ECMWF COPERNICUS REPORT

Copernicus Climate Change Service

Product User Guide and Specification (PUGS) – ANNEX A for products CO2_GOS_OCFP, CH4_GOS_OCFP & CH4_GOS_OCPR (v7.2, 2009-2018)

C3S_312b_Lot2_IUP-UB – Atmosphere

Issued by: Hartmut Boesch, Jasdeep Anand, and Antonio Di Noia, University of Leicester, Leicester, UK

Date: 03/11/2019

Ref: C3S_D312b_Lot2.3.2.3-v1.0_PUGS-GHG_ANNEX-A_v3.1

Official reference number service contract: 2018/C3S_312b_Lot2_DLR/SC1

This document has been produced in the context of the Copernicus Climate Change Service (C3S).
The activities leading to these results have been contracted by the European Centre for Medium-Range Weather Forecasts, operator of C3S on behalf of the European Union (Delegation Agreement signed on 11/11/2014). All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission and the European Centre for Medium-Range Weather Forecasts has no liability in respect of this document, which is merely representing the authors view.

Contributors

INSTITUTE OF ENVIRONMENTAL PHYSICS (IUP), UNIVERSITY OF BREMEN, BREMEN, GERMANY
(IUP)
M. Buchwitz

UNIVERSITY OF LEICESTER, LEICESTER, UK
(UoL)
H. Boesch
J. Anand
P. Somkuti
R. Parker
A. Di Noia

Table of Contents

History of modifications

Version	Date	Description of modification	Chapters / Sections
1.3	20-October-2017	New document for data set CDR1 (2009-2016)	All
2.0	4-October-2018	Updated product description, filtering criteria, and bias correction information for v7.2 Update for CDR2 (2009-2017)	All
3.0	12-August-2019	Update for CDR3 (2009-2018)	All
3.1	03-November-2019	Update after review by Assimila: Only date and version number changed (for consistency reasons) as otherwise OK as is.	First few pages

Related documents

Reference ID	Document
D1	Main PUGS: Buchwitz, M., et al., Product User Guide and Specification (PUGS) – Main document for Greenhouse Gas (GHG: CO2 & CH4) data set CDR 3 (2003-2018), project C3S_312b_Lot2_DLR – Atmosphere, 3.1, 2019. *(this document is an ANNEX to the Main PUGS)*

Acronyms

Acronym	Definition
CAR	Climate Assessment Report
C3S	Copernicus Climate Change Service
CCDAS	Carbon Cycle Data Assimilation System
ECMWF	European Centre for Medium Range Weather Forecasting
ECV	Essential Climate Variable
EU	European Union
FP	Full Physics retrieval method
FTS	Fourier Transform Spectrometer
GHG	GreenHouse Gas
GMES	Global Monitoring for Environment and Security
GOSAT	Greenhouse Gases Observing Satellite
IUP	Institute of Environmental Physics (IUP) of the University of Bremen, Germany
JAXA	Japan Aerospace Exploration Agency
L1	Level 1
L2	Level 2
L3	Level 3
L4	Level 4
LMD	Laboratoire de Météorologie Dynamique
LMDZ	Laboratoire de Météorologie Dynamique Zoom (Global climate model)
MACC	Monitoring Atmospheric Composition and Climate, EU GMES project
NA	Not applicable
NetCDF	Network Common Data Format
NIR	Near Infra Red
NOAA	National Oceanic and Atmospheric Administration
OCO	Orbiting Carbon Observatory
OE	Optimal Estimation
PCA	Principal Component Analysis
ppb	Parts per billion
ppm	Parts per million
PR	(light path) PRoxy retrieval method
PQAR	Product Quality Assessment Report
SWIR	Short Wave Infra Red
TANSO	Thermal And Near infrared Sensor for carbon Observation
TANSO-FTS	Fourier Transform Spectrometer on GOSAT
TCCON	Total Carbon Column Observing Network
TIR	Thermal Infra Red

TR	Target Requirements
TRD	Target Requirements Document
UoL	University of Leicester, United Kingdom

General definitions

Table 1 lists some general definitions relevant for this document.

Table 1: General definitions.

Item	Definition
XCO_2	Column-averaged dry-air mixing ratios (mole fractions) of CO_2
XCH_4	Column-averaged dry-air mixing ratios (mole fractions) of CH_4
L1	Level 1 satellite data product: geolocated radiance (spectra)
L2	Level 2 satellite-derived data product: Here: CO_2 and CH_4 information for each ground-pixel
L3	Level 3 satellite-derived data product: Here: Gridded CO_2 and CH_4 information, e.g., 5 deg times 5 deg, monthly
L4	Level 4 satellite-derived data product: Here: Surface fluxes (emission and/or uptake) of CO_2 and CH_4

Scope of document

This document is a Product User Guide and Specification (PUGS) for the Copernicus Climate Change Service (C3S, https://climate.copernicus.eu/) greenhouse gas (GHG) component as covered by project C3S_312b_Lot2.

Within this project satellite-derived atmospheric carbon dioxide (CO_2) and methane (CH_4) Essential Climate Variable (ECV) data products will be generated and delivered to ECMWF for inclusion into the Copernicus Climate Data Store (CDS) from which users can access these data products and the corresponding documentation.

The satellite-derived GHG data products are:

- Column-averaged dry-air mixing ratios (mole fractions) of CO_2 and CH_4, denoted XCO_2 (in parts per million, ppm) and XCH_4 (in parts per billion, ppb), respectively.
- Mid/upper tropospheric mixing ratios of CO_2 (in ppm) and CH_4 (in ppb).

This document describes the C3S products CO2_GOS_OCFP (v 7.2), CH4_GOS_OCFP (v 7.2) and CH4_GOS_OCPR (v 7.2).

These products are XCO_2 and XCH_4 Level 2 products as retrieved from GOSAT using algorithms developed at the University of Leicester, UK.

Executive summary

This document summarises the data and metadata stored in the Copernicus Climate Change Service (C3S) Level 2 CO_2 and CH_4 data products developed by the University of Leicester (UoL). These products provide the column-averaged dry-air mixing ratios (mole fractions) of CO_2 (XCO_2) and (XCH_4), derived from short-wave infrared (SWIR) spectra from the JAXA Greenhouse gases Observing SATellite (GOSAT). The datasets discussed in this work cover the entire satellite operational period (2009-2018), and are published as daily netCDF files available from the C3S website: https://climate.copernicus.eu/.

