ECMWF COPERNICUS REPORT

Copernicus Climate Change Service

# Product User Guide and Specification (PUGS) – ANNEX B for products CO2_GO2_SRFP, CH4_GO2_SRFP (v2.0.0, 2019-2022)

## C3S2_312a_Lot2_DLR – Atmosphere

Issued by: Andrew Barr and Tobias Borsdorff, SRON, The Netherlands

Date: 13/11/2023

Ref: C3S2_312a_Lot2_D-WP2_PUGS-2023-GHG_ANNEX-B_v7.2

Official reference number service contract: 2021/C3S2_312a_Lot2_DLR/SC1

PROGRAMME OF THE EUROPEAN UNION

IMPLEMENTED BY ECMWF

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 on the European Union (Contribution Agreement signed on 22/07/2021). All information in this document is provided “as is” and no guarantee of warranty is given that the information is fit for any particular purpose.
The users thereof use the information at their sole risk and liability. For the avoidance of all doubt, the European Commission and the European Centre for Medium-Range Weather Forecasts have no liability in respect of this document, which is merely representing the author’s view.

# Contributors

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

**SRON NETHERLANDS INSTITUTE FOR SPACE RESEARCH, LEIDEN, THE NETHERLANDS**
**(SRON)**
T. Borsdorff
A. G. Barr
O. P. Hasekamp

# 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 | 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: Correction of typos. | All |
| 3.1 | 29-November-2019 | One misplaced string removed after 2nd review by Assimila | Title page |
| 4.0 | 18-August-2020 | Update for CDR4 (2009-2019) | All |
| 5.0 | 18-February-2021 | Update for CDR5 (2009-mid2020) | All |
| 6.0 | 4-August-2022 | Update for CDR6: GOSAT-2 retrievals (2019 - 2021) | All |
| 6.1 | 6-December-2022 | Update after review (use of new template, several improvements at various places) | All |
| 6.3 | 18-April-2023 | Update after 2nd review. Several improvements at various places. | All |
| 7.0 | 24-August-2023 | Update for CDR7 (until end of 2022) | All |
| 7.1 | 30-October-2023 | Minor modifications after review | All |
| 7.2 | 13-November-2023 | Minor modifications after review | All |

## List of datasets covered by this document

| Deliverable ID | Product title | Product type (CDR, ICDR) | Version number | Delivery date |
|---|---|---|---|---|
| WP2-FDDP-GHG-v2 | CO2_GO2_SRFP | CDR 7 | 2.0.0 | 31-Aug-2023 |
| WP2-FDDP-GHG-v2 | CH4_GO2_SRFP | CDR 7 | 2.0.0 | 31-Aug-2023 |

## Related documents

| Reference ID | Document |
|---|---|
| D1 | **Main PUGS:**<br>Buchwitz, M., et al., Product User Guide and Specification (PUGS) – Main document for Greenhouse Gas (GHG: CO2 & CH4) data set CDR 7 (2003-2022), project C3S2_312a_Lot2_DLR – Atmosphere, 2023.<br><br>*(this document is an ANNEX to the Main PUGS)* |
| D2 | **TRD GAD GHG, 2020:** Buchwitz, M., Aben, I., Armante, R., Boesch, H., Crevoisier, C., Hasekamp, O. P., Wu, L.., Reuter, M., Schneising-Weigel, O., Target Requirement and Gap Analysis 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_312b_Lot2), Version 2.11, 9-April-2020, pp. 80, 2020. |
| D3 | Barr, A. G., et al., Algorithm Theoretical Basis Document (ATBD) – ANNEX B for product CH4 GO2 SRFP and CO2 SRFP (v2.0.0, 2019-2022), Technical Report C3S project C3S2 312a Lot2 DLR – Atmosphere, 2023. |

