O_2$ column divided by the prior derived from the ECMWF surface pressure estimate.

### 4.2 Machine Learning Approach

An alternative approach to quality flagging with threshold criteria, as applied on GOSAT, is to use machine learning in the form of a random forest classifier (RFC). To this purpose, we use the RandomForestClassifier tool within Python's SciKit Learn package (Pedregosa et al., 2011).

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


**Table 1.** List of threshold conditions to quality filter GOSAT-2 data.

| Filter | Description | criteria |
|---|---|---|
| 1 | $\chi^2$ of spectral fit | $\chi^2$ < 12.0 |
| 2 | number of iterations $N_{\mathrm{iter}}$ | $N_{\mathrm{iter}} < 31$ |
| 3 | signal-to-noise ratio (SNR) at band continuum | SNR > 50 |
| 4 | variance $\sigma_{\mathrm{surf}}$ of surface elevation | $\sigma_{\mathrm{surf}}$ < 100 m |
| 5 | solar zenith angle (SZA) | SZA< 75° |
| 6 | aerosol optical thickness ($\tau_{\mathrm{aer}}$) in NIR window | $\tau_{\mathrm{aer}} < 1.0$ |
| 7 | aerosol size parameter $\alpha$ | 3 < $\alpha$ < 6 |
| 8 | aerosol layer height $z_{\mathrm{aer}}$ | 0 < $z_{\mathrm{aer}}$ < 10,000 m |
| 9 | blended albedo $A_{\mathrm{bld}}$ | 0 < $A_{\mathrm{bld}}$ < 1.4 |
| 10 | cirrus radiance signal $I_{\mathrm{cir}}$ | $0 < I_{\mathrm{cir}} < 2.0 \cdot 10^{-9}$ [W cm/(m$^2$ s str] |
| 11 | $CO_2$ column ratio $r_{\mathrm{CO_2}}$ | 0.99 < $R_{\mathrm{CO_2}}$ < 1.018 |
| 12 | $H_2O$ column ratio $r_{\mathrm{H_2O}}$ | 0.95 < $R_{\mathrm{H_2O}}$ < 1.08 |
| 13 | $O_2$ column ratio $r_{\mathrm{O_2}}$ | 0.96 < $R_{\mathrm{O_2}}$ < 1.04 |

A random forest model utilises an ensemble of $N$ decision trees, which take a random subset of the available features and each make a decision on the target classification (Breiman, 2001). The final result of the model is taken as the majority chosen class. This is ultimately applied to each ground pixel of GOSAT-2 data using a set of features consisting of RemoTeC retrieval parameters, to predict the quality of the retrieval. We use separately trained models for each of the three data products, which will be described in more detail in section 4.2.1.

#### 4.2.1 RFC training using TCCON data

For the quality classification of our data product, we use a trained RFC. The supervised form of learning requires a labeled training dataset. To this end, we need knowledge of a "ground truth" and the best estimate of the true value of $XCH_4$ and $XCO_2$ comes from TCCON. Therefore, in order to determine the true label for the quality flag, we use GOSAT-2 level 2 data from measurements that are colocated in space and time with TCCON sites, and classify the training sample via the bias:

$$|\Delta X| < X_T: \quad \text{label } L_{\mathrm{X_T}} = 0 \quad (\text{good}) \tag{9}$$

$$|\Delta X| > X_T: \quad \text{label } L_{\mathrm{X_T}} = 1 \quad (\text{bad}) \tag{10}$$

with the biases $\Delta XCH_4$ = $XCH_{4,\text{GOSAT-2}}$ - $XCH_{4,\text{TCCON}}$, and $\Delta XCO_2$ = $XCO_{2,\text{GOSAT-2}}$ - $XCO_{2,\text{TCCON}}$. $X_T$ we name the training threshold and takes the form of e.g. $\pm$ 18 ppb for $\Delta XCH_4$. A label $L$ of 0 corresponds to a good-quality retrieval, and a label of 1 means a bad-quality retrieval.

A consequence of training the random forest model on GOSAT-2 colocations with TCCON is that retrievals with surface albedo $\gtrapprox$ 0.4 were underrepresented in the training sample, due to the lack of TCCON stations in high albedo areas. This

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


would lead to albedo-related biases when using such models to filter the global dataset. To avoid this, we defined a subsample of retrievals with albedo > 0.4 to include in the training set using the threshold filtering criteria in Table 1.

In this study, we limit the quality filtering of GOSAT-2 data using the machine learning approach to retrievals over land only, due to the lack of available training data over ocean. We note that this limiting factor would not apply to satellite data from push-broom spectrometers that have better spatial coverage, such as TROPOMI (Hu et al., 2016, 2018). Instead, we filter retrievals performed in glint mode over the ocean also following Table 1.

TCCON $XCH_4$ and $XCO_2$ data are also used to validate the final product (see section 5.1). Due to the supervised learning approach, utilising TCCON data in training the filtering models means that these data can no longer be quality filtered without receiving what was defined during training. Since the training data also comprise the validation data, this would lead to artificially choosing validation results, therefore compromising any independent validation with TCCON using the assigned quality flags. To avoid this, we train different filtering models, one year at a time, where the data from the year to be predicted are excluded from training. This results in one filtering model per year of data. Here we assume that the relationship between retrieval quality and features is temporally independent.

#### 4.2.2 RFC Prediction Performance

To evaluate our classification for the three products, we consider the performance of the RFCs by comparing its predicted labels to the true labels given by the elements of the confusion matrix (Liang, 2022): the False-Positive (FP), the False-Negative (FN), the True-Positive (TP), and True-Negative (TN). Here the terms 'true' and 'false' refer to a correct or wrong prediction, 'positive' and 'negative' to the bad and good label of the predicted data. From these, we evaluate the classification using the following metrics:

1. The true-positive-rate TPR is defined as

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \,. \tag{11}$$

   and measures the number of correctly identified positive instances out of all true positive instances.

2. The False-Positive rate (FPR) is the corresponding rate of False Positive with respect to all true negative instances,

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} \,. \tag{12}$$

A binary classification model predicts the probability of an instance belonging to one of the two classes depending on the classification threshold, which we name $p_t$. Varying $p_t$, leads to the Receiver-Operating Characteristic curve (ROC) (Bradley, 1997), which is a parametric curve of $\mathrm{FPR}(p_t)$ versus $\mathrm{TPR}(p_t)$. For a large threshold ($p_t \to 1$), TPR goes to one, but so does the FPR. In the other extreme for $p_t \to 0$, both TPR and FPR go to zero. Therefore, the more the ROC curve goes through the top-left quadrant of the diagram, the better the classifier. This is characterized by the area under the ROC curve (AUC). We assume a value of 0.5 for $p_t$ for all classification models.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


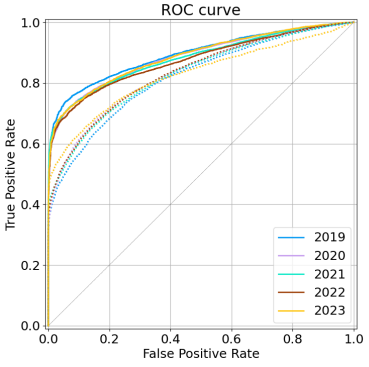


**Figure 1.** ROC curves for the different filtering models of the Full Physics $XCH_4$ product and Proxy $XCH_4$ product, in solid and dotted lines respectively. The solid grey line indicates the performance of a randomly guessed prediction, with a 50 % chance of being correct.

**Table 2.** Summary of classification metrics averaged over all years for $p_t$=0.5.

| Product | TPR | FPR | AUC |
|---|---|---|---|
| $XCO_2$ Full Physics | 0.90 | 0.42 | 0.89 |
| $XCH_4$ Full Physics | 0.89 | 0.44 | 0.89 |
| $XCH_4$ Proxy | 0.64 | 0.14 | 0.83 |

Figure 1 compares the ROC curve for one of the $XCH_4$ Full Physics classification models to one of the $XCH_4$ Proxy ones. There is no clear differentiation between the ROC curves of each product implying that the models for each year perform comparably to each other. Average metrics over all models per product are given in Table 2. From this we see that the performance of the RFC models for the two Full Physics products are similar - which is intuitive given that these come from the same retrieval - whereas the diagnostics for the Proxy product are slightly worse.

Such an effect can be understood by the nature of the Proxy approach and as a consequence of equation 5, where most of the systematic error is divided out by dividing the two columns of $CH_4$ and $CO_2$. Consequently, the distinction between high quality and low quality retrievals is much more obvious in the Full Physics case. Quantitatively, the ratio of good to bad retrievals in the training data is about 0.3 for the Full Physics products, whereas for the Proxy it is 1.6. It is therefore easier to accurately label the training sample in the Full Physics RFC models, leading to better performance metrics.

#### 4.2.3 The QA value

In the random forest models, the strictness of the training threshold, $X_T$, defined when labelling the training dataset (equations 9 & 10) has a directly proportional effect to the number of retrievals ultimately classed as good quality, as well as the scatter of the total column mixing ratio with respect to TCCON. Figure 2 shows the number of good retrievals as a function of the RMSE

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


derived for different training thresholds $X_T$ and the depicted positive trend is intuitively expected. This allows us to define a non-binary quality flag that is grounded in TCCON validation.

Starting with a set of $n$ threshold values $X_{T_1}, \cdots X_{T_n}$ we can assign the label vector $L = (L_1, \cdots L_n)$ with $L_i = L_{X_{T_i}}$ as defined in Eq. 9. We define the QA value of a data point by the mean value of the components of the corresponding label vector $L$,

$$QA = \langle L \rangle \tag{13}$$

QA can have $n+1$ discrete values in the range $[0,1]$ depending on the number $n$ of used threshold values $X_T$. For GOSAT-2, we use $n = 5$ with $QA \in [0, 1/5, 2/5, \cdots, 1]$.

Figure 2 also shows that in comparison to filtering GOSAT-2 data using thresholding defined in Table 1, for the Full Physics products, the new filtering can achieve an increase in data yield of $\sim$ 48 % and 85 % for $XCH_4$ and $XCO_2$ respectively, for equivalent RMSE. Alternatively, an improvement in RMSE of 2.2 ppb and 0.7 ppm for $XCH_4$ and $XCO_2$ respectively can be achieved for equivalent data yield. For the Proxy product this can be 1.6 ppb for the same amount of data, or conversely, an increase of 29 % for the same RMSE. Thus user can therefore choose the option of more data, which may be advantageous to plume detection where better coverage is desirable, or better quality data, where as small as possible systematic biases are required by atmospheric modellers. Furthermore, with Figure 2, the user may choose the QA value which corresponds to their acceptable RMSE with TCCON.

## 5 Validation and Satellite Inter-comparison

### 5.1 TCCON Validation

TCCON is central to the work presented here as it provides both the ground truth in labeling training data, as well as one of the main validation sources. In this article, all references to TCCON are for the GGG2020 TCCON release (Laughner et al., 2023). The TCCON stations used in the analysis are summarised in the Appendix in Table B1.

In this section, we present the validation of our GOSAT-2 data products with respect to TCCON. TCCON sites are considered only if there are more than 50 spatio-temporal colocations with GOSAT-2 over the whole time-series, defined as overlying within a radius of 300 km and time range of $\pm$ 2.5 hrs. We evaluate the data products for the QA value of 0 (strictest filtering with RFC models; see section 4.2.3).

Figure 3 shows the correlation between colocated GOSAT-2 and TCCON data for $XCO_2$ and both $XCH_4$ products. These are single soundings of GOSAT-2 over land compared to an average of the TCCON measurements that coincide spatially and temporally. For $XCH_4$ we derive a RMSE of 15.2 ppb and 15.7 ppb for the Full Physics and Proxy products respectively, and Pearson's correlation coefficient of 0.89 and 0.88 for the Full Physics and Proxy products respectively. For $XCO_2$ these are 2.1 ppm and 0.88 respectively.

Time-series of GOSAT-2 colocations with TCCON for each product are shown in the Appendix section B. Following Noël et al. (2022), we further parameterise the bias over time as:

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


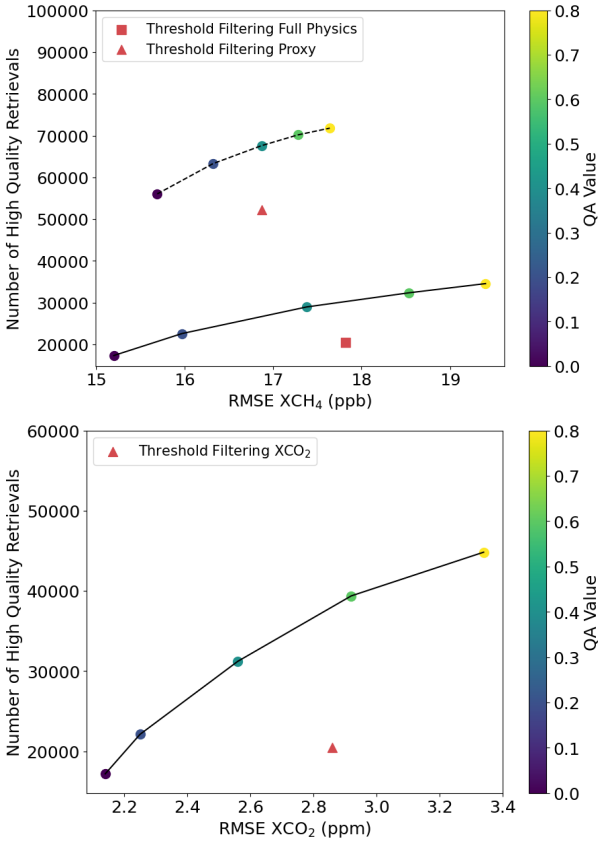


**Figure 2.** Number of retrievals flagged as good for five different thresholds as a function of the RMSE derived by the TCCON validation. The mean QA value per data ensemble is given by the color code. Results for $XCH_4$ are shown on the top panel, where the RFC filtering models for the Proxy and Full Physics product are represented by the dashed and solid lines respectively. On the bottom panel are the results for the $XCO_2$ filtering models. The red squares and triangles mark the parameter space for the statistics of filtering the data product according to Table 1.

$$\Delta X = a_0 + a_1 t + a_2 \sin(2\pi t + a_3) + \epsilon \tag{14}$$

where equation 14 is fit to the time-series of the bias of with each station individually. $a_0$ is a constant bias term, $a_1$ represents a linear term, $a_2$ measures the amplitude of the seasonal variation of the bias, $a_3$ measures the temporal shift of the seasonal term, and $\epsilon$ is an error term.