This aim of this document is to clearly describe to users the quality flags and metadata, data format, product grid and geographical projection, known limitations, available tools for decoding and interpreting the data. In addition, this document also briefly discusses the validation of these datasets against the C3S target requirements stated in, *TRD GHG, 2017*, through comparisons with highly accurate ground-based measurements provided by the Total Carbon Column Observing Network (TCCON). Through these comparisons, we are confident that the datasets in this work at least meet the breakthrough requirements set in the TRD. Overall, the UoL datasets meet the stringent precision and accuracy requirements set by C3S, and are thought to offer information on regional surface fluxes of CO_2 and CH_4.

1. Product description

1.1 The GOSAT-FTS Instrument

The Japanese Greenhouse gases Observing SATellite (GOSAT) was launched on 23rd January 2009, *Yokota et al., 2009*, by JAXA, the Japanese Space Agency. GOSAT provides the first dedicated global measurements of total column CO_2 and CH_4 from its SWIR bands, *Yoshida et al., 2013*. It is equipped with two instruments; the Thermal And Near-infrared Sensor for carbon Observations - Fourier Transform Spectrometer (TANSO-FTS), and a dedicated Cloud and Aerosol Imager (TANSO-CAI).

TANSO-FTS measures in four spectral bands with a high spectral resolution of 0.2 cm^{-1} , three of which operate in the SWIR at around 0.76, 1.6 and 2.0 µm providing sensitivity to the near-surface absorbers with the fourth channel operating in the thermal infrared between 5.5 and 14.3 µm providing mid-tropospheric sensitivity, *Saitoh et al., 2009*.

The measurement strategy of TANSO-FTS is optimised for the characterisation of continental-scale sources and sinks, with the aim of achieving a 0.3—1% relative accuracy for 3-month averages of CO_2 at a 100-1000 km spatial resolution, *Kuze et al., 2009*. The aim for CO_2 is to achieve an accuracy of better than 2% on the same spatial and temporal scales. In order to achieve this, TANSO-FTS utilises a pointing mirror to perform off-nadir measurements at the same location on each 3-day repeat cycle. The pointing mirror allows TANSO-FTS to observe up to ±35° across track and ±20° along-track. These measurements nominally consist of 5 across track points spaced ~100 km apart (although measurements are possible with 1, 3, 5, 7 or 9 across track points) with a ground footprint diameter of approximately 10.5 km and a 4 second exposure duration. Whilst the majority of data is limited to measurements over land where surface reflectance is high, TANSO-FTS also observes in sun-glint mode over the ocean within ±20° of the subsolar latitude.

1.2 The University of Leicester Products

The UoL have retrieved several datasets from GOSAT TANSO-FTS NIR and SWIR spectra, which are discussed in this section:

XCO_2:

- CO2_GOS_OCFP (v 7.2)

XCH_4:

- CH4_GOS_OCFP (v 7.2)
- CH4_GOS_OCPR (v 7.2)

All products mentioned in this document are retrieved using the University of Leicester (UoL) Full-Physics Retrieval Algorithm, based on the original Orbiting Carbon Observatory (OCO) Full Physics retrieval algorithm, modified for use with GOSAT spectra (OCFP).

The retrieval algorithm uses an iterative retrieval scheme based on Bayesian optimal estimation to retrieve a set of atmospheric, surface and instrument parameters, referred to as the state vector, from measured spectral radiances, *Boesch et al., 2011*; *Connor et al., 2008*. The forward model, used to relate the state vector to the measured radiances, includes the LIDORT, *Spurr, 2008*, and TWOSTR, *Spurr et al., 2011*, radiative transfer models combined with a fast 2 orders of scattering vector radiative transfer code *Natraj et al., 2008*. In order to accelerate the radiative transfer component of the retrieval algorithm, the code uses the low stream interpolation (LSI) method described in *O'Dell, 2010,* or the principal component analysis (PCA)-based fast RT scheme described in *Somkuti et al.,* 2017.

In addition to the Full-Physics retrieval products, we also offer a separate product for CH_4, which is retrieved using the Full-Physics algorithm modified by the "proxy" technique (OCPR) as discussed in, *Parker et al., 2011* and, *Parker et al., 2015*. CO_2 is known to vary in the atmosphere much less than CH_4 and as the CO_2 absorption band is spectrally close to that of CH_4 we can use the CO_2 as a proxy for the light path to minimize common spectral artefacts due to aerosol scattering and instrumental effect. CH_4 and CO_2 retrievals are carried out sequentially with channels at 1.65 µm and 1.61 µm respectively.

In order to obtain a volume mixing ratio (VMR) of CH_4, it is necessary to multiply the retrieved XCH_4/XCO_2 ratio by a model XCO_2. We obtain the CO_2 VMRs from the median of a model CO_2 ensemble that comprises GEOS-Chem (University of Edinburgh), LMDZ/MACC-II and NOAA CarbonTracker, convolved with scene-dependent instrument averaging kernels obtained from the GOSAT 1.6 µm CO_2 retrieval.

Figures 1-3 show the global seasonal variation of XCO_2 and XCH_4 over all three data products between April 2009 and December 2017.

Figure 1: Global seasonal maps of UoL GOSAT XCO_2 (CO2_GOS_OCFP) retrieved between April 2009 and December 2017.