# Acronyms

| Acronym | Definition |
|---|---|
| ATBD | Algorithm Theoretical Basis Document |
| CAR | Climate Assessment Report |
| C3S | Copernicus Climate Change Service |
| CCI | Climate Change Initiative |
| CDR | Climate Data Record |
| CDS | (Copernicus) Climate Data Store |
| CRG | Climate Research Group |
| D/B | Data base |
| EC | European Commission |
| ECMWF | European Centre for Medium Range Weather Forecasting |
| ECV | Essential Climate Variable |
| EO | Earth Observation |
| ESA | European Space Agency |
| EU | European Union |
| EUMETSAT | European Organisation for the Exploitation of Meteorological Satellites |
| FP | Full Physics retrieval method |
| FTIR | Fourier Transform InfraRed |
| FTS | Fourier Transform Spectrometer |
| GCOS | Global Climate Observing System |
| GEOSS | Global Earth Observation System of Systems |
| GHG | GreenHouse Gas |
| GOSAT | Greenhouse Gases Observing Satellite |
| GOSAT-2 | Greenhouse Gases Observing Satellite 2 |
| IPCC | International Panel in Climate Change |
| IUP | Institute of Environmental Physics (IUP) of the University of Bremen, Germany |
| JAXA | Japan Aerospace Exploration Agency |
| KIT | Karlsruhe Institute of Technology |
| L1 | Level 1 |
| L2 | Level 2 |
| L3 | Level 3 |
| L4 | Level 4 |
| LMD | Laboratoire de Météorologie Dynamique |
| MACC | Monitoring Atmospheric Composition and Climate, EU GMES project |
| NA | Not applicable |

| | |
|---|---|
| NetCDF | Network Common Data Format |
| NIES | National Institute for Environmental Studies |
| NIR | Near Infra Red |
| NOAA | National Oceanic and Atmospheric Administration |
| Obs4MIPs | Observations for Climate Model Intercomparisons |
| ppb | Parts per billion |
| ppm | Parts per million |
| PR | (light path) PRoxy retrieval method |
| PVIR | Product Validation and Intercomparison Report |
| QA | Quality Assurance |
| QC | Quality Control |
| RemoTeC | Retrieval algorithm developed by SRON |
| REQ | Requirement |
| RMS | Root-Mean-Square |
| RTM | Radiative transfer model |
| SNR | Signal-to-Noise Ratio |
| SRON | SRON Netherlands Institute for Space Research |
| SWIR | Short Wave Infra Red |
| SZA | Solar Zenith Angle |
| TANSO | Thermal And Near infrared Sensor for carbon Observation |
| TANSO-FTS | Fourier Transform Spectrometer on GOSAT |
| TANSO-FTS-2 | Fourier Transform Spectrometer on GOSAT-2 |
| TBC | To be confirmed |
| TBD | To be defined / to be determined |
| TCCON | Total Carbon Column Observing Network |
| TIR | Thermal Infra Red |
| TR | Target Requirements |
| TRD | Target Requirements Document |
| URD | User Requirements Document |
| WMO | World Meteorological Organization |
| Y2Y | Year-to-year (bias variability) |

# General definitions

**Essential climate variable (ECV)**

An ECV is a physical, chemical, or biological variable or a group of linked variables that critically contributes to the characterization of Earth's climate.

**Climate data record (CDR)**

The US National Research Council (NRC) defines a CDR as a time series of measurements of sufficient length, consistency, and continuity to determine climate variability and change.

**Fundamental climate data record (FCDR)**

A fundamental climate data record (FCDR) is a CDR of calibrated and quality-controlled data designed to allow the generation of homogeneous products that are accurate and stable enough for climate monitoring.

**Thematic climate data record (TCDR)**

A thematic climate data record (TCDR) is a long time series of an essential climate variable (ECV).

**Intermediate climate data record (ICDR)**

An intermediate climate data record (ICDR) is a TCDR which undergoes regular and consistent updates, for example because it is being generated by a satellite sensor in operation.

**Satellite data processing levels**

The NASA Earth Observing System (EOS) distinguishes six processing levels of satellite data, ranging from Level 0 (L0) to Level 4 (L4) as follows.

| | |
|---|---|
| L0 | Unprocessed instrument data |
| L1A | Unprocessed instrument data alongside ancillary information |
| L1B | Data processed to sensor units (geo-located calibrated spectral radiance and solar irradiance) |
| L2 | Derived geophysical variables (e.g., $XCO_2$) over one orbit |
| L3 | Geophysical variables averaged in time and mapped on a global longitude/latitude horizontal grid |
| L4 | Model output derived by assimilation of observations, or variables derived from multiple measurements (or both) |

# Table of Contents

## Scope of document

This document is a Product User Guide and Specification (PUGS) for the Copernicus Climate Change Service (C3S, http://climate.copernicus.eu/) greenhouse gas (GHG) component as covered by project C3S2_312a_Lot2.

Within this project, satellite-derived atmospheric carbon dioxide ($CO_2$) and methane ($CH_4$) Essential Climate Variable (ECV) data products have been 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 GHG satellite-derived 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 C3S2 products CO2_GO2_SRFP and CH4_GO2_SRFP v2.0.0.