The parameters in Table 3 are extracted from fits of equation 14 to the time-series of the bias for each TCCON station. We illustrate an overview of the per station statistics in terms of site and seasonal bias, as well as linear drift in the bias, in Figures 4 to 6. $\Delta_{site}$ is the site bias and defined as the mean of $\Delta X$ from equation 14 and $\Delta_{seas}$ is the seasonal bias and defined as

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


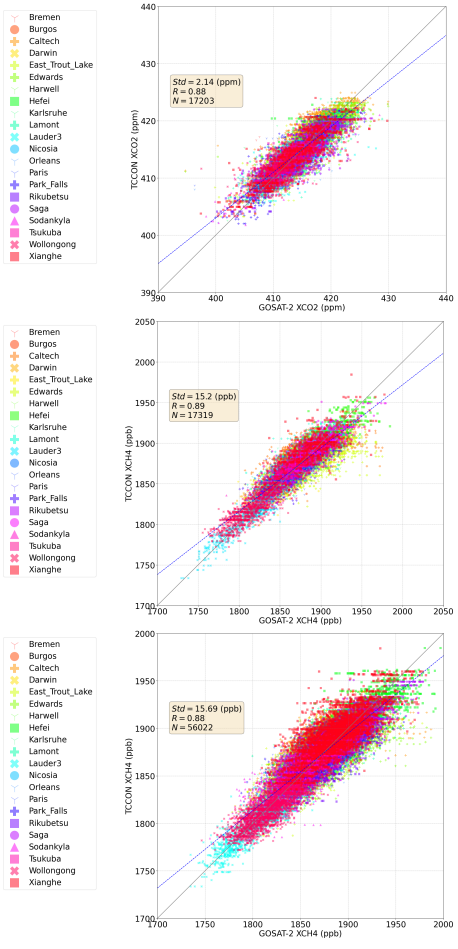


**Figure 3.** GOSAT-2 $XCO_2$ (top), $XCH_4$ Full Physics (middle) and $XCH_4$ Proxy (bottom) plotted against TCCON. Data are compared only if they are fully colocated in space and time. The standard deviation of the population, Pearson's correlation coefficient and number of retrievals are given in the inset. The legend plots the different TCCON stations where markers are as follows. Stations that are along the coast and also sensitive to glint mode (ocean) measurements are indicated as circles. Those that have high latitudes in the northern and southern hemispheres are upward triangles and crosses, respectively. Stations in Asia, North America and Europe are indicated by squares, pluses and downward triangles respectively.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


the standard deviation of the seasonal (sine) term in equation 14. Finally, $\Delta_{dri}$ is the linear drift and is calculated as $a_1$ from equation 14.

For the Full Physics products, we derive average values of the site bias of -0.1 ppb and -0.2 ppm for $XCH_4$ and $XCO_2$ respectively, after bias correction. We exclude the station-averaged $\Delta_{site}$ from Table 3 as it is by definition close to zero due to the bias correction. The seasonal bias term is higher for both products with 4.0 ppb and 0.6 ppm for $XCH_4$ and $XCO_2$ respectively. For the Proxy product, the average site bias and seasonal bias are 0.2 ppb and 3.1 ppb respectively.

From Table 3 we also report a linear drift of 0.6 ppb $yr^{-1}$ and 0.2 ppm $yr^{-1}$ for $XCH_4$ and $XCO_2$ respectively, for the Full Physics products. For the Proxy product the average linear drift is 1.2 ppb $yr^{-1}$. Another important metric of the systematic error is the station-to-station bias. This is defined as the standard deviation of the individual site biases, in contrast to the RMSE which is the standard deviation of all the differences together. We report station-to-station bias of 0.5 ppm, 4.2 ppb and 3.7 ppb for $XCO_2$, $XCH_4$ Full Physics and $XCH_4$ Proxy, respectively. The site-to-site biases and linear drift terms are low, and below the breakthrough systematic error threshold requirements (GCOS, 2016), which is an essential characteristic of the data product for determining regional scale sources and sinks through flux inversion modelling.

Comparing the TCCON validation for GOSAT-2 to those of GOSAT in Figures 4 to 6, we find that generally the RMSE is lower for GOSAT-2 than GOSAT across all stations while the number of retrievals is higher for GOSAT-2. We observe that the site bias is smaller for GOSAT-2 with GOSAT showing some significant biases with respect to TCCON, whereas the linear drift is more variable between the two satellites.

We note the difference between the average RMSE and that calculated from the sample of GOSAT-2/TCCON differences as a whole, which is particularly clear for $XCH_4$. The RMSE for $XCH_4$ Full Physics taking all data as one sample is 15.2 ppb, however the average of the individual station RMSE is 13.1 ppb. From Table 3, only three stations have a RMSE higher than 15 ppb: Caltech, Edwards and Xianghe. The disproportionately high number of collocations from Caltech and Edwards (together constituting 40 % of the data points) skew significantly the statistics towards these stations. Due to the difficult nature of accurately measuring GHG concentrations at all three of these stations, we consider the values for $XCH_4$ quoted in Figure 3 an upper limit. For the Proxy product the difference is less although still notable, with 15.7 ppb compared to 14.7 ppb when taking the average RMSE over all stations.

Despite the lower performance of the Proxy product with respect to the Full Physics ones observed in section 4.2.2, the level 2 quality of the Proxy $XCH_4$ product presented in Table 3 is effectively as good as the Full Physics $XCH_4$ product, with the advantage of better data coverage. This can be understood by in the ratio of FP/FN, for which in the case of Full Physics is 1:1, is 2:1 for the Proxy. The higher number of FPs lead to a poorer ROC curve, however in terms of the problem of quality filtering, FNs are more detrimental to the level 2 quality, since they correspond to ground truth bad data flagged as good.

In addition to the data we present in this article, two other $XCO_2$ and $XCH_4$ data products are also available from GOSAT-2. The first are the operational GOSAT-2 data products from NIES (Yoshida et al., 2023). The authors report RMSE with respect to TCCON of 1.8 ppm and 8.9 ppb for $XCO_2$ and $XCH_4$ respectively, across a time range of March 2019 to Dec 2020. Also, they derive station-to-station bias of 0.71 ppm and 2 ppb for $XCO_2$ and $XCH_4$ respectively. We note the short time-serie these values are derived from.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


**Table 3.** Summary of the main statistics of GOSAT-2 product validation with TCCON. RMSE is the root mean square error, $\Delta_{dri}$ is the linear drift and $\Delta_{seas}$ is the seasonal bias, averaged over all stations. $\sigma_{site}$ is the station-to-station bias.

| $XCH_4$ Full Physics | | | | $XCO_2$ Full Physics | | | | $XCH_4$ Proxy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| RMSE (ppb) | $\Delta_{seas}$ (ppb) | $\Delta_{dri}$ (ppb $yr^{-1}$) | $\sigma_{site}$ (ppb) | RMSE (ppm) | $\Delta_{seas}$ (ppm) | $\Delta_{dri}$ (ppm $yr^{-1}$) | $\sigma_{site}$ (ppm) | RMSE (ppb) | $\Delta_{seas}$ (ppb) | $\Delta_{dri}$ (ppb $yr^{-1}$) | $\sigma_{site}$ (ppb) |
| 13.1 | 4.0 | 0.6 | 4.2 | 2.0 | 0.6 | 0.2 | 0.5 | 14.7 | 3.1 | 1.2 | 3.7 |

Secondly are $XCO_2$ and $XCH_4$ products from IUP-Bremen (Noël et al., 2021, 2022). As for Yoshida et al. (2023), the time range of the dataset presented is until the end of 2020. The authors find, with respect to TCCON, a RMSE and station-to-station bias of 1.86 ppm and 1.14 ppm respectively, for $XCO_2$ (Noël et al., 2021). For $XCH_4$, station-to-station biases of 4.7 ppb and 6.2 ppb for the and Full Physics and Proxy products respectively, and RMSE of around 12 ppb for both, are reported (Noël et al., 2022). The authors note that, due to the short time-series, these results are drawn from only seven TCCON stations, some of which span only a few months.

### 5.2 GOSAT Inter-comparison

The similarity in the setup of GOSAT and GOSAT-2, along with the wide use of GOSAT in the scientific literature, make them ideal candidates for satellite inter-comparison. We compare our GOSAT-2 Full Physics products with the corresponding GOSAT products from RemoTeC, version 2.3.8, over time frame of 2019 to 2023. For the Proxy product comparison, we compare our GOSAT-2 $XCH_4$ Proxy product to that of GOSAT version 2.3.9. We present comparisons only for the GOSAT-2 QA value of 0.

From the global maps of the $XCH_4$ Full Physics product in Figure 7, the superior coverage of GOSAT-2 is striking; a consequence of the intelligent pointing system to avoid cloudy scenes. We further analyse how the GHG concentrations compare, illustrated as kernel density estimation (KDE) plots, analysing data over land only. The scatter of satellite differences is 14.5 ppb, similar to the RMSE of the bias with TCCON (13.1 ppb; see section 5.1). We find a large average global bias of -15.2 ppb, which we discuss further in section 5.3.

For the Proxy $XCH_4$ product, the comparison between GOSAT and GOSAT-2 is better compared to the Full Physics product. The average global bias is only -5.3 ppb, and the standard deviation and correlation coefficients are 13.5 ppb and 0.9 respectively.

For $XCO_2$, the correlation between GOSAT and GOSAT-2 is weaker, with a coefficient of 0.64 compared to 0.88 for $XCH_4$. This difference is expected, as $CO_2$'s longer atmospheric lifetime leads to greater large-scale diffusion, reducing correlation strength. The scatter of the differences is 2.9 ppm, slightly higher than the GOSAT-2 RMSE with respect to TCCON of 2.0 ppm, and we find a bias of 0.9 ppm.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


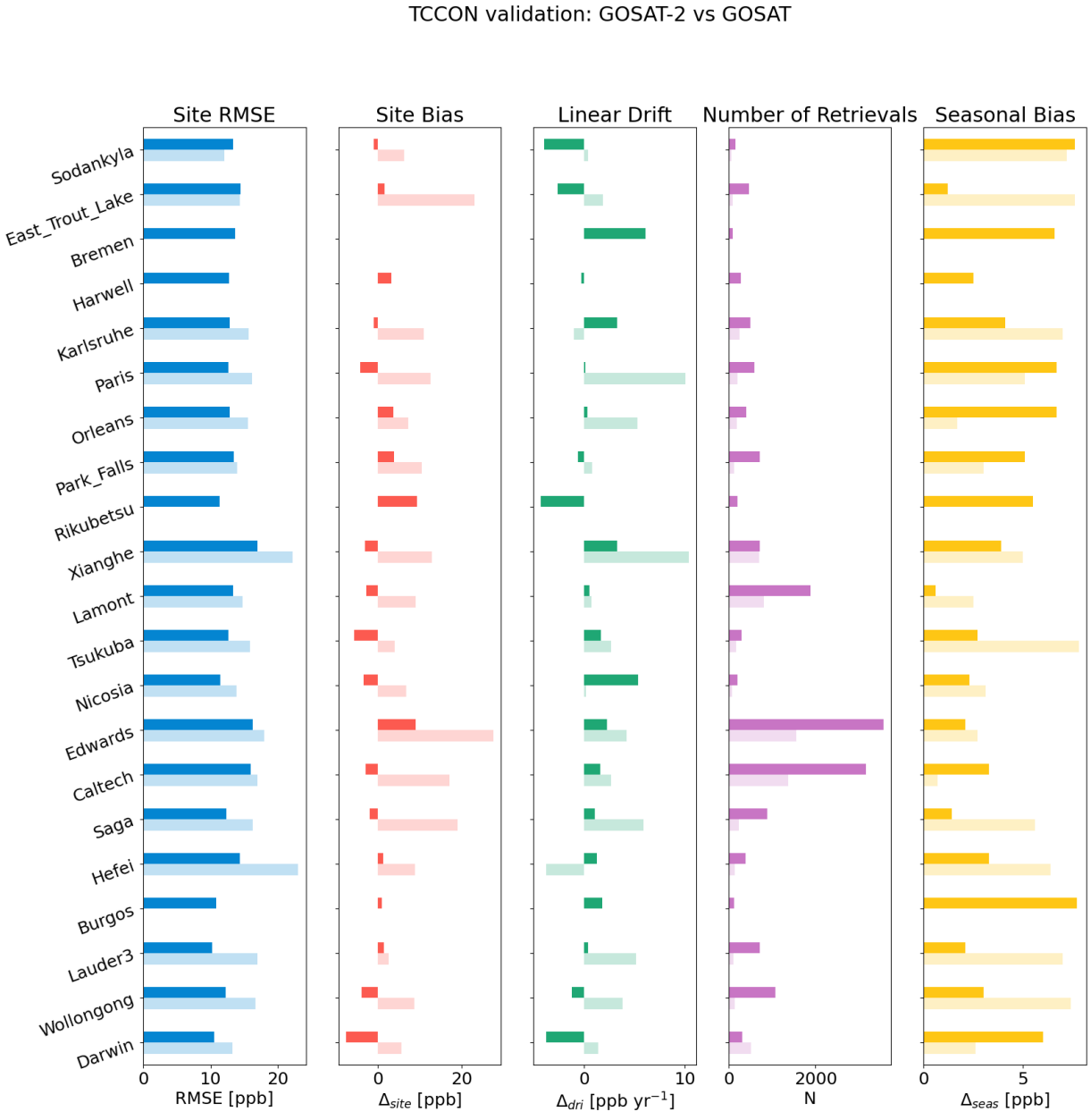


**Figure 4.** Overview of the bias parametrisation for the Full Physics $XCH_4$ product, per station. Shown in blue is the RMSE, red the site bias $\Delta_{site}$, green the linear drift $\Delta_{dri}$, yellow the seasonal bias $\Delta_{seas}$ and in purple the number of retrievals. Values for GOSAT-2 are shown in bold bars, and those of GOSAT are in light bars. Stations are listed in order of decreasing latitude. Missing bars correspond to less than 50 colocations for that station, therefore we do not calculate the values there. We note that the site bias for GOSAT-2 at Bremen is close to zero.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