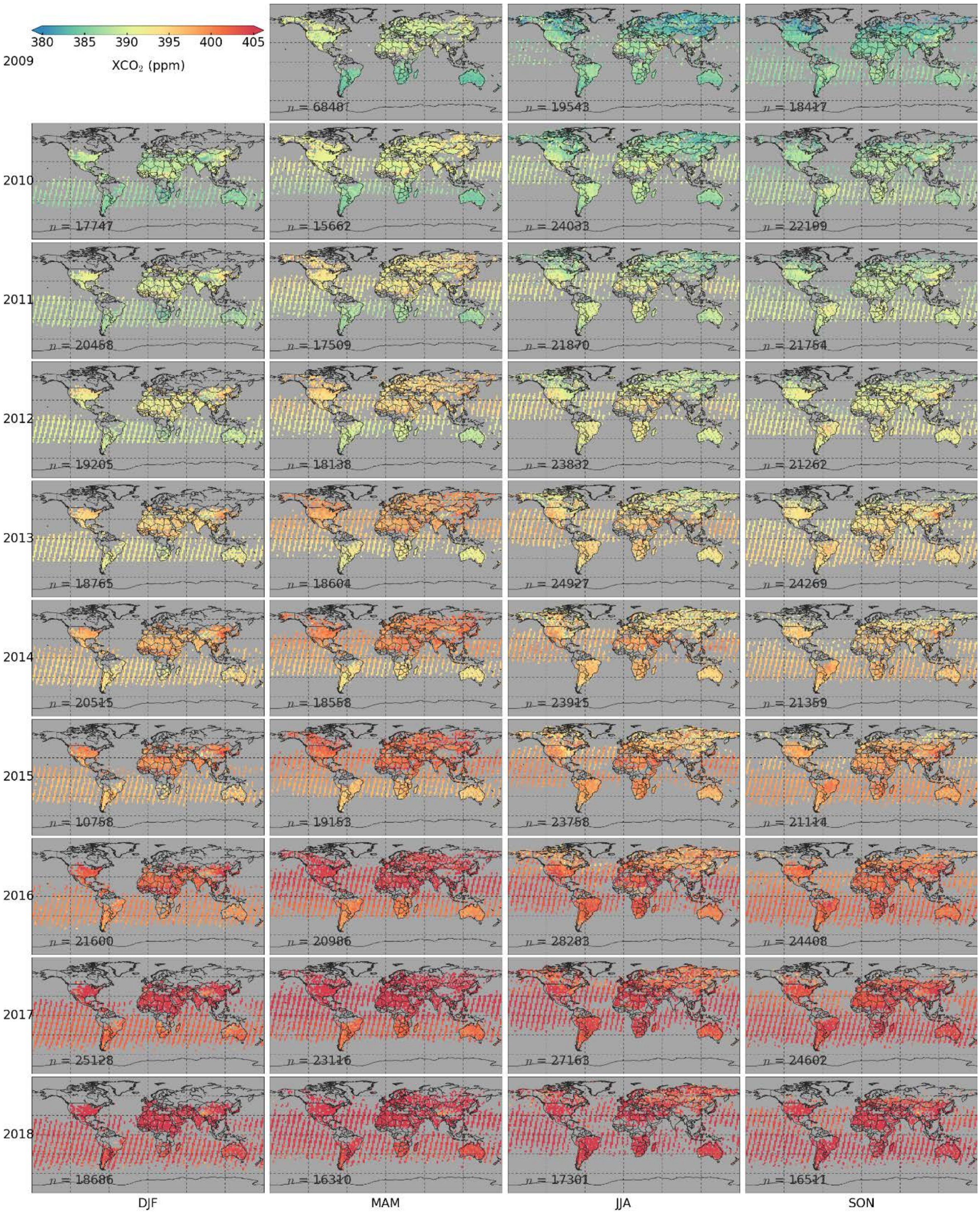


Figure 2: Global seasonal maps of UoL GOSAT XCH_4 (CH4_GOS_OCFP) retrieved between April 2009 and December 2017.

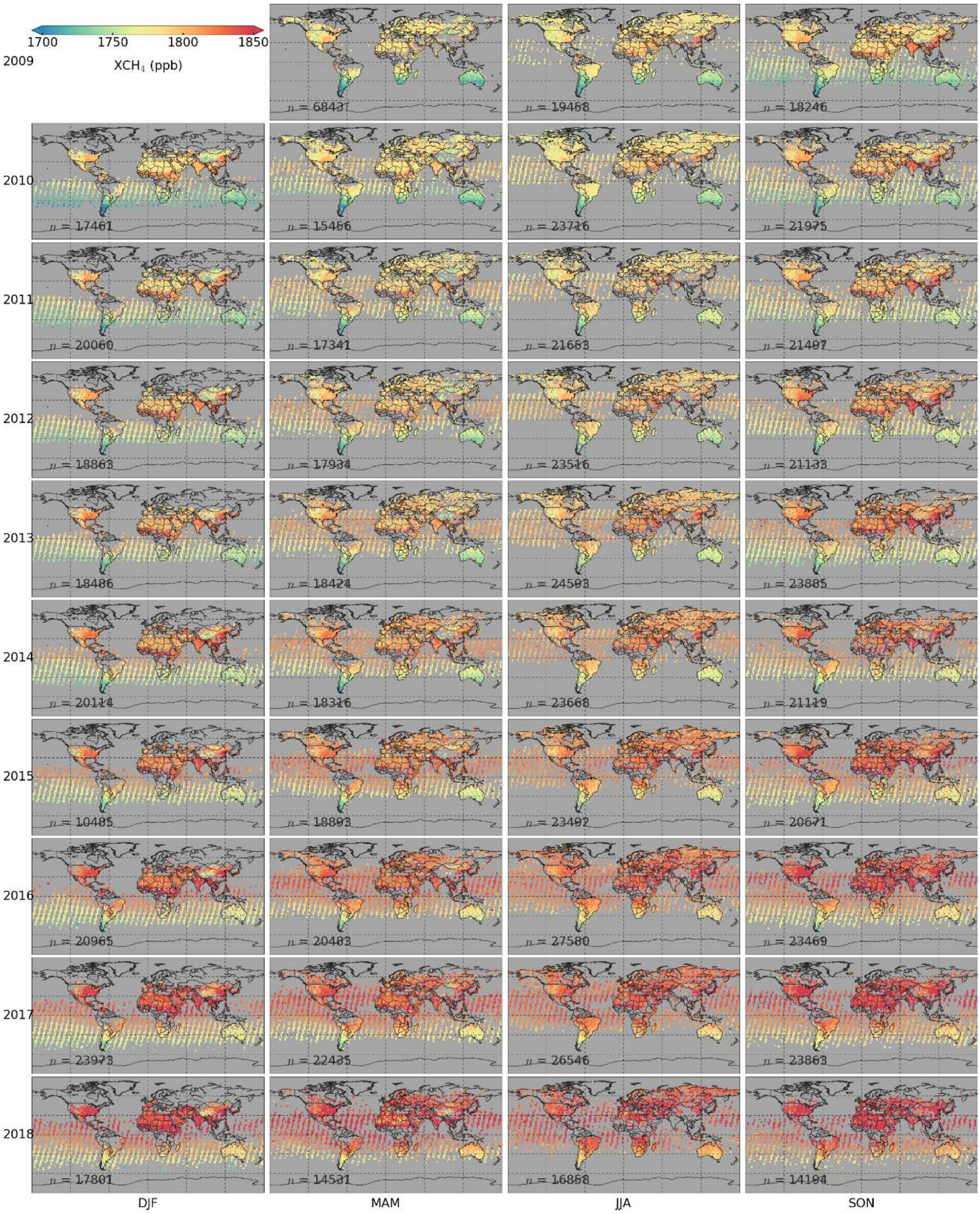
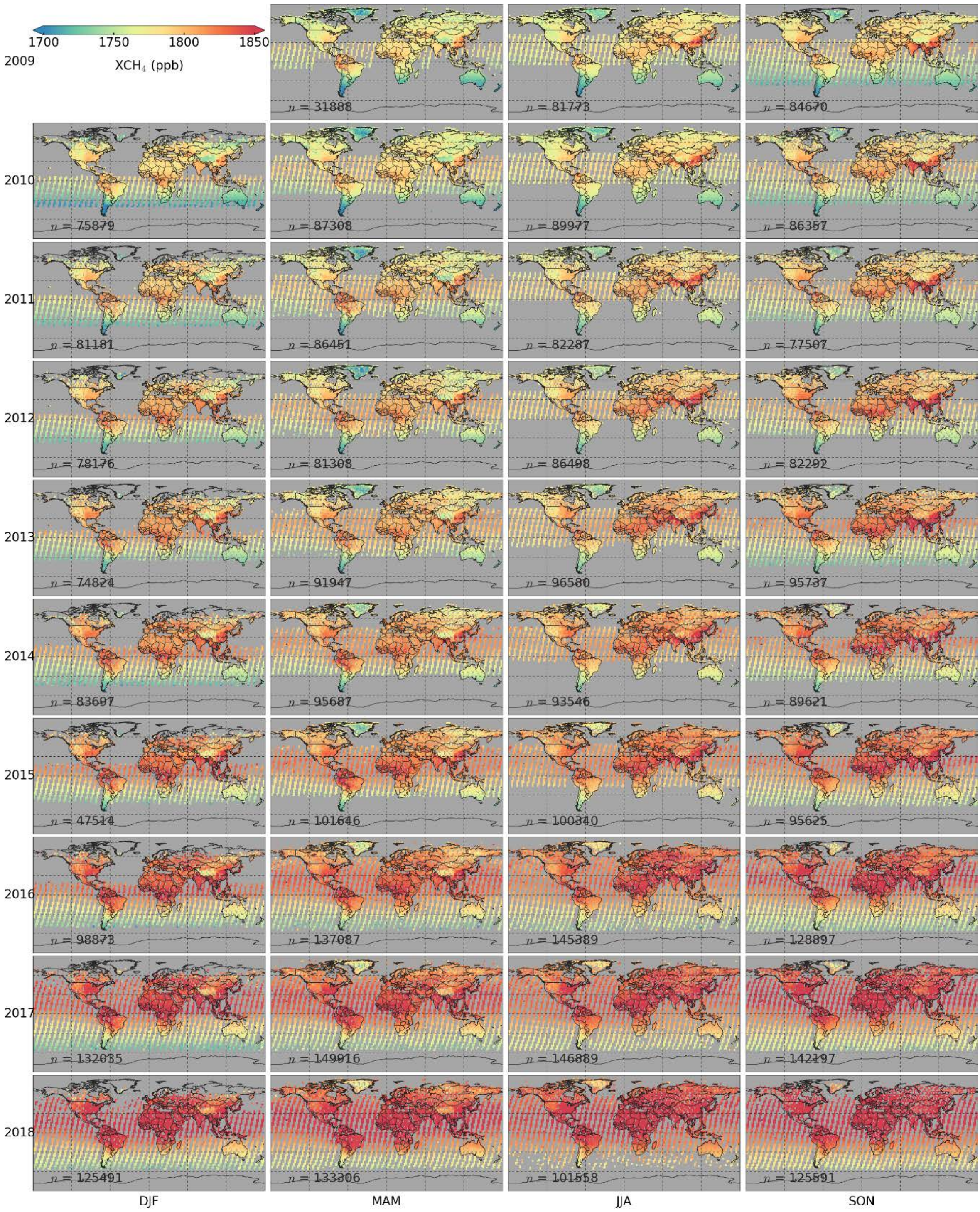