These products are $XCO_2$ and $XCH_4$ Level 2 products as retrieved from GOSAT-2 using algorithms developed at SRON, The Netherlands.

# Executive summary

Because of their important role for climate, they are classified as Essential Climate Variables (ECVs). The ECV GHG as formulated by GCOS (Global Climate Observing System) is defined as follows: “Retrievals of greenhouse gases, such as $CO_2$ and $CH_4$, of sufficient quality to estimate regional sources and sinks” (*GCOS-154*). This definition contains already the main application of these atmospheric data products; namely to use them (in combination with appropriate modelling) to obtain (improved) information on their (primarily surface) sources and sinks.

Both gases ($CO_2$ and $CH_4$) have a long lifetime in the atmosphere. As a consequence of this fact and related human emissions, the atmospheric concentrations of these gases are relatively high compared to other atmospheric trace gases. As a result of this, even a moderate to strong (surface) source or sink typically only results in a relatively small local or regional change (enhancement or depletion relative to the surrounding region) in their vertical columns or their mid/upper tropospheric concentration. The observational requirements are therefore very demanding in particular with respect to random and systematic errors and stability.

This document is Annex B of the Product User Guide (PUG), which is a deliverable of the C3S2 project. This document describes the $XCO_2$ and $XCH_4$ Full Physics data products generated by the RemoTeC algorithm (CO2_GO2_SRFP and CH4_GO2_SRFP). The description includes quality flags, metadata, data format, product grid, defined limitations, bias correction, and the product (column) averaging kernels, as well as a description of how to use them appropriately.

The greenhouse gas (GHG) activities of this C3S project and its C3S pre-cursor projects are essentially the operational continuation of the research and development (R&D) pre-cursor projects GHG-CCI and GHG-CCI+ of ESA’s Climate Change Initiative (CCI). R&D for the GOSAT-2 products is currently an ongoing activity of the ESA GHG-CCI+ project*.

The description includes quality flags and metadata, data format, product grid, known limitations, bias correction, and the product (column) averaging kernels and a description of how to use them. The two data products ($XCO_2$ and $XCH_4$) are produced by the same retrieval, and separated into different products in the post-processing therefore in the following sections, we treat both products as one, and give the necessary details for each where necessary.

Section 1.1 describes the product, covering the relevant information for the algorithm and input data, as well as describing the bias correction process. Section 1.2 presents the target requirements for these products, and section 1.3 outlines how to use the data. Finally, section 2 provides information on where to find, and how to access, the data.

*http://climate.esa.int/en/projects/ghgs/ - 16 Jan 2023

# 1. Full Physics Products

## 1.1 Data product description

The Japanese Greenhouse gases Observing SATellite-2 (GOSAT-2) was launched on 29th October 2018 and started operational observations form February 2019. GOSAT-2 provides dedicated global measurements of total column $CO_2$ and $CH_4$ from its SWIR bands. It is equipped with two instruments, the Thermal And Near Infrared Sensor for carbon Observations - Fourier Transform Spectrometer-2 (TANSO-FTS2) with spectral channels presented in Table 1. A more extensive description of the instrument can be found in the ATBD Annex-B (D3 in the related documents section).

Table 1: GOSAT-2-FTS bands.

| Channel | Wavelength range [nm] | Resolution [$cm^{-1}$] |
|---|---|---|
| 1 | 758-775 | 0.2 |
| 2 | 1460-1720 | 0.2 |
| 3 | 1920-2330* | 0.2 |
| 4 | 5560-8400 | 0.2 |
| 5 | 8400-14300 | 0.2 |

***GOSAT-1 only had a spectral range up to 2080nm.**

The CH4_GO2_SRFP and CO2_GO2_SRFP products are retrieved from GOSAT-2 TANSO-FTS-2 NIR and SWIR spectra using the RemoTeC algorithm that is being jointly developed at SRON and KIT. The algorithm retrieves simultaneously $XCH_4$ and $XCO_2$. For the retrieval, we analyze four spectral regions: the 0.77 µm oxygen band, two $CO_2$ bands at 1.61 and 2.06 µm, as well as a $CH_4$ band at 1.64 µm. Within the retrieval procedure the sub-columns of $CO_2$ and $CH_4$ in different altitude layers are being retrieved. To obtain the column averaged dry air mixing ratios $XCO_2$ and $XCH_4$ the sub-columns are summed up to get the total column which is divided by the dry-air columns obtained from ECMWF model data in combination with a surface elevation data base. The retrieved $XCH_4$ and $XCO_2$ have been extensively validated with ground based Total Carbon Column Observing Network (TCCON) measurements. To further improve accuracy a bias correction has been developed based on TCCON comparisons. We use the GGG2020 TCCON dataset.