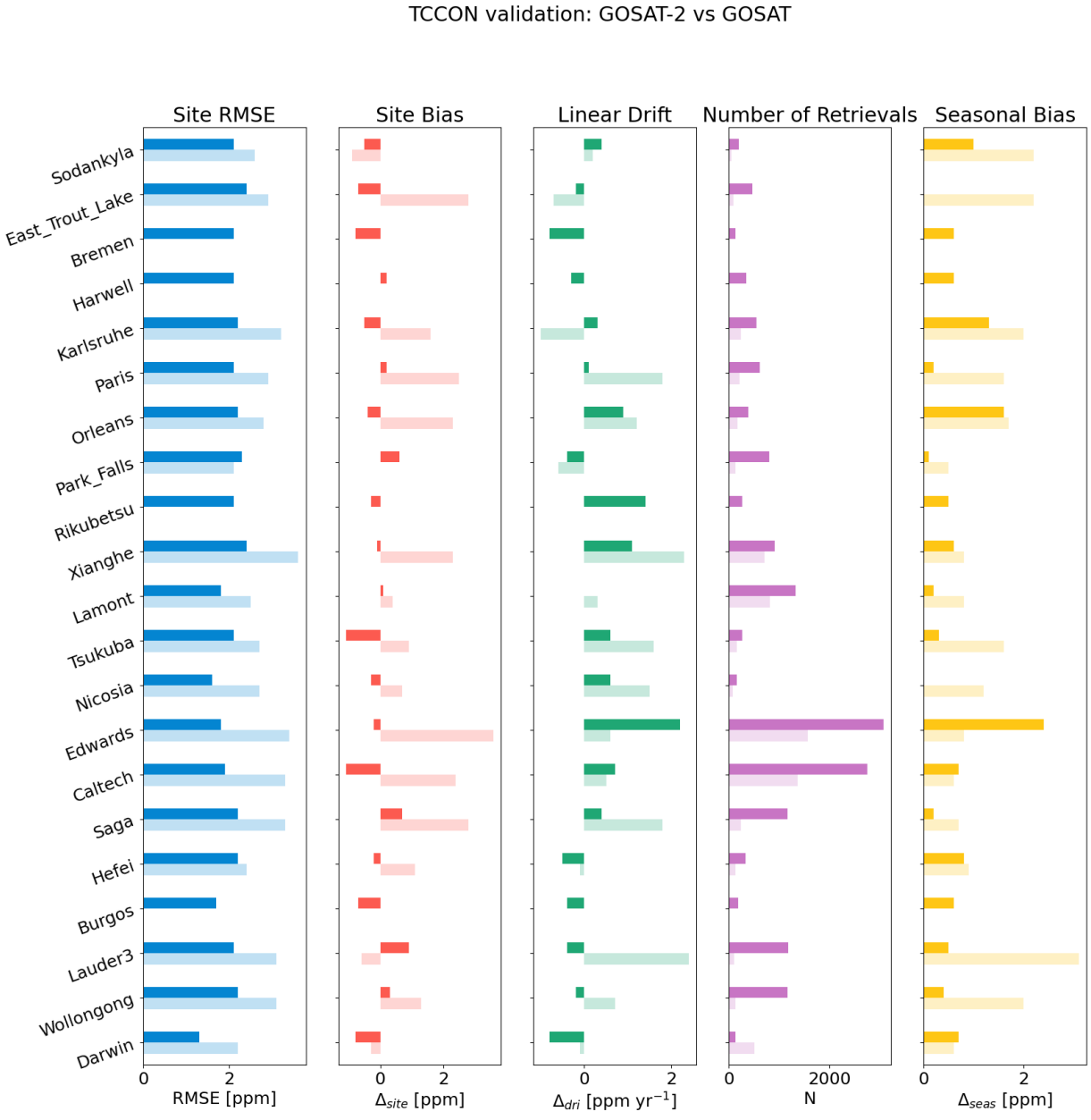


**Figure 5.** Overview of the bias parametrisation for the Full Physics $XCO_2$ product, per station. Shown in blue is the RMSE, red the site bias $\Delta_{site}$, green the linear drift $\Delta_{dri}$, yellow the seasonal bias $\Delta_{seas}$ and in purple the number of retrievals. Values for GOSAT-2 are shown in bold bars, and those of GOSAT are in light bars. Stations are listed in order of decreasing latitude. Missing bars correspond to less than 50 colocations for that station, therefore we do not calculate the values there.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


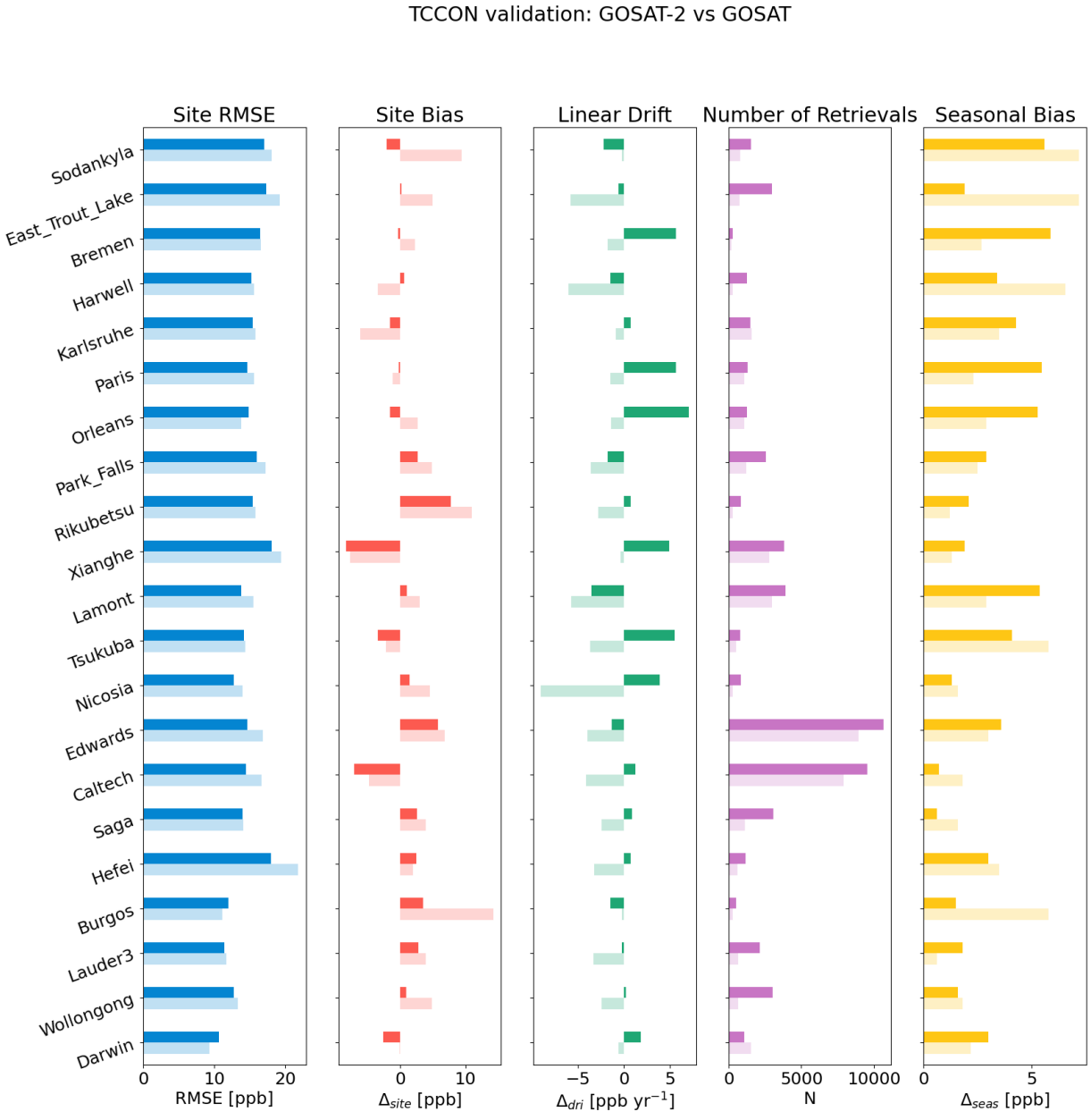


**Figure 6.** Overview of the bias parametrisation for the Proxy $XCH_4$ product, per station. Shown in blue is the RMSE, red the site bias $\Delta_{site}$, green the linear drift $\Delta_{dri}$, yellow the seasonal bias $\Delta_{seas}$ and in purple the number of retrievals. Values for GOSAT-2 are shown in bold bars, and those of GOSAT are in light bars. Stations are listed in order of decreasing latitude. Missing bars correspond to less than 50 colocations for that station, therefore we do not calculate the values there.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


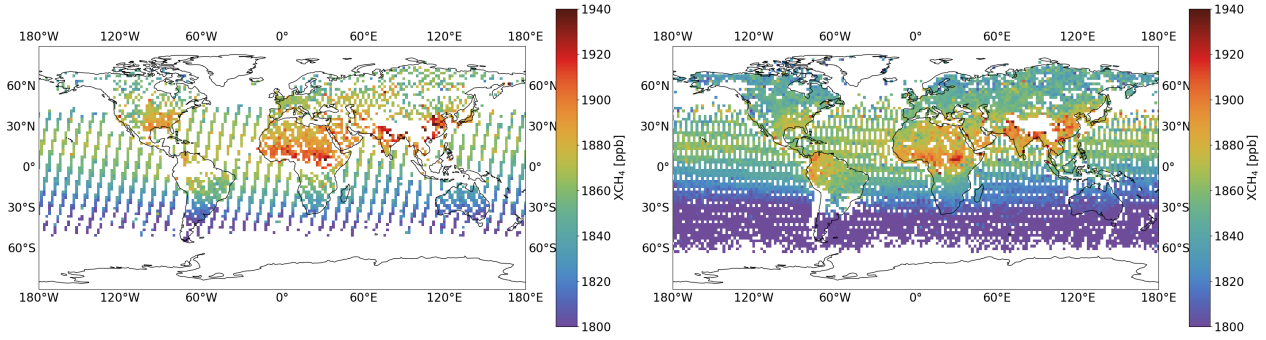


**Figure 7.** GOSAT-GOSAT-2 comparison for the GOSAT-2 $XCH_4$ Full Physics product. Maps are shown of $XCH_4$ over the year 2020 averaged onto $2^\circ \times 2^\circ$ boxes for GOSAT and GOSAT-2 on the left and right respectively.

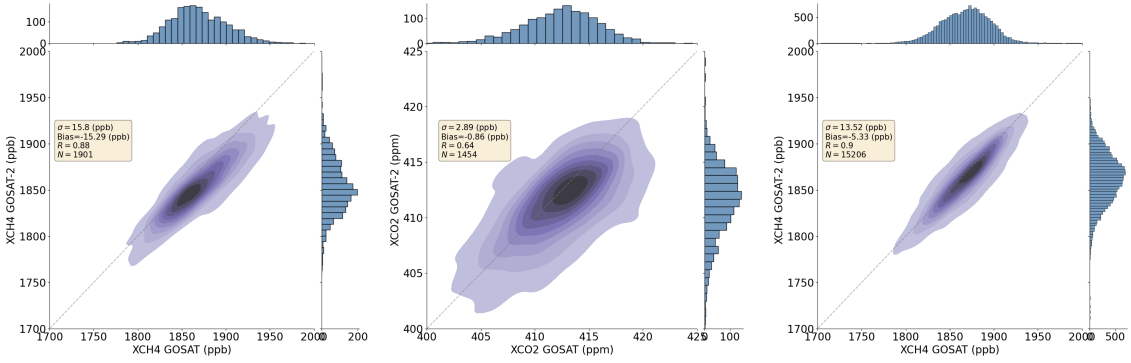


**Figure 8.** Correlation between GOSAT and GOSAT-2 shown as a kernel density estimation (KDE) plots for each data product. Plots for $XCH_4$ Full Physics, $XCO_2$ Full Physics and $XCH_4$ Proxy are shown from left to right, respectively. The mean bias, standard deviation, number of points and correlation coefficient of the population are also quoted. Histograms of the number of counts are shown around the margin, along with the linear regression and the 1-to-1 lines in black and grey respectively. Results are for soundings over land.

Furthermore, we plot time-series of GOSAT and GOSAT-2 globally, and for the three latitude bands of Northern/Southern Hemispheres (NH & SH) and the Tropics in Figure 9. These are defined as $0^\circ$ to $60^\circ$N for NH, $-25^\circ$N to $25^\circ$N for the Tropics, and $-60^\circ$N to $0^\circ$ for SH.

The globally averaged seasonal cycles of $XCH_4$ Full Physics follow each other well between April and August, but from September to March, the GOSAT one peaks at higher values. This characteristic is representative of the Tropics and SH time-series, however for the NH time-series, the GOSAT time-series is consistently higher by approximately 15 ppb.