Figure 3: Global seasonal maps of UoL GOSAT XCH_4 (CH4_GOS_OCPR) retrieved between April 2009 and December 2017.



1.3 Post-retrieval processing

1.3.1 Filtering

To ensure data quality, the GOSAT data is filtered for anomalously high or low retrieval fit statistics, along with anomalous values in its geophysical or final state vector parameters. The filtering criteria were empirically determined through analysis of the fit statistics, along with comparisons made with co-located ground-based measurements from the Total Carbon Column Observing Network (TCCON, See Section 2).

1.3.1.1 Pre-retrieval screening

Before a retrieval is performed the GOSAT soundings are subjected to several tests for measurement noise and other issues. For CO2_GOS_OCFP and CH4_GOS_OCFP, only soundings that pass the criteria shown in Table 2 are used in the retrieval. For the CH4_GOS_OCPR product only the cloud screening and geographic criteria shown in Table 2 are applied.

Table 2: The pre-retrieval filtering criteria used in the CO2_GOS_OCFP and CH4_GOS_OCFP products.

Parameter	Filtering criteria
SNR (all bands)	≥ 20
SZA	≤ 75°
Latitude	≥ 60° S
Δ(Surface pressure): difference between retrieved and a priori value (cloud screen)	≤ 30 hPa
Weak/strong CO_2 column ratio	≥ 0.98, ≤ 1.05

1.3.1.2 Post-retrieval screening

After the retrieval, the datasets are subsequently screened to determine if the retrieval was successful. Data retrieved from glint and land measurements are filtered separately, as viewing conditions are markedly different over oceans. The post-filtering criteria used in the CO2_GOS_OCFP are shown in Table 3. For CH4_GOS_OCFP only soundings which had previously passed the CO2_GOS_OCFP filtering were considered useful. The CH_4 retrievals from these soundings were then subsequently filtered again using the criteria shown in Table 4 before being flagged as good data.
It is important to point out that retrievals for 2018 use a new version of GOSAT L1B data (V210.210 instead of V201.202, Kitaoka et al., 2019). As we estimate the noise from the GOSAT L1B spectra itself (see ATBD), this resulted in a different estimate of the noise and subsequently we found the χ^2 values of our full physics retrievals to be generally larger than those observed up to 2017. As a consequence, the χ^2 filter used in the post-processing of retrievals from 2018 had to be adjusted compared to that used up to 2017. Exceptions are the band 3 filter for CO_2 land retrievals and the band 2 filter for CH_4 land retrievals, which did not require any update.

Table 3: The post-retrieval filtering criteria used in the CO2_GOS_OCFP product.

Parameter	Filtering criteria	
	Land	Glint
Retrieval outcome	Converged	Converged
SNR (all bands)	≥ 45	≥ 45
n retrieval iterations	≤ 7	≤ 7
SZA	≤ 65°	NA
n diverging retrieval steps	≤ 2	≤ 2
χ^2 (Band 1)	≥ 0.5, ≤ 1.55 (2018: ≥ 1.5, ≤ 2.55)	≥ 0.9, ≤ 1.45 (2018: ≥ 1.7, ≤ 2.95)
χ^2 (Band 2)	≥ 0.6, ≤ 2.0 (2018: ≥ 0.6, ≤ 2.0)	≥ 0.8, ≤ 1.70 (2018: ≥ 1.3, ≤ 2.3)
χ^2 (Band 3)	≥ 0.5, ≤ 1.55	≥ 0.65, ≤ 1.25 (2018: ≥ 0.8, ≤ 1.7)
Weak/strong CO_2 column ratio	≥ 0.99, ≤ 1.01	≥ 0.99, ≤ 1.01
XCO_2 a posteriori error	≤ 2.5 ppm	≤ 1.15 ppm
Total AOD (cirrus + small + large aerosols)	≤ 0.5	≤ 0.17
AOD (small aerosol)	≤ 0.3	≤ 0.3
AOD (large aerosol)	≤ 0.15	≤ 0.08
ΔAOD (large aerosol): difference between retrieved and a priori value	≥ -1.8	≥ -1.25
ΔAOD (cirrus): difference between retrieved and a priori value	≥ -6.25	≥ -7.0
σ surface pressure within ground pixel	≤ 20 hPa	NA
Δ(Surface pressure): difference between retrieved and a priori value (cloud screen)	NA	≥ -3.32, ≤ 1.0
Albedo slope (Band 1)	≤ 2.5×10^{-5}	≥ 2.6×10^{-6}, ≤ 1.75×10^{-5}
Albedo slope (Band 2)	NA	≥ 0.0, ≤ 5.0×10^{-6}
Albedo slope (Band 3)	≥ -2.0×10^{-4}	≥ 0.0, ≤ 2.5×10^{-5}
Albedo ratio between band 1 and band 2	≤ 2.75	≥ 0.98, ≤ 1.2
Albedo ratio between band 1 and band 3	NA	≥ 1.09, ≤ 1.2

Retrieved CO_2 profile gradient between the surface and retrieval level 15	NA	≥ 0.9, ≤ 1.01

The soundings which pass the filtering criteria in the CO2_GOS_OCFP product are subsequently filtered again using the criteria in Table 4 to provide filtered data for CH4_GOS_OCFP.

Table 4: The post-retrieval filtering criteria used in the CH4_GOS_OCFP product. Note that soundings are first filtered using the XCO_2 retrieval parameters shown in Table 3. Each parameter indicates whether the considered value is taken from the XCO_2 or XCH_4 retrieval.

Parameter	Filtering criteria	
	Land	Glint
Retrieval outcome (XCH_4)	Converged	Converged
SNR (all bands)	≥ 45	≥ 45
n retrieval iterations (XCH_4)	≤ 7	≤ 7
n diverging retrieval steps (XCH_4)	≤ 2	≤ 2
χ^2 (Band 2, XCH_4)	≤ 3.6	≥ 0.86, ≤ 1.44 (2018: ≥ 1.05, ≤ 1.63)
Albedo slope (Band 2, XCH_4)	≥ -5.5×10^{-5}	NA
XCH_4 a posteriori error	≤ 9.5 ppb	NA
σ surface pressure within ground pixel (XCO_2)	NA	≤ 2.38 hPa

Table 5: The post-retrieval filtering criteria used in the CH4_GOS_OCPR product.

Parameter	Filtering criteria
χ^2 (XCH_4 retrieval)	≥ 0.4, ≤ 1.9 (2018: ≥ 0.6, ≤ 2.1)
χ^2 (XCO_2 retrieval)	≥ 0.4, ≤ 1.9 (2018: ≥ 0.6, ≤ 2.1)
XCH_4 a posteriori error	≤ 20 ppb
XCO_2 a posteriori error	≤ 3 ppm
Retrieved XCH_4	≥ 1650 ppb
Retrieved XCO_2	≥ 350 ppm

1.3.2 Bias correction

For these data products, a bias correction based on several state vector parameters is calculated via a regression analysis of the difference between collocated GOSAT and TCCON XCH_4 and XCO_2 observations. Land and glint measurements were corrected separately for each product.

For CO2_GOS_OCFP and CH4_GOS_OCFP, the correction takes the form of a linear equation of *n* state vector parameters (***x***) multiplied by a unique coefficient (*m*) along with a single offset (*c*), such that:

$$\text{correction} = c + m_0 x_0 + m_1 x_1 + \ldots + m_{n-1} x_{n-1}$$

The correction is then subtracted from the original XCO_2 or XCH_4 to give the final value:

$$\text{XCO}_{2_{\text{final}}} = \text{XCO}_2 - \text{correction}$$

The regression analysis makes use of the RANSAC method to avoid statistical outliers affecting the fit. As such, the total mean bias against TCCON remaining in the data after this correction is not zero (see Section 2).

Tables 6-9 show the values of *m* and *c* used to correct the land and glint data in the CO2_GOS_OCFP and CH4_GOS_OCFP products.

Table 6: The parameters and coefficient values used in the bias correction for the CO2_GOS_OCFP product (land soundings only). An offset of: -17.96 ppm is also applied.

Parameter	Coefficient
Retrieved CO_2 profile gradient between the surface and retrieval level 15	19.75
AOD (large aerosol)	-25.57

Table 7: The parameters and coefficient values used in the bias correction for the CO2_GOS_OCFP product (glint soundings only). An offset of: -23.44 ppm is also applied.

Parameter	Coefficient
Albedo slope (Band 3)	1.39×10^5
Retrieved CO_2 profile gradient between the surface and retrieval level 15	24.80
Total AOD (cirrus + small + large aerosols)	-13.20
Albedo slope (Band 1)	-4.01×10^4

Table 8: The parameters and coefficient values used in the bias correction for the CH4_GOS_OCFP product (land soundings only). An offset of: -0.125 ppb is also applied. Each parameter indicates whether the considered value is taken from the XCO_2 or XCH_4 retrieval.

Parameter	Coefficient
Albedo ratio between band 1 and band 3 (XCO_2)	4.58
Total AOD (cirrus + small + large aerosols, XCO_2)	-68.35
Albedo slope (Band 2, XCO_2)	1.25×10^5

Table 9: The parameters and coefficient values used in the bias correction for the CH4_GOS_OCFP product (glint soundings only). An offset of: 10.79 ppb is also applied. Each parameter indicates whether the considered value is taken from the XCO_2 or XCH_4 retrieval.