For the time-series of the Proxy products, we find that the satellite time-series correlate extremely well with each other, and the seasonal cycles follow each other closely in all latitude bands, in contrast to the Full Physics $XCH_4$ product where significant biases are observed. There is a small systematic bias of around 5 ppb that begins positive but then switches around the halfway point of the time-series.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


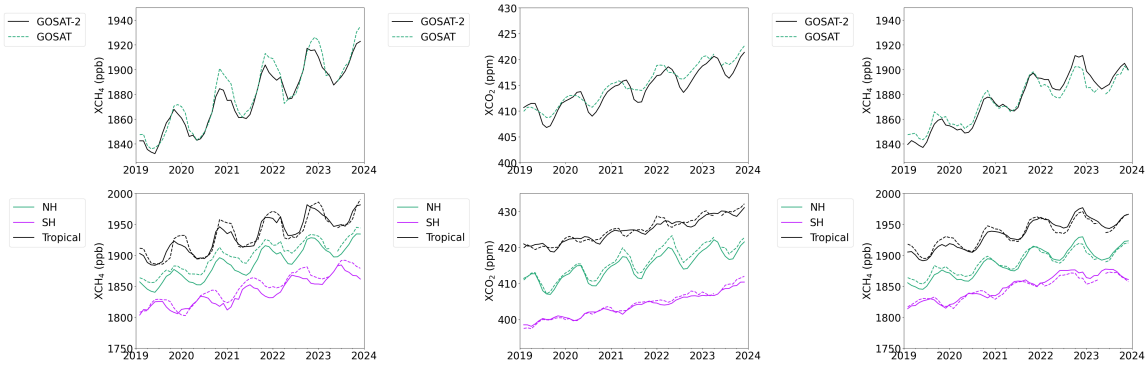


**Figure 9.** Time-series of the GOSAT and GOSAT-2 RemoTeC products. The $XCH_4$ Full Physics, $XCO_2$ Full Physics and $XCH_4$ Proxy products are shown from left to right. GOSAT-2 data are shown as solid lines, whereas GOSAT data are shown as dashed lines. The upper panels give the globally averaged monthly time-series. The lower panels give the same but split into the different latitude bands of NH, SH and Tropics. For $XCH_4$, the time-series of the Tropics are shifted up by a constant factor of +50 ppb for better visualisation. For $XCO_2$, the time-series of the Tropics is shifted by +20 ppm and the SH by -10 ppm.

For $XCO_2$, the time-series in the NH follow each other closely until mid-2020 after which the GOSAT time-series in consistently higher than GOSAT-2. The SH time-series agree well over the whole time-series, but that of the Tropics is less pronounced in GOSAT-2 with larger seasonal fluctuations exhibited for GOSAT.

We note here that we compare data products from GOSAT and GOSAT-2. We do not comment on the performance of one satellite over another as the comparison of the data products also depends on the different quality filtering applied.

### 5.3 TROPOMI Intercomparision

The fact that the RFC quality filtering models are trained on the spatially limited dataset of TCCON implies that understanding how well the models - and thus also the filtering - generalise to the global GOSAT-2 dataset, is of high priority. This is reinforced when considering that the validation data and the training data constitute essentially the same representation of data, which may lead to biases that would not be picked up by validation only with TCCON. Central to the performance of the models is the behaviour exhibited in Figure 2. Therefore if such behaviour is exhibited also on global scales, this would be good confirmation that the quality filtering performs equivalently on global dataset as it does on data colocated with TCCON.

Here we inter-compare our GOSAT-2 product against the TROPOMI operational product, version 2.4.0, and evaluate the performance of the quality filtering on global scales. The TROPOMI product was pre-filtered with VIIRS cloud product using the strictest filter of cloud fraction $< 0.001$, and quality filtered using nominal quality flags. We note that since no $XCO_2$ product exists for TROPOMI, the inter-comparison here is limited to the $XCH_4$ products. Figures 10 and 11 illustrate results for the whole of the year 2020, taking GOSAT-2 QA value equal to 0, for the Full Physics and Proxy $XCH_4$ products respectively. To evaluate the generalisation of the RFC quality filtering to global scales, we give results for the other QA values of the GOSAT-2

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


product in Table 4. Furthermore, because the RFC filtering in GOSAT-2 is only applied to soundings over land, we restrict the analysis to satellite data over land.

We find that, when considering GOSAT-2 data with QA value of 0, the global systematic bias between GOSAT-2 and TROPOMI, which we define as $XCH_{4,GOSAT-2}$-$XCH_{4,TROPOMI}$, is extremely low. We derive a global average of the bias of -5.53 ppb and 0.03 ppb for the Full Physics and Proxy products respectively. The satellite products are highly correlated with correlation coefficients above 0.87 and 0.86, and standard deviations less than 16.32 and 17.43 ppb of $XCH_4$ Full-Physics and Proxy data. Here we note that the TROPOMI operational product uses a different bias correction to GOSAT-2. The TROPOMI bias correction is based on the small area approximation (Lorente et al., 2021; O'Dell et al., 2018) taking a uniform $XCH_4$ distribution as a function of albedo in multiple regions, whereas the GOSAT-2 bias correction is based on TCCON data (equation 8).

A key result shown in Table 4 is that the QA value increases proportional to the scatter of the GOSAT2-TROPOMI differences and number of data points, while the bias remains effectively constant. This is a good reflection of the behaviour represented in Figure 2, meaning that, despite the fact that the quality filtering models are trained on the spatially limited dataset of TCCON, they generalise well to the global ensemble.

**Table 4.** Overview of inter-comparison of $XCH_4$ between GOSAT-2 and TROPOMI. Information for all QA values available to GOSAT-2 are given. $\Delta XCH_4$ is the mean bias with TROPOMI, $\sigma_{TROPOMI}$ and $N_{TROPOMI}$ are the scatter and number of TROPOMI-GOSAT-2 colocations respectively, and $N_{TCCON}$ is the number of TCCON-GOSAT-2 colocations, with $\sigma_{TCCON}$ the RMSE of the bias between GOSAT-2 and TCCON.

| $XCH_4$ Full Physics | | | | | |
|---|---|---|---|---|---|
| QA value | $\Delta XCH_4$ (ppb) | $\sigma_{TROPOMI}$ (ppb) | $N_{TROPOMI}$ | $N_{TCCON}$ | $\sigma_{TCCON}$ (ppb) |
| 0 | -4.6 | 15.0 | 22,863 | 17,250 | 15.19 |
| 0.2 | -4.8 | 15.4 | 29,539 | 22,635 | 15.98 |
| 0.4 | -5.3 | 16.5 | 38,049 | 28,943 | 17.36 |
| 0.6 | -5.7 | 17.4 | 43,059 | 32,309 | 18.53 |
| 0.8 | -6.3 | 18.5 | 47,896 | 34,578 | 19.42 |
| $XCH_4$ Proxy | | | | | |
| QA value | $\Delta XCH_4$ (ppb) | $\sigma_{TROPOMI}$ (ppb) | $N_{TROPOMI}$ | $N_{TCCON}$ | $\sigma_{TCCON}$ (ppb) |
| 0 | 1.7 | 16.6 | 76,353 | 55,915 | 15.65 |
| 0.2 | 1.8 | 17.3 | 77,540 | 63,248 | 16.30 |
| 0.4 | 1.5 | 18.2 | 88,983 | 67,607 | 16.87 |
| 0.6 | 1.2 | 19.0 | 93,385 | 70,198 | 17.28 |
| 0.8 | 1.0 | 19.7 | 97,451 | 71,884 | 17.66 |

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


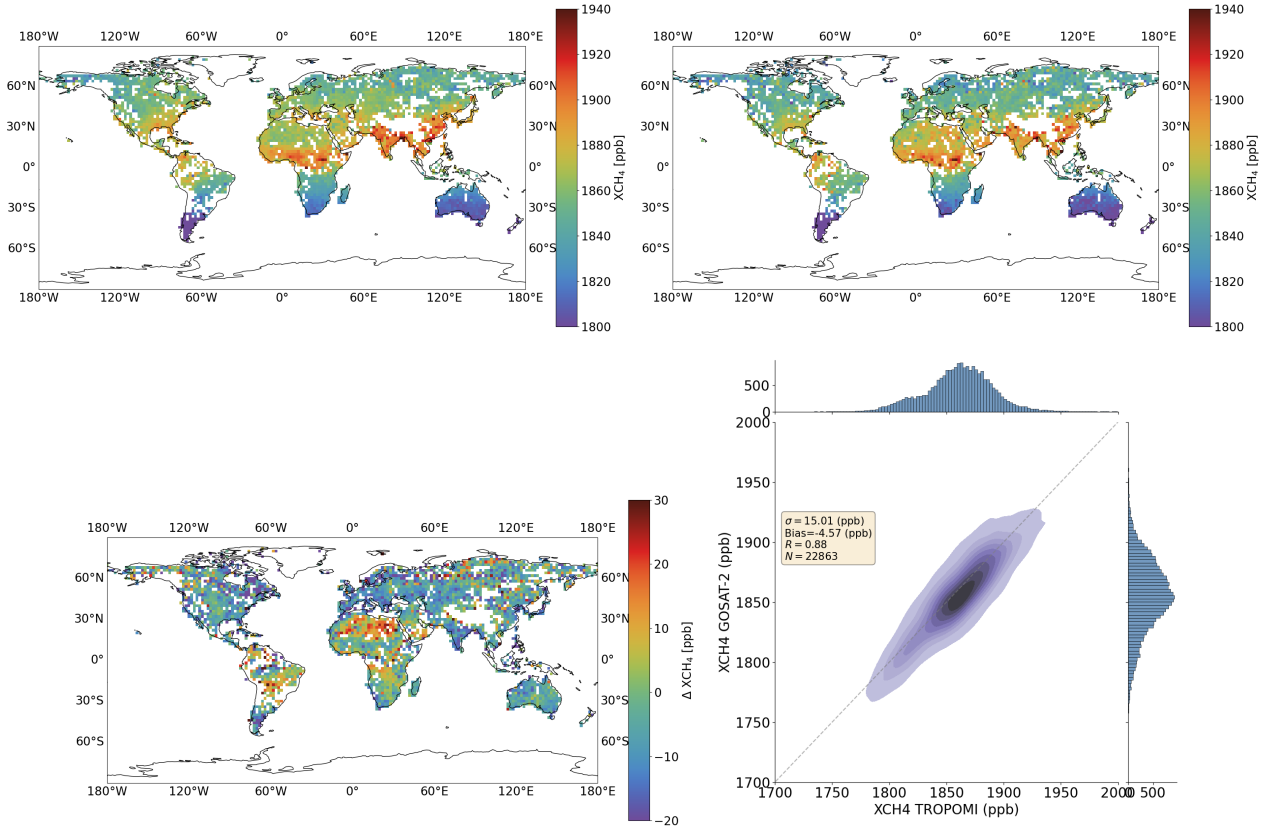


**Figure 10.** GOSAT2-TROPOMI comparison with QA value equal to 0 for the GOSAT-2 Full Physics product. *upper left:* Map of TROPOMI $XCH_4$ daily averages colocated with GOSAT-2, over the year 2020, sampled on $2^\circ \times 2^\circ$ boxes. *upper right:* Map of GOSAT-2 $XCH_4$ daily averages colocated with TROPOMI, over the year 2020, sampled on $2^\circ \times 2^\circ$ boxes. *lower left:* Map of the difference between satellite data defined as $XCH_{4,\text{GOSAT-2}}$-$XCH_{4,\text{TROPOMI}}$. *lower right:* Correlation between all colocated $XCH_4$ measurements over 2020, shown as a kernel density estimation (KDE) plot. The mean bias, standard deviation, number of points and correlation coefficient of the population are also quoted. Histograms of the number of counts are shown around the margins, along with the linear regression and the 1-to-1 lines in black and grey respectively.

From the global maps, significant differences between TROPOMI and GOSAT-2 are obvious over Northern. Central Africa has a similar characteristic of high bias, and we speculate that these differences are a consequence of dust and burning events that lead to high aerosol load complicating the retrieval. This conclusion would be consistent with the fact that the biases are larger for the Proxy GOSAT-2 product than the Full Physics product, in which aerosols are better characterised. The reason for low coverage and high bias over the Amazon can be a result of low surface albedo and observations that are contaminated by high water vapour.

The global average is close to zero, in contrast to what we observe for GOSAT. Systematic biases of -13 ppb are found between TROPOMI and GOSAT for a global average (Hu et al., 2018; Lorente et al., 2021). A similar bias is found, in both

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


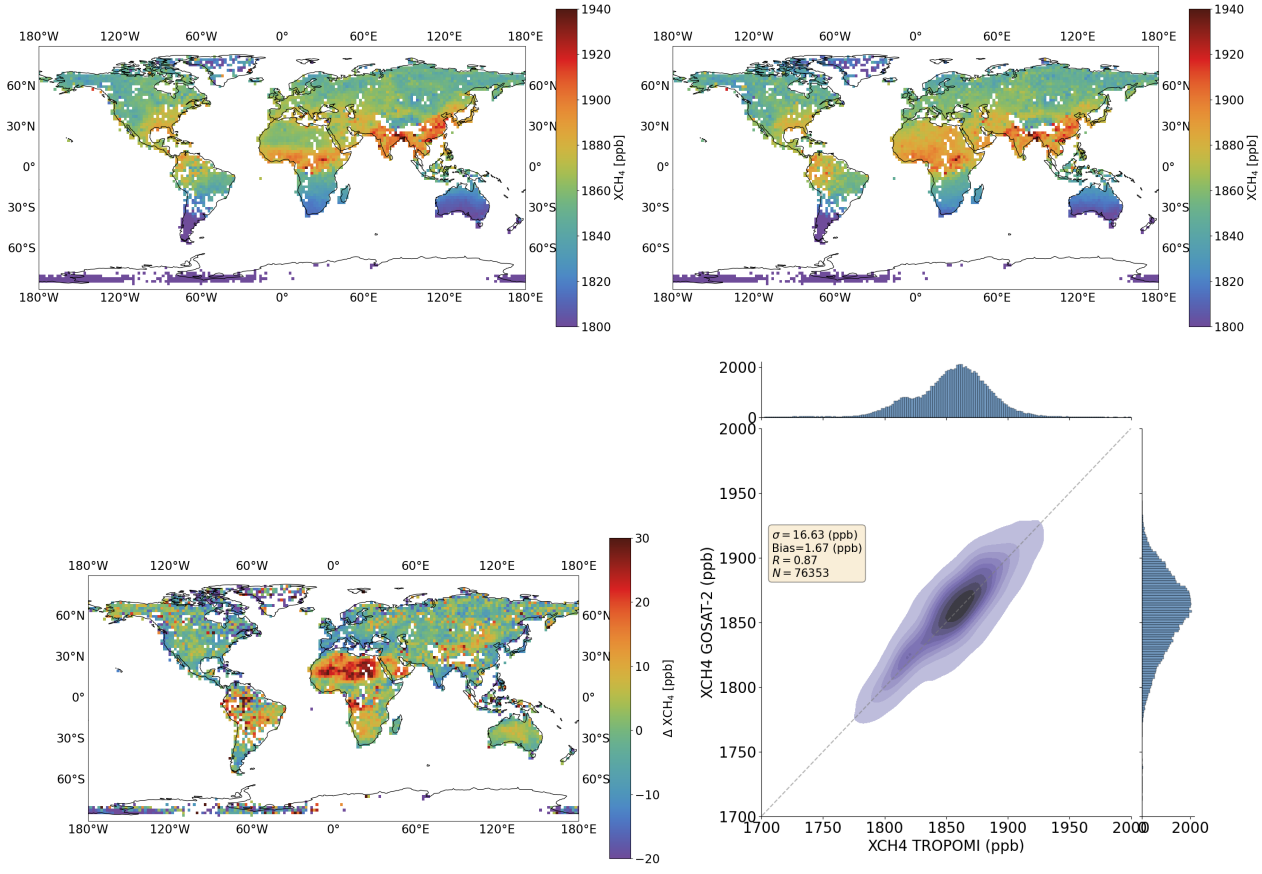


**Figure 11.** The same as Figure 10 but for the GOSAT-2 Proxy product.

sign and magnitude, between GOSAT and GOSAT-2 (section 5.2). Both, TROPOMI and GOSAT-2 $XCH_4$ are globally below the corresponding XCH4 GOSAT data. We propose therefore that the bias observed between GOSAT and GOSAT-2 comes from systematic biases in the GOSAT $XCH_4$ products, consistent with the results of TCCON validation presented in Figures 4 and 6.