Parameter	Coefficient
Total AOD (cirrus + small + large aerosols, XCO_2)	-89.77
Albedo slope (Band 1, XCO_2)	-1.44×10^6

For CH4_GOS_OCPR a simple global bias correction of -7.71 ppb is applied to all data to remove the mean bias to TCCON.

2. Target requirements

Products submitted to C3S must fulfill a number of stringent quality requirements, which are further discussed in the Target Requirements Document; *TRD GHG, 2017*. A full summary of these requirements, and how far our products fulfil them, is available in the PQAR Document. In this section we briefly summarise the requirements for random and systematic errors, and validate our products using TCCON data. Table 10 shows the random and systematic errors stated in the TRD.

Table 10: XCO_2 and XCH_4 random ("precision") and systematic retrieval error requirements for measurements over land. Abbreviations: G=Goal, B=Breakthrough, T=Threshold requirement. §) Required systematic error after an empirical bias correction, that does not use the verification data. #) Required systematic error and stability after bias correction, where bias correction is not limited to the application of a constant offset / scaling factor.

Random and systematic error requirements for XCO_2 and XCH_4					
Parameter	Req. type	Random error ("Precision")		Systematic error	Stability
		Single obs.	1000^2 km^2 monthly		
XCO_2	G	< 1 ppm	< 0.3 ppm	< 0.2 ppm (absolute)	As systematic error but per year
	B	< 3 ppm	< 1.0 ppm	< 0.3 ppm (relative §))	-"-
	T	< 8 ppm	< 1.3 ppm	< 0.5 ppm (relative #))	-"-
XCH_4	G	< 9 ppb	< 3 ppb	< 1 ppb (absolute)	< 1 ppb/year (absolute)
	B	< 17 ppb	< 5 ppb	< 5 ppb (relative §))	< 2 ppb/year (relative §))
	T	< 34 ppb	< 11 ppb	< 10 ppb (relative #))	< 3 ppb/year (relative #))

For both full-physics products, we have considered the land and glint measurements separately (see PQAR Document). Table 11 shows a summary of the statistics generated from direct comparisons between GOSAT and TCCON. The mean GOSAT-TCCON bias is a representation of the true systematic error, while the standard deviation is a representation of the true random error. Therefore, all datasets achieve at least the breakthrough requirements for XCO_2 and XCH_4 stated in Table 10.

Table 11: The results of direct comparisons between the UoL products and TCCON for GOSAT soundings between April 2009 and December 2017.

Dataset	Number of measurements	Pearson coefficient (r)	Mean bias	Standard deviation
XCO_2 (land)	21965	0.96	0.00 ppm	1.85 ppm
XCO_2 (glint)	1255	0.97	0.24 ppm	1.19 ppm
XCH_4 (OCFP, land)	21863	0.91	-0.31 ppb	13.94 ppb
XCH_4 (OCFP, glint)	1189	0.92	0.40 ppb	10.05 ppb
XCH_4 (OCPR, land)	64454	0.92	-1.25 ppb	13.81 ppb
XCH_4 (OCPR, glint)	8060	0.92	3.80 ppb	12.35 ppb

3. Data usage information

For all data products, **the xco2_quality_flag or xch4_quality_flag variable must be applied to the data before use; a value of 0 indicates that the data has passed our quality control**. All vertically resolved data is provided on levels (as opposed to layers). This is especially important when applying UoL averaging kernels to model data.

For the CO2_GOS_OCFP and CH4_GOS_OCFP products, most users will be interested in the **xch4** or **xco2** variables, which store the column-averaged dry-air mixing ratios of the required gas. We also provide the values of the mixing ratios before any bias correction is applied, which are stored in the **xco2_no_bias_correction/xch4_no_bias_correction** variable.

For CH4_GOS_OCPR, the final proxy data product is stored in the **xch4** variable. It is recommended that users use this variable unless explicitly interested in the retrieved XCH_4/XCO_2 ratio. Users interested in the raw XCH_4 and XCO_2 retrieved from the 1.6 µm band uncorrected for aerosol scattering can find these values stored in the **raw_xch4** and **raw_xco2** variables.

We also include other important variables, such as averaging kernels, errors, and geolocation data in the netCDF files. Please see Section 3.3 for the full data file content.

3.1 Tools for reading the data

The datasets are stored in netCDF format, which can be read with standard tools in common programming languages.

3.2 Known limitations and issues

Users must be aware of the following caveats when using these datasets:

- As discussed in Section 1.3.2 we apply a bias correction to the data based on linear regression of geophysical parameters against the observed GOSAT-TCCON bias.
- A preliminary comparison of our XCO_2 and XCH_4 a posteriori errors against the standard deviation of the GOSAT-TCCON differences has indicated that our error estimates are potentially too small. For the xco2_uncertainty reported in the CO2_GOS_OCFP data product, we have multiplied the a posteriori error by a factor of 1.75 for land observations and 1.17 for glint observations so that it is a more realistic value. Similarly, the xch4_uncertainty reported in the CH4_GOS_OCFP product has been multiplied by a factor of 1.73 for land observations and 1.23 for glint observations. Further exploration of this will be performed as part of the validation exercises.
- For the CH4_GOS_OCPR product, more information about the models used to estimate the true XCO_2 column can be found in Section 1.2. If you wish to renormalize the XCH_4/XCO_2 ratio with your own model XCO_2 data, please be aware that you should first apply the provided averaging kernels to your model data.

3.3 Data file content

netCDF data files contain all of the common parameters for the C3S data products, as well as additional product-specific parameters. A dimension of *n* refers to the number of retrievals per file, whilst a dimension of *m* refers to the number of levels retrieved for each sounding. For CO2_GOS_OCFP, CH4_GOS_OCFP and most of CH4_GOS_OCPR *m* is always 20.

However, for some soundings made over high terrain in CH4_GOS_OCPR the lowest level is removed to ensure that the remaining levels are above the surface. In this case, values in the 20th level are replaced with the fill value: -9999.99. Users reading averaging kernel and pressure level information will always see 20 levels, but will need to check whether the lowest level has been replaced with the fill value.

Table 12: Variables present in the CO2_GOS_OCFP product.