## 6 Conclusions

In this article, we have presented total column mixing ratio data products from GOSAT-2, retrieved with the RemoTeC algorithm. From the two retrieval approaches of RemoTeC, three products are extracted; $XCH_4$ and $XCO_2$ from the Full Physics retrieval and $XCH_4$ from the Proxy retrieval. The time-series of these products span five years, from 2019 to 2023. All three products are validated with TCCON and inter-compared to GOSAT and TROPOMI and the long time-series ensures robust results from each.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


The RMSE between GOSAT-2 and TCCON of both the $XCH_4$ products is below 15 ppb, with the Proxy product having more data by a factor of 3, and the RMSE of $XCO_2$ is 2 ppm. We derive station-to-station biases of 4.2 ppb and 0.5 ppm for the $XCH_4$ and $XCO_2$ Full Physics products respectively, and 3.7 ppb for the Proxy product. Finally we quantify the linear drift as 0.6 ppb $yr^{-1}$, 0.2 ppm $yr^{-1}$ and or $XCH_4$ Full Physics, $XCO_2$ Full Physics and $XCH_4$ Proxy products respectively.

In comparison to GOSAT, the GOSAT-2 $XCH_4$ Full Physics product shows large differences, with a global average bias of -15 ppb. This is less for the Proxy product and on the order of -5 ppb. Compared to TROPOMI, GOSAT-2 is in excellent agreement, with average global biases of -5 ppb and 0 ppb for the Full Physics and Proxy GOSAT-2 products respectively. High correlation coefficients above 0.85, and standard deviations less than 17 ppb are derived for GOSAT-2 compared to TROPOMI.

Finally, we present a new quality filtering based on a machine learning approach. Training data for the random forest classifier models are taken from TCCON colocations with GOSAT-2, where we classify good/bad quality retrievals through the bias with TCCON. Since TCCON data are also used to validate the products, we train separate models to quality filter each year of data, to avoid compromising any independent validation.

Multiple QA values are implemented by training models with different training thresholds. Increasing the QA value leads to more data at the cost of worsening the RMSE with TCCON. In this way, users can choose between higher data yield or better quality data, which may have different advantages depending on the use of the data product.

## Appendix A: GOSAT-2 data over Ocean

Despite the low surface albedo, satellite measurements over ocean are possible when operating the satellite in sunglint mode. Sunglint observations take advantage of specific viewing angles where the radiance of back-scattered sunlight is higher due to reflection from waves. This amplifies the albedo, allowing retrievals over ocean to be carried out, where the albedo is generally too low to retrieve accurate concentrations.

TCCON stations are located only on land, therefore validation of sunglint observations are only possible using stations that are close to shorelines, or on islands. In this section, the results of the TCCON validation for sunglint mode, for all three RemoTeC GOSAT-2 products, are presented.

The RMSE for ocean measurements is higher than over land, although correlation coefficients are comparable. For $XCH_4$, this is more obvious for the Full Physics product, compared to the Proxy product, with 3 ppb difference in RMSE between ocean and land. We note that such statistics are drawn only from a handful of TCCON stations due to the limited availability of TCCON data close to the ocean. As mentioned in section 4.2.1, GOSAT-2 measurements in sunglint mode are quality filtered using the threshold criteria described in section 4.1.

## Appendix B: Supplementary Material of TCCON Validation

Here we provide additional information on the validation of GOSAT-2 products with TCCON. Table B1 lists all the TCCON stations used in the analysis. Data from all stations are also used as input to train the RFC quality filtering networks. Figures

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


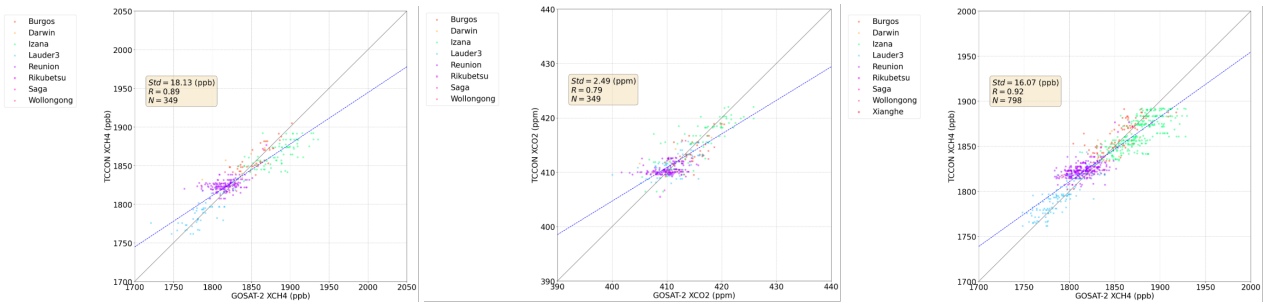


**Figure A1.** GOSAT-2 plotted against TCCON for the Full Physics $XCH_4$, Full Physics $XCO_2$ and Proxy $XCH_4$ products from left to right respectively. Data are compared only if they are fully colocated in space and time. The standard deviation of the population, Pearson's correlation coefficient and number of retrievals are given in the inset. The legend plots the different TCCON stations.

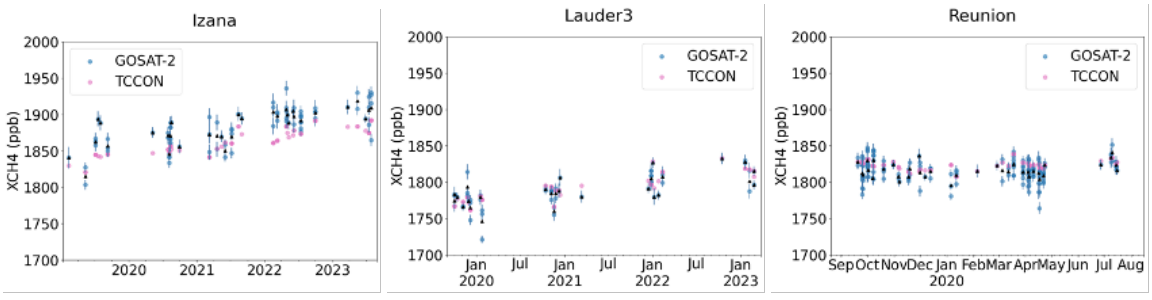


**Figure A2.** Time-series of GOSAT-2 colocated measurements over ocean with TCCON stations for the $XCH_4$ Full Physics retrievals. All GOSAT-2 observations are taken in sunglint mode. Pink circles correspond to the daily average of TCCON soundings that are spatio-temporally colocated with GOSAT-2. All individual GOSAT-2 sounding coloated with TCCON are plotted as blue circles, and the daily average of these are given as black triangles.

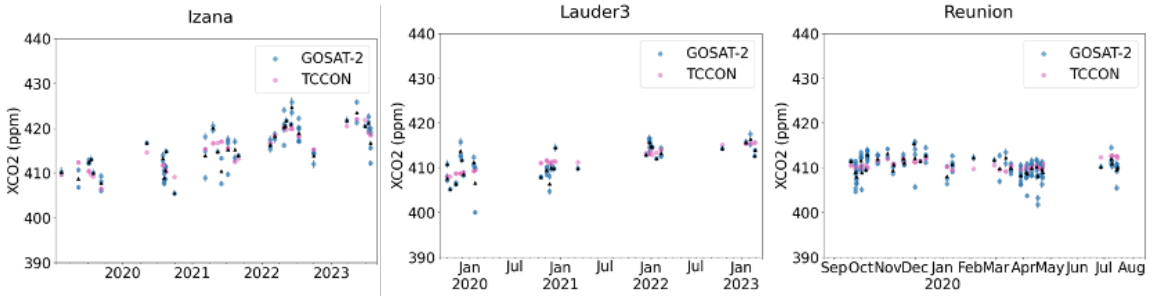


**Figure A3.** Same as Figure A2 but for the $XCO_2$ Full Physics product.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


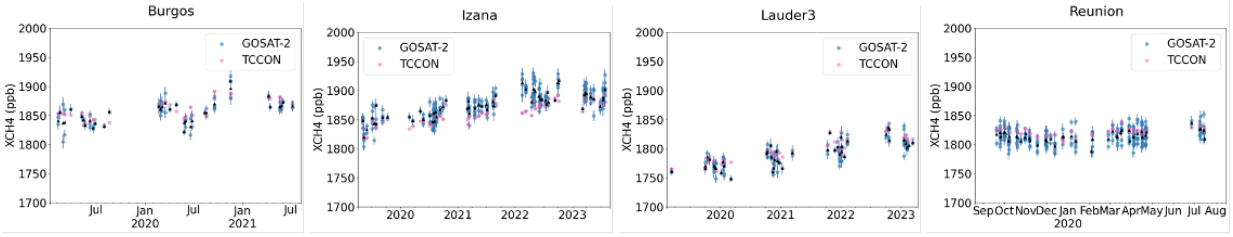


**Figure A4.** Same as Figure A2 but for the $XCH_4$ Proxy product.

B1 to B3 present time-series of GOSAT-2 compared to TCCON for all stations for the $XCH_4$ Full Physics, $XCO_2$ Full Physics and $XCH_4$ Proxy products respectively. When enough TCCON data is available, time-series span the full 5 year period from 2019 to 2023.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


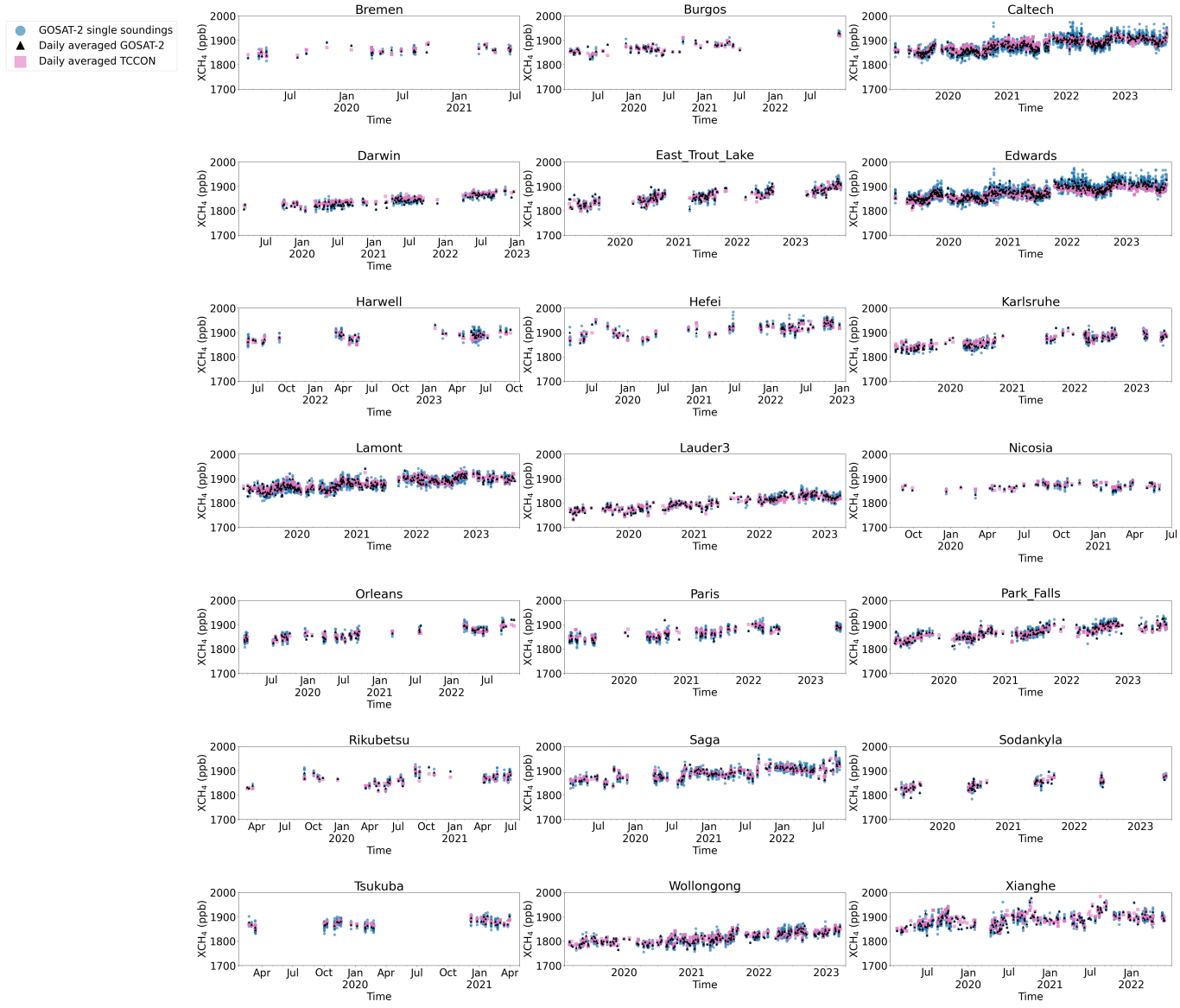


**Figure B1.** Time-series of GOSAT-2 colocated measurements with TCCON stations for the $XCH_4$ Full Physics retrievals. Pink squares correspond to the daily average of TCCON soundings that are spatio-temporally colocated with GOSAT-2. All individual GOSAT-2 sounding coloated with TCCON are plotted as blue circles, and the daily average of these are given as black triangles.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