Name	Type	Dimensions	Units	Description
solar_zenith_angle	float	n	degree	Angle between line of sight to the sun and local vertical
sensor_zenith_angle	float	n	degree	Angle between the line of sight to the sensor and the local vertical
time	double	n	seconds since 1970-01-01 00:00:00	Measurement time
longitude	float	n	degrees_east	Centre longitude
latitude	float	n	degrees_north	Centre latitude
pressure_levels	float	n, m	hPa	Vertical altitude coordinate in pressure units as used for averaging kernels
pressure_weight	float	n, m		Pressure weights as used for averaging kernels
xco2	float	n	1e-6	Retrieved column-averaged dry-air mole fraction of atmospheric carbon dioxide (XCO_2) in ppm.
xco2_no_bias_correction	float	n	1e-6	Retrieved column-averaged dry-air mole fraction of atmospheric carbon dioxide (XCO_2) in ppm. No bias correction is applied
xco2_uncertainty	float	n	1e-6	Statistical uncertainty of XCO_2 in ppm (1σ)
xco2_averaging_kernel	float	n, m		XCO_2 averaging kernel (a profile = vector for each single observation). Quantifies the

				altitude sensitivity of the XCO_2 retrieval
co2_profile_apriori	float	n, m	1e-6	A-priori mole fraction profile of atmospheric CO_2 in ppm
exposure_id	char	n, 22		Exposure identification number of the sounding
surface_altitude	float	n	metres	Altitude is the (geometric) height above the geoid, which is the reference geopotential surface
surface_altitude_stdev	float	n	metres	Standard deviation of the surface elevation within the area of the GOSAT sounding, as derived from the SRTM database
surface_air_pressure_apriori	float	n	hPa	A-priori surface pressure value
surface_air_pressure_apriori_std	float	n	hPa	A-priori surface pressure standard deviation
gain	byte	n		GOSAT TANSO-FTS instrument gain mode. 1 indicates high gain. 0 indicates medium gain
air_temperature_apriori	float	n, m	K	Air temperature is the bulk temperature of the air, not the surface (skin) temperature
h2o_profile_apriori	float	n, m	ppm	A-priori mole fraction profile of atmospheric H_2O in ppm
total_aod	float	n		Retrieved total aerosol optical depth
aod_type1	float	n		Retrieved AOD (small)
aod_type2	float	n		Retrieved AOD (large)
cirrus	float	n		Retrieved AOD (cirrus)
retr_flag	byte	n		Retrieval type flag (0 = land, 1 = glint)

Table 13: Variables present in the CH4_GOS_OCFP product

Name	Type	Dimensions	Units	Description
solar_zenith_angle	float	n	degree	Angle between line of sight to the sun and local vertical
sensor_zenith_angle	float	n	degree	Angle between the line of sight to the sensor and the local vertical
time	double	n	seconds since 1970-01-01 00:00:00	Measurement time
longitude	float	n	degrees_east	Centre longitude
latitude	float	n	degrees_north	Centre latitude
pressure_levels	float	n, m	hPa	Vertical altitude coordinate in pressure units as used for averaging kernels
pressure_weight	float	n, m		Pressure weights as used for averaging kernels
xch4	float	n	1e-9	Retrieved column-averaged dry-air mole fraction of atmospheric methane (XCH_4) in ppb
xch4_no_bias_correction	float	n	1e-9	Retrieved column-averaged dry-air mole fraction of atmospheric methane (XCH_4) in ppb. No bias correction is applied
xch4_uncertainty	float	n	1e-9	Statistical uncertainty of XCH_4 in ppb (1σ)
xch4_averaging_kernel	float	n, m		XCH_4 averaging kernel (a profile = vector for each single observation). Quantifies the altitude sensitivity of the XCH_4 retrieval
co2_profile_apriori	float	n, m	1e-6	A-priori mole fraction profile of atmospheric CO_2 in ppm
ch4_profile_apriori	float	n, m	1e-9	A-priori mole fraction profile of atmospheric CH_4 in ppb
exposure_id	char	n, 22		Exposure identification number of the sounding
surface_altitude	float	n	metres	Altitude is the (geometric) height above the geoid, which is the reference geopotential surface
surface_altitude_stdev	float	n	metres	Standard deviation of the surface elevation within the

				area of the GOSAT sounding, as derived from the SRTM database
surface_air_pressure_apriori	float	n	hPa	A-priori surface pressure value
surface_air_pressure_apriori_std	float	n	hPa	A-priori surface pressure standard deviation
gain	byte	n		GOSAT TANSO-FTS instrument gain mode. 1 indicates high gain. 0 indicates medium gain
air_temperature_apriori	float	n, m	K	Air temperature is the bulk temperature of the air, not the surface (skin) temperature
h2o_profile_apriori	float	n, m	ppm	A-priori mole fraction profile of atmospheric H_2O in ppm
total_aod	float	n		Retrieved total aerosol optical depth
aod_type1	float	n		Retrieved AOD (small)
aod_type2	float	n		Retrieved AOD (large)
cirrus	float	n		Retrieved AOD (cirrus)
retr_flag	byte	n		Retrieval type flag (0 = land, 1 = glint)

Table 14: Variables present in the CH4_GOS_OCPR product.

Name	Type	Dimensions	Units	Description
solar_zenith_angle	float	n	degree	Angle between line of sight to the sun and local vertical
sensor_zenith_angle	float	n	degree	Angle between the line of sight to the sensor and the local vertical
time	double	n	seconds since 1970-01-01 00:00:00	Measurement time
longitude	float	n	degrees_east	Centre longitude
latitude	float	n	degrees_north	Centre latitude
pressure_levels	float	n, m	hPa	Vertical altitude coordinate in pressure units as used for averaging kernels
pressure_weight	float	n, m		Pressure weights as used for averaging kernels
xch4	float	n	1e-9	Retrieved column-averaged dry-air mole fraction of atmospheric methane (XCH_4) in ppb.
xch4_uncertainty	float	n	1e-9	Statistical uncertainty of XCH_4 in ppb (1σ)
xch4_averaging_kernel	float	n, m		XCH_4 averaging kernel (a profile = vector for each single observation). Quantifies the altitude sensitivity of the XCH_4 retrieval
co2_profile_apriori	float	n, m	1e-6	A-priori mole fraction profile of atmospheric CO_2 in ppm
ch4_profile_apriori	float	n, m	1e-9	A-priori mole fraction profile of atmospheric CH_4 in ppb
raw_xco2	float	n	ppm	Retrieved 1.6µm XCO_2
raw_xch4	float	n	ppb	Retrieved 1.6µm XCH_4
raw_xco2_error	float	n	ppm	Retrieved 1.6µm XCO_2 error
raw_xch4_error	float	n	ppb	Retrieved 1.6µm XCH_4 error
model_xco2	float	n	ppm	Model XCO_2 component of the final proxy data product
model_xco2_range	float	n	ppm	Maximum difference (in ppm) between model XCO_2 from GEOS-Chem, CarbonTracker and LMDZ
exposure_id	char	n, 22		Exposure identification number of the sounding

surface_altitude	float	n	metres	Altitude is the (geometric) height above the geoid, which is the reference geopotential surface
surface_altitude_stdev	float	n	metres	Standard deviation of the surface elevation within the area of the GOSAT sounding, as derived from the SRTM database
surface_air_pressure_apriori	float	n	hPa	A-priori surface pressure value
surface_air_pressure_apriori_std	float	n	hPa	A-priori surface pressure standard deviation
gain	byte	n		GOSAT TANSO-FTS instrument gain mode. 1 indicates high gain. 0 indicates medium gain
air_temperature_apriori	float	n, m	K	Air temperature is the bulk temperature of the air, not the surface (skin) temperature
h2o_profile_apriori	float	n, m	ppm	A-priori mole fraction profile of atmospheric H_2O in ppm
total_aod	float	n		Retrieved total aerosol optical depth
aod_type1	float	n		Retrieved AOD (small)
aod_type2	float	n		Retrieved AOD (large)
cirrus	float	n		Retrieved AOD (cirrus)
retr_flag	byte	n		Retrieval type flag (0 = land, 1 = glint)

References

Boesch et al., 2011: Boesch, H., D. Baker, B. Connor, D. Crisp, and C. Miller, Global characterization of CO_2 column retrievals from shortwave-infrared satellite observations of the Orbiting Carbon Observatory-2 mission, Remote Sensing, 3 (2), 270-304, 2011.

Chédin et al. 2003: Chédin, A., Saunders, R., Hollingsworth, A., Scott, N. A., Matricardi, M., Etcheto, J., Clerbaux, C., Armante, R. and Crevoisier, C.: The feasibility of monitoring CO_2 from high resolution infrared sounders. J. Geophys. Res., 108, ACH 6-1–6-19, doi: 10.1029/2001JD001443, 2003.

Connor et al., 2008: Connor, B. J., Boesch, H., Toon, G., Sen, B., Miller, C., and Crisp, D.: Orbiting Carbon Observatory: Inverse method and prospective error analysis, J. Geophys. Res., 113, D05305, doi:10.1029/2006JD008336, 2008.

Corbin et al., 2008: Corbin, K. D., A. S. Denning, L. Lu, J.-W. Wang, and I. T. Baker, Possible representation errors in inversions of satellite CO_2 retrievals, J. Geophys. Res., 113, D02301, doi:10.1029/2007JD008716, 2008.

Kitaoka et al., 2019: Kitaoka, F., Knuteson, R. O., Kuze, A., Shiomi, K., Suto, H., Yoshida, J., Kondo, S., and Saitoh, N., Calibration, Level 1 processing, and radiometric validation for TANSO-FTS TIR on GOSAT, IEEE Trans. Geosci. Remote Sens., 57, 3490-3500, doi: 10.1109/TGRS.2018.2885162, 2019.

Kuze et al., 2009: Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T. (2009), 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, 2009.

Natraj et al., 2008: Natraj, V., Boesch, H., Spurr, R. J. D., and Yung, Y. L.: Retrieval of XCO_2 from simulated Orbiting Carbon Observatory measurements using the fast linearized R-2OS radiative transfer model, Journal of Geophysical Research D: Atmospheres, 113(11), 2008.

O'Dell, 2010: O'Dell, C. W.: Acceleration of multiple-scattering, hyper-spectral radiative transfer calculations via low-streams interpolation, Journal of Geophysical Research D: Atmospheres, 115(10), 2010.

Parker et al., 2011: Parker, R., Boesch, H., Cogan, A., et al., Methane Observations from the Greenhouse gases Observing SATellite: Comparison to ground-based TCCON data and Model Calculations, *Geophys. Res. Lett*., doi:10.1029/2011GL047871, 2011.

Parker et al., 2015: Parker, R. J., Boesch, H., Byckling, K., Webb, A. J., Palmer, P. I., Feng, L., Bergamaschi, P., Chevallier, F., Notholt, J., Deutscher, N., Warneke, T., Hase, F., Sussmann, R., Kawakami, S., Kivi, R., Griffith, D. W. T., and Velazco, V.: Assessing 5 years of GOSAT Proxy XCH4 data and associated uncertainties, Atmos. Meas. Tech., 8, 4785- 4801, doi:10.5194/amt-8-4785-2015, 2015.

Saitoh et al., 2009: Saitoh, N., Imasu, R., Ota, Y., and Niwa, Y.: CO_2 retrieval algorithm for the thermal infrared spectra of the greenhouse gases observing satellite: Potential of retrieving CO_2 vertical profile from high-resolution FTS sensor, Journal of Geophysical Research D: Atmospheres, 114(17), 2009.

Somkuti et al., 2017: Somkuti, P., Boesch, H., Natraj, V., Kopparla, P., Application of a PCA‐Based Fast Radiative Transfer Model to XCO2 Retrievals in the Shortwave Infrared, Journal of Geophysical Research: Atmospheres, 122(19), doi: 10.1002/2017JD027013, 2017.

Spurr, 2008: Spurr, R.: LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems. In Light Scattering Reviews 3 (pp. 229-275). Springer Berlin Heidelberg, 2008.

Spurr et al., 2011: Spurr, R., & Natraj, V.: A linearized two-stream radiative transfer code for fast approximation of multiple-scatter fields. Journal of Quantitative Spectroscopy and Radiative Transfer, 112(16), 2630-2637, 2011.

TRD GHG, 2017: Buchwitz, M., Aben, I., Anand, J., Armante, R., Boesch, H., Crevoisier, C., Detmers, R. G., Hasekamp, O. P., Reuter, M., Schneising-Weigel, O., Target Requirement Document, Copernicus Climate Change Service (C3S) project on satellite-derived Essential Climate Variable (ECV) Greenhouse Gases (CO_2 and CH_4) data products (project C3S_312a_Lot6), Version 1, 28-March-2017, pp. 52, 2017.

Yokota et al., 2009: Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global concentrations of CO_2 and CH_4 retrieved from GOSAT: First preliminary results, Sola, 5, 160-163, doi: 10.2151/sola.2009-041, 2009.

Copernicus Climate Change Service

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: info@copernicus-climate.eu