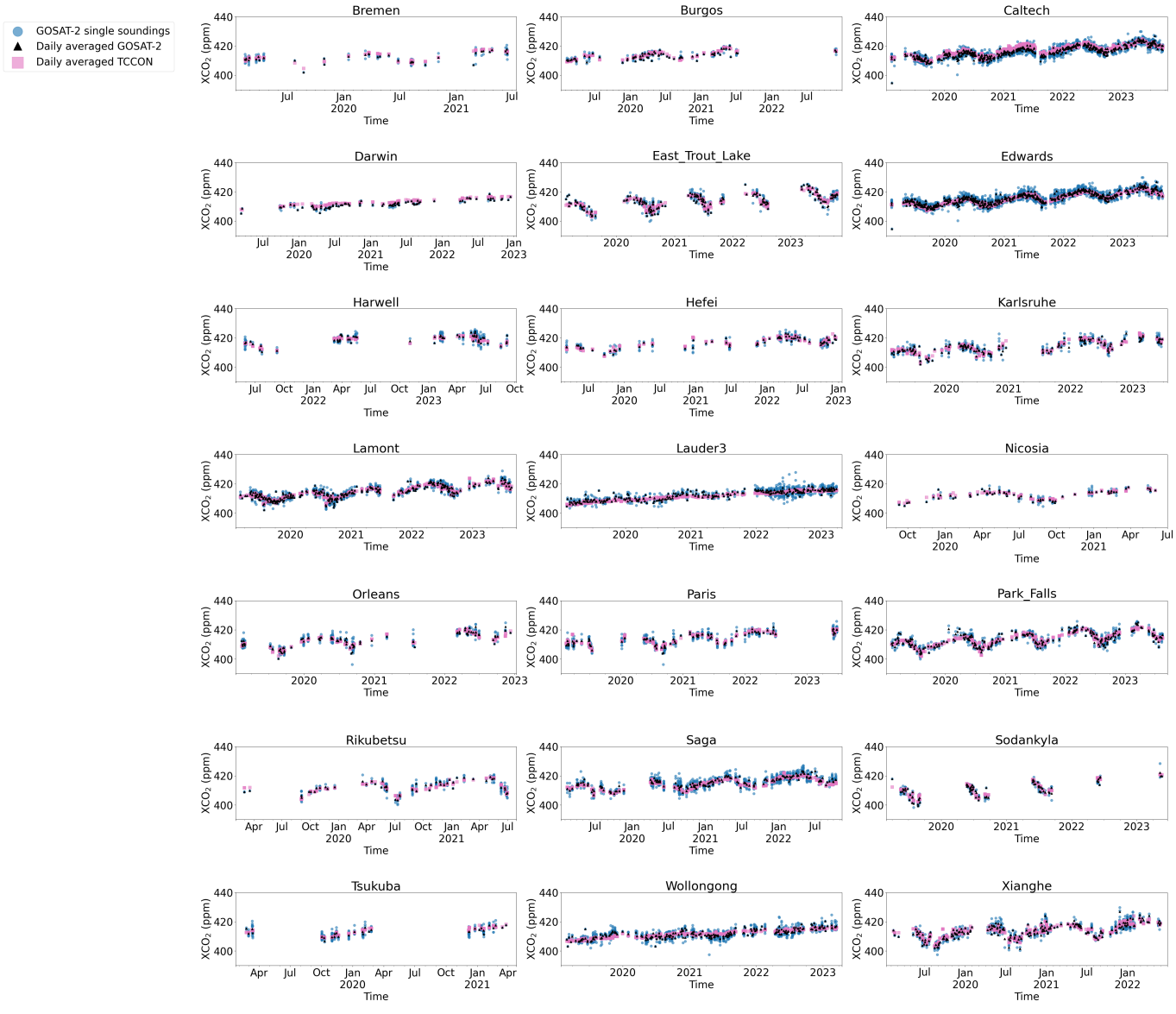


**Figure B2.** Same as Figure B1 but for the $XCO_2$ Full Physics product.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


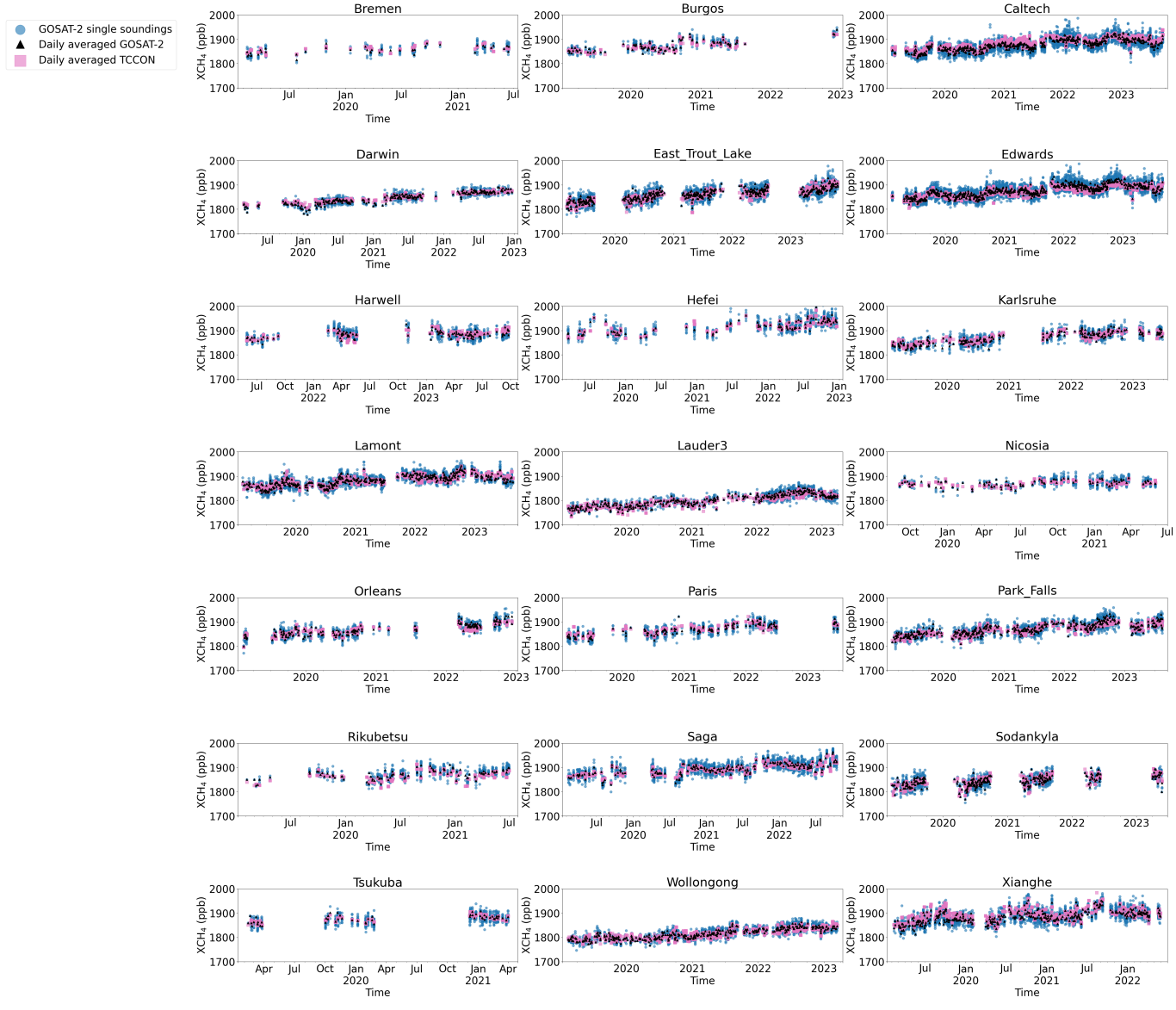


**Figure B3.** Same as Figure B1 but for the $XCH_4$ Proxy product.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


**Table B1.** List of TCCON stations used in training the quality filtering model and/or validation.

| Site (Country) | Coordinates (lat, lon°) | Reference |
|---|---|---|
| Bremen (Germany) | [53, 8.85] | (Notholt et al., 2022) |
| Burgos (Phillipines) | [18.53, 120.65] | (Morino et al., 2022c) |
| Caltech (USA) | [34.14, -118.13] | (Wennberg et al., 2022a) |
| Darwin (Australia) | [-12.42, 130.89] | (Deutscher et al., 2023b) |
| East Trout Lake (Canada) | [54.35, -104.99] | (Wunch et al., 2022) |
| Edwards (USA) | [34.96, -117.88] | (Iraci et al., 2022) |
| Eureka (Canada) | [80.05, -86.42] | (Strong et al., 2022) |
| Garmisch (Germany) | [47.48, 11.06] | (Sussmann and Rettinger, 2023) |
| Harwell (UK) | [51.57, -1.32] | (Weidmann et al., 2023) |
| Hefei (China) | [31.9, 119.17] | (Liu et al., 2023) |
| Izana (Spain) | [28.30, 16.50] | (García et al., 2022) |
| Karlsruhe (Germany) | [49.10, 8.44] | (Hase et al., 2023) |
| Lamont (USA) | [36.60, -97.49] | (Wennberg et al., 2022c) |
| Lauder (New Zealand) | [-45.04, 169.68] | (Pollard et al., 2022) |
| Nicosia (Cyprus) | [35.14, 33.38] | (Petri et al., 2024) |
| Ny Ålesund (Norway) | [78.92,11.92] | (Buschmann et al., 2022) |
| Orleans (France) | [47.97, 2.11] | (Warneke et al., 2022) |
| Paris (France) | [48.49, 2.36] | (Té et al., 2022) |
| Park Falls (USA) | [45.95, -90.27] | (Wennberg et al., 2022b) |
| Reunion Island (France) | [-20.9, 55.48] | (De Mazière et al., 2022) |
| Rikubetsu (Japan) | [43.46, 143.77] | (Morino et al., 2022a) |
| Saga (Japan) | [33.24, 130.29] | (Shiomi et al., 2022) |
| Sodankylå (Finland) | [67.37, 26.63] | (Kivi et al., 2022) |
| Tsukuba (Japan) | [36.05, 140.12] | (Morino et al., 2022b) |
| Wollongong (Australia) | [-34.41, 150.88] | (Deutscher et al., 2023a) |
| Xianghe (China) | [39.80, 116.96] | (Zhou et al., 2022) |

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


*Author contributions.* AB ran the retrievals to obtain level 2 data from GOSAT-2, conducted the analysis on all three products and wrote the paper. MM provided the TROPOMI operational prodcut data. TB and JL gave guidance in writing the paper and significant input into the configuration and architecture of the retrieval. All other co-authors provided data from the TCCON network.

*Competing interests.* Three of the co-authors are members of the editorial board for Atmospheric Measurement Techniques in the subject area of Gases.

*Acknowledgements.* The research in this manuscript was supported by funding from the European Space Agency (ESA) via the project GHG-CCI+ (ESA contract no. 4000126450/19/I-NB). We thank the European Center for Medium Range Weather Forecasts (ECMWF) for providing the ERA5 reanalysis data. We also thank ESA and JAXA/NIES for providing us with the GOSAT-2 level 1b data. MV acknowledges Horizon 2020 (grant no. EMME-CARE – Eastern Mediterranean and Middle East – Climate and Atmosphere Research Centre (856612)).

*Data availability.* All three GOSAT-2 products are operationally provided as part of the Copernicus Climate Change Service (C3S) and can be downloaded from https://zenodo.org/records/12180512 under version 2.1.0 (DOI:10.5281/zenodo.12180512). Alternatively all three products will be made available on the CEDA Data Archive, https://catalogue.ceda.ac.uk/, as part of Climate Change Initiative plus (CCI+) under version v2.0.3.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


## References

Banks, H. and Wood, R.: Where to Look for Anthropogenic Climate Change in the Ocean., Journal of Climate, 15 (8), 879–891, 2002.

Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P., Langenfelds, R., Torn, M., Biraud, S., Stephens, B., Andrews, A., and Worthy, D.: Global $CO_2$ fluxes estimated from GOSAT retrievals of total column $CO_2$, Atmospheric Chemistry and Physics, 13, 8695–8717, https://doi.org/10.5194/acp-13-8695-2013, 2013.

Bergamaschi, P., Frankenberg, C., Meirink, J. F., Krol, M., Villani, M. G., Houweling, S., Dentener, F., Dlugokencky, E. J., Miller, J. B., Gatti, L. V., Engel, A., and Levin, I.: Inverse modeling of global and regional CH4 emissions using SCIAMACHY satellite retrievals, Journal of Geophysical Research: Atmospheres, 114, https://doi.org/https://doi.org/10.1029/2009JD012287, 2009.

Bradley, A. P.: The use of the area under the ROC curve in the evaluation of machine learning algorithms, Pattern Recognition, 30, 1145–1159, https://doi.org/https://doi.org/10.1016/S0031-3203(96)00142-2, 1997.

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Buchwitz, M., Reuter, M., Schneising, O., Boesch, H., Guerlet, S., Dils, B., Aben, I., Armante, R., Bergamaschi, P., Blumenstock, T., Bovensmann, H., Brunner, D., Buchmann, B., Burrows, J., Butz, A., Chédin, A., Chevallier, F., Crevoisier, C., Deutscher, N., Frankenberg, C., Hase, F., Hasekamp, O., Heymann, J., Kaminski, T., Laeng, A., Lichtenberg, G., De Mazière, M., Noël, S., Notholt, J., Orphal, J., Popp, C., Parker, R., Scholze, M., Sussmann, R., Stiller, G., Warneke, T., Zehner, C., Bril, A., Crisp, D., Griffith, D., Kuze, A., O'Dell, C., Oshchepkov, S., Sherlock, V., Suto, H., Wennberg, P., Wunch, D., Yokota, T., and Yoshida, Y.: The Greenhouse Gas Climate Change Initiative (GHG-CCI): Comparison and quality assessment of near-surface-sensitive satellite-derived CO2 and CH4 global data sets, Remote Sensing of Environment, 162, 344–362, https://doi.org/https://doi.org/10.1016/j.rse.2013.04.024, 2015.

Buschmann, M., Petri, C., Palm, M., Warneke, T., and Notholt, J.: TCCON data from Ny-Ålesund, Svalbard (NO), Release GGG2020.R0, https://doi.org/10.14291/tccon.ggg2020.nyalesund01.R0, 2022.

Butz, A., Hasekamp, O. P., Frankenberg, C., and Aben, I.: Retrievals of atmospheric CO2 from simulated space-borne measurements of backscattered near-infrared sunlight: accounting for aerosol effects, Appl. Opt., 48, 3322–3336, https://doi.org/10.1364/AO.48.003322, 2009.

Butz, A., Hasekamp, O. P., Frankenberg, C., Vidot, J., and Aben, I.: CH4 retrievals from space-based solar backscatter measurements: Performance evaluation against simulated aerosol and cirrus loaded scenes, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/https://doi.org/10.1029/2010JD014514, 2010.

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, Geophysical Research Letters, 38, https://doi.org/https://doi.org/10.1029/2011GL047888, 2011.

Chevallier, F.: Contribution to documentation of products and services as provided within the scope of this contract – 2023 – Part CO2, ECMWF, https://atmosphere.copernicus.eu/sites/default/files/2023-10/CAMS255_2021SC1_D55.5.2.1-2023-PartCO2_Contribution%20to%20documentation%20of%20products%20and%20services%20as%20provided%20within%20the%20scope%20of%20this%20contract%20-%202023_v1_APPROVED_Ver1.pdf, 2023.

Chevallier, F., Engelen, R. J., and Peylin, P.: The contribution of AIRS data to the estimation of CO2 sources and sinks, Geophysical Research Letters, 32, https://doi.org/https://doi.org/10.1029/2005GL024229, 2005.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


Chevallier, F., Bréon, F.-M., and Rayner, P. J.: Contribution of the Orbiting Carbon Observatory to the estimation of CO2 sources and sinks: Theoretical study in a variational data assimilation framework, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/https://doi.org/10.1029/2006JD007375, 2007.

Chevallier, F., Maksyutov, S., Bousquet, P., Bréon, F.-M., Saito, R., Yoshida, Y., and Yokota, T.: On the accuracy of the CO2 surface fluxes to be estimated from the GOSAT observations, Geophysical Research Letters, 36, https://doi.org/https://doi.org/10.1029/2009GL040108, 2009.

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, funding by Belgian Federal Science Policy Office GRID grid.425119.a., 2022.

Detmers, R. G., Hasekamp, O., Aben, I., Houweling, S., van Leeuwen, T. T., Butz, A., Landgraf, J., Köhler, P., Guanter, L., and Poulter, B.: Anomalous carbon uptake in Australia as seen by GOSAT, Geophysical Research Letters, 42, 8177–8184, https://doi.org/https://doi.org/10.1002/2015GL065161, 2015.

Deutscher, N. M., Griffith, D. W. T., Bryant, G. W., Wennberg, P. O., Toon, G. C., Washenfelder, R. A., Keppel-Aleks, G., Wunch, D., Yavin, Y., Allen, N. T., Blavier, J.-F., Jiménez, R., Daube, B. C., Bright, A. V., Matross, D. M., Wofsy, S. C., and Park, S.: Total column CO2 measurements at Darwin, Australia; site description and calibration against in situ aircraft profiles, Atmospheric Measurement Techniques, 3, 947–958, https://doi.org/10.5194/amt-3-947-2010, 2010.

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., and et al.: TCCON data from Wollongong (AU), Release GGG2020.R0, https://doi.org/10.14291/tccon.ggg2020.wollongong01.R0, funding by Australian Research Council GRID grid.413452.5., 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., and et al.: TCCON data from Darwin (AU), Release GGG2020.R0, https://doi.org/10.14291/tccon.ggg2020.darwin01.R0, funding by Australian Research Council GRID grid.413452.5., 2023b.

Dils, B., De Mazière, M., Müller, J. F., Blumenstock, T., Buchwitz, M., de Beek, R., Demoulin, P., Duchatelet, P., Fast, H., Frankenberg, C., Gloudemans, A., Griffith, D., Jones, N., Kerzenmacher, T., Kramer, I., Mahieu, E., Mellqvist, J., Mittermeier, R. L., Notholt, J., Rinsland, C. P., Schrijver, H., Smale, D., Strandberg, A., Straume, A. G., Stremme, W., Strong, K., Sussmann, R., Taylor, J., van den Broek, M., Velazco, V., Wagner, T., Warneke, T., Wiacek, A., and Wood, S.: Comparisons between SCIAMACHY and ground-based FTIR data for total columns of CO, $CH_4$, $CO_2$ and $N_2O$, Atmospheric Chemistry and Physics, 6, 1953–1976, https://doi.org/10.5194/acp-6-1953-2006, 2006.

Dils, B., Buchwitz, M., Reuter, M., Schneising, O., Boesch, H., Parker, R., Guerlet, S., Aben, I., Blumenstock, T., Burrows, J. P., Butz, A., Deutscher, N. M., Frankenberg, C., Hase, F., Hasekamp, O. P., Heymann, J., De Mazière, M., Notholt, J., Sussmann, R., Warneke, T., Griffith, D., Sherlock, V., and Wunch, D.: The Greenhouse Gas Climate Change Initiative (GHG-CCI): comparative validation of GHG-CCI SCIAMACHY/ENVISAT and TANSO-FTS/GOSAT $CO_2$ and $CH_4$ retrieval algorithm products with measurements from the TCCON, Atmospheric Measurement Techniques, 7, 1723–1744, https://doi.org/10.5194/amt-7-1723-2014, 2014.

Frankenberg, C., Platt, U., and Wagner, T.: Iterative maximum a posteriori (IMAP)-DOAS for retrieval of strongly absorbing trace gases: Model studies for $CH_4$ and $CO_2$ retrieval from near infrared spectra of SCIAMACHY onboard ENVISAT, Atmospheric Chemistry and Physics, 5, 9–22, https://doi.org/10.5194/acp-5-9-2005, 2005.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


Frankenberg, C., Aben, I., Bergamaschi, P., Dlugokencky, E. J., van Hees, R., Houweling, S., van der Meer, P., Snel, R., and Tol, P.: Global column-averaged methane mixing ratios from 2003 to 2009 as derived from SCIAMACHY: Trends and variability, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/https://doi.org/10.1029/2010JD014849, 2011.

García, O. E., Schneider, M., Herkommer, B., Gross, J., Hase, F., Blumenstock, T., and Sepúlveda, E.: TCCON data from Izana (ES), Release GGG2020.R1, https://doi.org/10.14291/tccon.ggg2020.izana01.R1, 2022.

GCOS: The global observing system for climate: implementation needs., World Meteo- rological Organization (WMO), 2016.

Geibel, M. C., Messerschmidt, J., Gerbig, C., Blumenstock, T., Chen, H., Hase, F., Kolle, O., Lavrič, J. V., Notholt, J., Palm, M., Rettinger, M., Schmidt, M., Sussmann, R., Warneke, T., and Feist, D. G.: Calibration of column-averaged $CH_4$ over European TCCON FTS sites with airborne in-situ measurements, Atmospheric Chemistry and Physics, 12, 8763–8775, https://doi.org/10.5194/acp-12-8763-2012, 2012.

Giorgi, F. and Bi, X.: Time of emergence (TOE) of GHG-forced precipitation change hot-spots, Geophysical Research Letters, 36, https://doi.org/https://doi.org/10.1029/2009GL037593, 2009.

Gogoi, M. M., Babu, S. S., Imasu, R., and Hashimoto, M.: Satellite (GOSAT-2 CAI-2) retrieval and surface (ARFINET) observations of aerosol carbon over India, Atmospheric Chemistry and Physics, 23, 8059–8079, https://doi.org/10.5194/acp-23-8059-2023, 2023.

Guerlet, S., Butz, A., Schepers, D., Basu, S., Hasekamp, O. P., Kuze, A., Yokota, T., Blavier, J.-F., Deutscher, N. M., Griffith, D. W., Hase, F., Kyro, E., Morino, I., Sherlock, V., Sussmann, R., Galli, A., and Aben, I.: Impact of aerosol and thin cirrus on retrieving and validating XCO2 from GOSAT shortwave infrared measurements, Journal of Geophysical Research: Atmospheres, 118, 4887–4905, https://doi.org/https://doi.org/10.1002/jgrd.50332, 2013.

Hase, F., Herkommer, B., Groß, J., Blumenstock, T., Kiel, M., and Dohe, S.: TCCON data from Karlsruhe (DE), Release GGG2020.R1, https://doi.org/10.14291/tccon.ggg2020.karlsruhe01.R1, 2023.

Hasekamp, O. P. and Landgraf, J.: Linearization of vector radiative transfer with respect to aerosol properties and its use in satellite remote sensing, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/https://doi.org/10.1029/2004JD005260, 2005.

Hawkins, E. and Sutton, R.: Time of emergence of climate signals, Geophysical Research Letters, 39, https://doi.org/https://doi.org/10.1029/2011GL050087, 2012.

Hawkins, E., Frame, D., Harrington, L., Joshi, M., King, A., Rojas, M., and Sutton, R.: Observed Emergence of the Climate Change Signal: From the Familiar to the Unknown, Geophysical Research Letters, 47, e2019GL086 259, https://doi.org/https://doi.org/10.1029/2019GL086259, e2019GL086259 2019GL086259, 2020.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803, 2020.

Hu, H., Hasekamp, O., Butz, A., Galli, A., Landgraf, J., Aan de Brugh, J., Borsdorff, T., Scheepmaker, R., and Aben, I.: The operational methane retrieval algorithm for TROPOMI, Atmospheric Measurement Techniques, 9, 5423–5440, https://doi.org/10.5194/amt-9-5423-2016, 2016.

Hu, H., Landgraf, J., Detmers, R., Borsdorff, T., Aan de Brugh, J., Aben, I., Butz, A., and Hasekamp, O.: Toward Global Mapping of Methane With TROPOMI: First Results and Intersatellite Comparison to GOSAT, Geophysical Research Letters, 45, 3682–3689, https://doi.org/https://doi.org/10.1002/2018GL077259, 2018.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


Huijnen, V., Williams, J., van Weele, M., van Noije, T., Krol, M., Dentener, F., Segers, A., Houweling, S., and Peters, W.: The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0, Geoscientific Model Development, 3, 445–473, https://doi.org/10.5194/gmd-3-445-2010, 2010.

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, T. 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.

IPCC AR5: Near-term Climate Change: Projections and Predictability, pp. 953–1028, Cambridge University Press, https://doi.org/10.1017/CBO9781107415324.023, 2014.

IPCC AR6 2021: Framing, Context, and Methods, pp. 147–286, Cambridge University Press, https://doi.org/10.1017/9781009157896.003, 2023a.

IPCC AR6 2021: Global Carbon and Other Biogeochemical Cycles and Feedbacks, pp. 673–816, Cambridge University Press, https://doi.org/10.1017/9781009157896.007, 2023b.

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, funding by National Aeronautics and Space Administration GRID grid.238252.c., 2022.

Jiang, F., Wang, H., Chen, J. M., Ju, W., Tian, X., Feng, S., Li, G., Chen, Z., Zhang, S., Lu, X., Liu, J., Wang, H., Wang, J., He, W., and Wu, M.: Regional $CO_2$ fluxes from 2010 to 2015 inferred from GOSAT $XCO_2$ retrievals using a new version of the Global Carbon Assimilation System, Atmospheric Chemistry and Physics, 21, 1963–1985, https://doi.org/10.5194/acp-21-1963-2021, 2021.

Kivi, R., Heikkinen, P., and Kyrö, E.: TCCON data from Sodankylä (FI), Release GGG2020.R0, https://doi.org/10.14291/tccon.ggg2020.sodankyla01.R0, funding by Finnish Meteorological Institute, 2022.

Kou, X., Peng, Z., Zhang, M., Hu, F., Han, X., Li, Z., and Lei, L.: The carbon sink in China as seen from GOSAT with a regional inversion system based on the Community Multi-scale Air Quality (CMAQ) and ensemble Kalman smoother (EnKS), Atmospheric Chemistry and Physics, 23, 6719–6741, https://doi.org/10.5194/acp-23-6719-2023, 2023.

Kurucz, R. L.: Synthetic Infrared Spectra, in: Infrared Solar Physics, edited by Rabin, D. M., Jefferies, J. T., and Lindsey, C., pp. 523–531, Springer Netherlands, Dordrecht, 1994.

Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T.: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, Appl. Opt., 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.

Laughner, J. L., Roche, S., Kiel, M., Toon, G. C., Wunch, D., Baier, B. C., Biraud, S., Chen, H., Kivi, R., Laemmel, T., McKain, K., Quéhé, P.-Y., Rousogenous, C., Stephens, B. B., Walker, K., and Wennberg, P. O.: A new algorithm to generate a priori trace gas profiles for the GGG2020 retrieval algorithm, Atmospheric Measurement Techniques, 16, 1121–1146, https://doi.org/10.5194/amt-16-1121-2023, 2023.

Liang, J.: Confusion Matrix: Machine Learning, POGIL Activity Clearinghouse, 3, https://pac.pogil.org/index.php/pac/article/view/304, 2022.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


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

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.

Lorente, A., Borsdorff, T., Butz, A., Hasekamp, O., aan de Brugh, J., Schneider, A., Wu, L., Hase, F., Kivi, R., Wunch, D., Pollard, D. F., Shiomi, K., Deutscher, N. M., Velazco, V. A., Roehl, C. M., Wennberg, P. O., Warneke, T., and Landgraf, J.: Methane retrieved from TROPOMI: improvement of the data product and validation of the first 2 years of measurements, Atmospheric Measurement Techniques, 14, 665–684, https://doi.org/10.5194/amt-14-665-2021, 2021.

Lyu, K., Zhang, X., Church, J. A., Slangen, A. B. A., and Hu, J.: Time of emergence for regional sea-level change, Nature Climate Change, 4, 1006–1010, https://doi.org/10.1038/nclimate2397, 2014.

Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Scarpelli, T. R., Nesser, H., Sheng, J.-X., Zhang, Y., Hersher, M., Bloom, A. A., Bowman, K. W., Worden, J. R., Janssens-Maenhout, G., and Parker, R. J.: Global distribution of methane emissions, emission trends, and OH concentrations and trends inferred from an inversion of GOSAT satellite data for 2010–2015, Atmospheric Chemistry and Physics, 19, 7859–7881, https://doi.org/10.5194/acp-19-7859-2019, 2019.

Mahlstein, I., Knutti, R., Solomon, S., and Portmann, R.: Early onset of significant local warming in low latitude countries., Environmental Research Letter., 6(3), 2011.

Malina, E., Veihelmann, B., Buschmann, M., Deutscher, N. M., Feist, D. G., and Morino, I.: On the consistency of methane retrievals using the Total Carbon Column Observing Network (TCCON) and multiple spectroscopic databases, Atmospheric Measurement Techniques, 15, 2377–2406, https://doi.org/10.5194/amt-15-2377-2022, 2022.

Meirink, J. F., Eskes, H. J., and Goede, A. P. H.: Sensitivity analysis of methane emissions derived from SCIAMACHY observations through inverse modelling, Atmospheric Chemistry and Physics, 6, 1275–1292, https://doi.org/10.5194/acp-6-1275-2006, 2006.

Messerschmidt, J., Geibel, M. C., Blumenstock, T., Chen, H., Deutscher, N. M., Engel, A., Feist, D. G., Gerbig, C., Gisi, M., Hase, F., Katrynski, K., Kolle, O., Lavrič, J. V., Notholt, J., Palm, M., Ramonet, M., Rettinger, M., Schmidt, M., Sussmann, R., Toon, G. C., Truong, F., Warneke, T., Wennberg, P. O., Wunch, D., and Xueref-Remy, I.: Calibration of TCCON column-averaged $CO_2$: the first aircraft campaign over European TCCON sites, Atmospheric Chemistry and Physics, 11, 10 765–10 777, https://doi.org/10.5194/acp-11-10765-2011, 2011.

Metz, E.-M., Vardag, S. N., Basu, S., Jung, M., Ahrens, B., El-Madany, T., Sitch, S., Arora, V. K., Briggs, P. R., Friedlingstein, P., Goll, D. S., Jain, A. K., Kato, E., Lombardozzi, D., Nabel, J. E. M. S., Poulter, B., Séférian, R., Tian, H., Wiltshire, A., Yuan, W., Yue, X., Zaehle, S., Deutscher, N. M., Griffith, D. W. T., and Butz, A.: Soil respiration–driven CO<sub>2</sub> pulses dominate Australia’s flux variability, Science, 379, 1332–1335, https://doi.org/10.1126/science.add7833, 2023.

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, funding by National Institute for Environmental Studies GRID grid.140139.e., 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, funding by National Institute for Environmental Studies GRID grid.140139.e., 2022b.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


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, funding by National Institute for Environmental Studies GRID grid.140139.e., 2022c.

Noël, S., Reuter, M., Buchwitz, M., Borchardt, J., Hilker, M., Bovensmann, H., Burrows, J. P., Di Noia, A., Suto, H., Yoshida, Y., Buschmann, M., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Hase, F., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Podolske, J. R., Pollard, D. F., Sha, M. K., Shiomi, K., Sussmann, R., Té, Y., Velazco, V. A., and Warneke, T.: $XCO_2$ retrieval for GOSAT and GOSAT-2 based on the FOCAL algorithm, Atmospheric Measurement Techniques, 14, 3837–3869, https://doi.org/10.5194/amt-14-3837-2021, 2021.

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.

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.

O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Mandrake, L., Merrelli, A., Natraj, V., Nelson, R. R., Osterman, G. B., Payne, V. H., Taylor, T. E., Wunch, D., Drouin, B. J., Oyafuso, F., Chang, A., McDuffie, J., Smyth, M., Baker, D. F., Basu, S., Chevallier, F., Crowell, S. M. R., Feng, L., Palmer, P. I., Dubey, M., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Roehl, C. M., Sha, M. K., Strong, K., Sussmann, R., Te, Y., Uchino, O., and Velazco, V. A.: Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, Atmospheric Measurement Techniques, 11, 6539–6576, https://doi.org/10.5194/amt-11-6539-2018, 2018.

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.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

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, 2024.

Philips, D. L.: A technique for the numerical solution of certain integral equations of the first kind., J. Assoc. Comput. Mach., 9, 84–97, 1962.

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.

Rohde, R. A. and Hausfather, Z.: The Berkeley Earth Land/Ocean Temperature Record, Earth System Science Data, 12, 3469–3479, https://doi.org/10.5194/essd-12-3469-2020, 2020.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


Rothman, L., Gordon, I., Barbe, A., Benner, D., Bernath, P., Birk, M., Boudon, V., Brown, L., Campargue, A., Champion, J.-P., Chance, K., Coudert, L., Dana, V., Devi, V., Fally, S., Flaud, J.-M., Gamache, R., Goldman, A., Jacquemart, D., Kleiner, I., Lacome, N., Lafferty, W., Mandin, J.-Y., Massie, S., Mikhailenko, S., Miller, C., Moazzen-Ahmadi, N., Naumenko, O., Nikitin, A., Orphal, J., Perevalov, V., Perrin, A., Predoi-Cross, A., Rinsland, C., Rotger, M., Šimečková, M., Smith, M., Sung, K., Tashkun, S., Tennyson, J., Toth, R., Vandaele, A., and Vander Auwera, J.: The HITRAN 2008 molecular spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 110, 533–572, https://doi.org/https://doi.org/10.1016/j.jqsrt.2009.02.013, hITRAN, 2009.

Schepers, D., Guerlet, S., Butz, A., Landgraf, J., Frankenberg, C., Hasekamp, O., Blavier, J.-F., Deutscher, N. M., Griffith, D. W. T., Hase, F., Kyro, E., Morino, I., Sherlock, V., Sussmann, R., and Aben, I.: Methane retrievals from Greenhouse Gases Observing Satellite (GOSAT) shortwave infrared measurements: Performance comparison of proxy and physics retrieval algorithms, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/https://doi.org/10.1029/2012JD017549, 2012.

Shi, C., Hashimoto, M., Shiomi, K., and Nakajima, T.: Development of an Algorithm to Retrieve Aerosol Optical Properties Over Water Using an Artificial Neural Network Radiative Transfer Scheme: First Result From GOSAT-2/CAI-2, IEEE Transactions on Geoscience and Remote Sensing, 59, 9861–9872, https://doi.org/10.1109/TGRS.2020.3038892, 2021.

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.

Strong, K., Roche, S., Franklin, J. E., Mendonca, J., Lutsch, E., Weaver, D., Fogal, P. F., Drummond, J. R., Batchelor, R., Lindenmaier, R., and et al.: TCCON data from Eureka (CA), Release GGG2020.R0, https://doi.org/10.14291/tccon.ggg2020.eureka01.R0, funding by Atlantic Innovation Fund, 2022.

Sussmann, R. and Rettinger, M.: TCCON data from Garmisch (DE), Release GGG2020.R0, https://doi.org/10.14291/tccon.ggg2020.garmisch01.R0, 2023.

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.

Tadić, J. M., Loewenstein, M., Frankenberg, C., Iraci, L. T., Yates, E. L., Gore, W., and Kuze, A.: A comparison of in-situ aircraft measurements of carbon dioxide to GOSAT data measured over Railroad Valley playa, Nevada, USA, Atmospheric Measurement Techniques Discussions, 5, 5641–5664, https://doi.org/10.5194/amtd-5-5641-2012, 2012.

Tans, P. and Keeling, R.: Trends in Atmospheric Carbon Dioxide. Global Monitoring Laboratory, National Oceanic & Atmospheric Administration Earth System Research Laboratories, (NOAA/ESRL), 2020.

Taylor, T. E., O'Dell, C. W., Frankenberg, C., Partain, P. T., Cronk, H. Q., Savtchenko, A., Nelson, R. R., Rosenthal, E. J., Chang, A. Y., Fisher, B., Osterman, G. B., Pollock, R. H., Crisp, D., Eldering, A., and Gunson, M. R.: Orbiting Carbon Observatory-2 (OCO-2) cloud screening algorithms: validation against collocated MODIS and CALIOP data, Atmospheric Measurement Techniques, 9, 973–989, https://doi.org/10.5194/amt-9-973-2016, 2016.

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, D. F., Rettinger, M., Schneider, M., Roehl, C. M., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Vrekoussis,

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


M., Warneke, T., and Wunch, D.: An 11-year record of $XCO_2$ 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.

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.

Tebaldi, C. and Friedlingstein, P.: Delayed detection of climate mitigation benefits due to climate inertia and variability., Proceedings of the National Academy of Sciences, 110(43), 2013.

Tikhonov, A. N.: Solution of incorrectly formulated problems and the regularization method., Soviet Math., 4, 1035–1038, 1963.

Toon, G., Blavier, J.-F., Washenfelder, R., Wunch, D., Keppel-Aleks, G., Wennberg, P., Connor, B., Sherlock, V., Griffith, D., Deutscher, N., and Notholt, J.: Total Column Carbon Observing Network (TCCON), in: Advances in Imaging, p. JMA3, Optica Publishing Group, https://doi.org/10.1364/FTS.2009.JMA3, 2009.

Turner, A. J., Jacob, D. J., Wecht, K. J., Maasakkers, J. D., Lundgren, E., Andrews, A. E., Biraud, S. C., Boesch, H., Bowman, K. W., Deutscher, N. M., Dubey, M. K., Griffith, D. W. T., Hase, F., Kuze, A., Notholt, J., Ohyama, H., Parker, R., Payne, V. H., Sussmann, R., Sweeney, C., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D.: Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data, Atmospheric Chemistry and Physics, 15, 7049–7069, https://doi.org/10.5194/acp-15-7049-2015, 2015.

Uno, T., Oishi, T., Fujii, Y., Imaki, M., Sato, A., Kataoka, F., Imai, H., Hashimoto, M., Shiomi, K., and Nakajima, M.: The development and on-orbit calibration status of GOSAT-2 TANSO-CAI-2 instrument, in: International Conference on Space Optics — ICSO 2020, edited by Cugny, B., Sodnik, Z., and Karafolas, N., vol. 11852, p. 1185257, International Society for Optics and Photonics, SPIE, https://doi.org/10.1117/12.2599936, 2021.

Warneke, T., Petri, C., Notholt, J., and Buschmann, M.: TCCON data from Orléans (FR), Release GGG2020.R0, https://doi.org/10.14291/tccon.ggg2020.orleans01.R0, 2022.

Washenfelder, R. A., Toon, G. C., Blavier, J.-F., Yang, Z., Allen, N. T., Wennberg, P. O., Vay, S. A., Matross, D. M., and Daube., B. C.: Carbon Dioxide Column Abundances at the Wisconsin Tall Tower Site., Journal of Geophysical Research, D 111 (D22), 2006.

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, funding by Science and Technology Facilities Council GRID grid.14467.30., 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, funding by NASA, 2022a.

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, funding by NASA, 2022b.

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.R0, https://doi.org/10.14291/tccon.ggg2020.lamont01.R0, funding by NASA, 2022c.

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., 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.

https://doi.org/10.5194/egusphere-2024-3990
Preprint. Discussion started: 24 April 2025


Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P., Biraud, S. C., Castano, R., Cressie, N., Crisp, D., Deutscher, N. M., Eldering, A., Fisher, M. L., Griffith, D. W. T., Gunson, M., Heikkinen, P., Keppel-Aleks, G., Kyrö, E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C. E., Morino, I., Notholt, J., Oyafuso, F. A., Rettinger, M., Robinson, J., Roehl, C. M., Salawitch, R. J., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D. R., Uchino, O., Warneke, T., and Wofsy, S. C.: A method for evaluating bias in global measurements of $CO_2$ total columns from space, Atmospheric Chemistry and Physics, 11, 12 317–12 337, https://doi.org/10.5194/acp-11-12317-2011, 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 (GGG2014.R0)., CaltechDATA., 2015.

Wunch, D., Mendonca, J., Colebatch, O., Allen, N. T., Blavier, J.-F., Kunz, K., Roche, S., Hedelius, J., Neufeld, G., Springett, S., and et al.: TCCON data from East Trout Lake, SK (CA), Release GGG2020.R0, https://doi.org/10.14291/tccon.ggg2020.easttroutlake01.R0, funding by Canada Foundation for Innovation GRID grid.439998.6., 2022.

Yoshida, Y., Kikuchi, N., and Yokota, T.: On-orbit radiometric calibration of SWIR bands of TANSO-FTS onboard GOSAT, Atmospheric Measurement Techniques, 5, 2515–2523, https://doi.org/10.5194/amt-5-2515-2012, 2012.

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, advpub, 2023–023, https://doi.org/10.2151/sola.2023-023, 2023.

Zhang, Y., Jacob, D. J., Lu, X., Maasakkers, J. D., Scarpelli, T. R., Sheng, J.-X., Shen, L., Qu, Z., Sulprizio, M. P., Chang, J., Bloom, A. A., Ma, S., Worden, J., Parker, R. J., and Boesch, H.: Attribution of the accelerating increase in atmospheric methane during 2010–2018 by inverse analysis of GOSAT observations, Atmospheric Chemistry and Physics, 21, 3643–3666, https://doi.org/10.5194/acp-21-3643-2021, 2021.

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.