More details on the technical aspects of the retrievals and auxiliary data can be found in the ATBD Annex-B (D3).

### 1.1.1 Bias Correction

From comparison with TCCON (Wunch et al. 2015) it was found that the error in $XCH_4$ correlates with the retrieved albedo $\alpha$ at 1.6 μm in band 2 and the ratio of $O_2$ retrieved relative to the prior. Based on this correlation the following bias correction for land retrievals has been developed for $XCO_2$:

$$XCO2_{corr} = XCO2 * (a + b * \alpha) \quad (1)$$

with *a*=0.9896 and *b*=0.0532 for land data. For ocean retrievals we apply a bias correction following:

$$XCO2_{corr} = XCO2 * (a + b * r) \quad (2)$$

with *a*=1.44743 and *b*=-0.45154, and *r* the retrieved $O_2$ divided by the prior $O_2$.

The bias correction parameters are obtained from fits to the GOSAT-2-TCCON differences and the subsequent correlation with TCCON is illustrated in Figure 1.

Similar correlations were found for $XCH_4$. Based on this correlation the following bias correction has been developed for $XCH_4$:

$$XCH4_{corr} = XCH4 * (a + b * \alpha) \quad (3)$$

with *a* = 0.99091, *b* = 0.03648 for land data. For ocean retrievals we apply a bias correction following:

$$XCH4_{corr} = XCH4 * (a + b * r) \quad (4)$$

with *a*=1.44648 and *b*=-0.45599, and *r* the retrieved $O_2$ divided by the prior $O_2$.

The bias correction parameters are obtained from fits to the GOSAT-2-TCCON differences and the subsequent correlation with TCCON is illustrated in Figure 2.

Figures 3 and 4 show global maps of $XCO_2$ and $XCH_4$, respectively, for land and ocean (sunglint) soundings separately. These have been sampled onto a 1x1 degree grid and cover the entire time range of available GOSAT-2 data.

Figure 1: Co-located GOSAT-2-TCCON $XCO_2$ measurements for land measurements

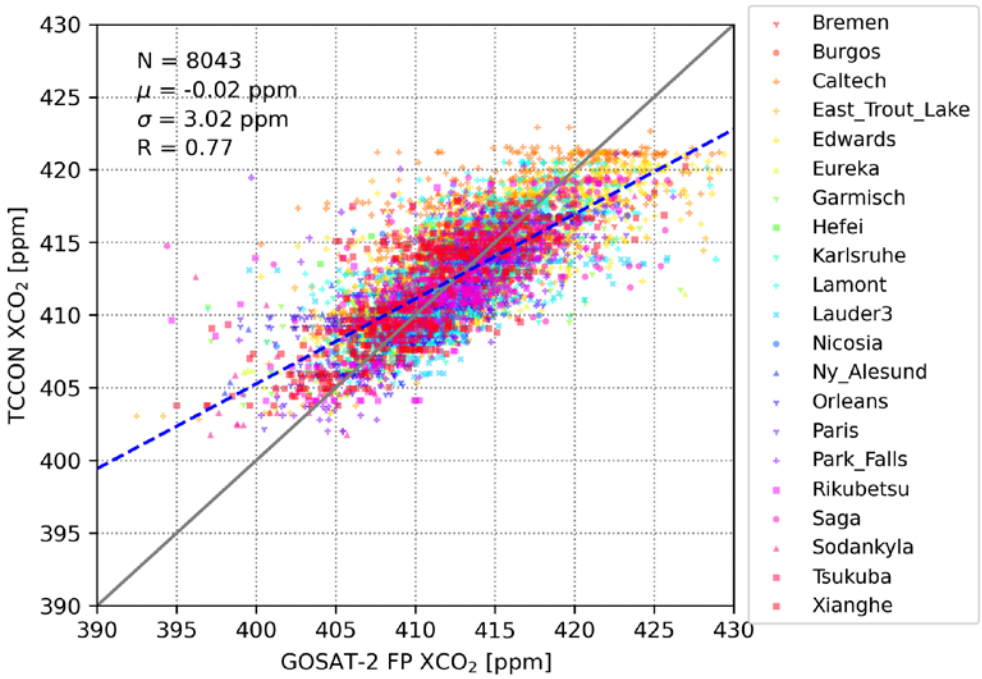



Figure 2: Co-located GOSAT-2-TCCON $XCH_4$ measurements for land measurements

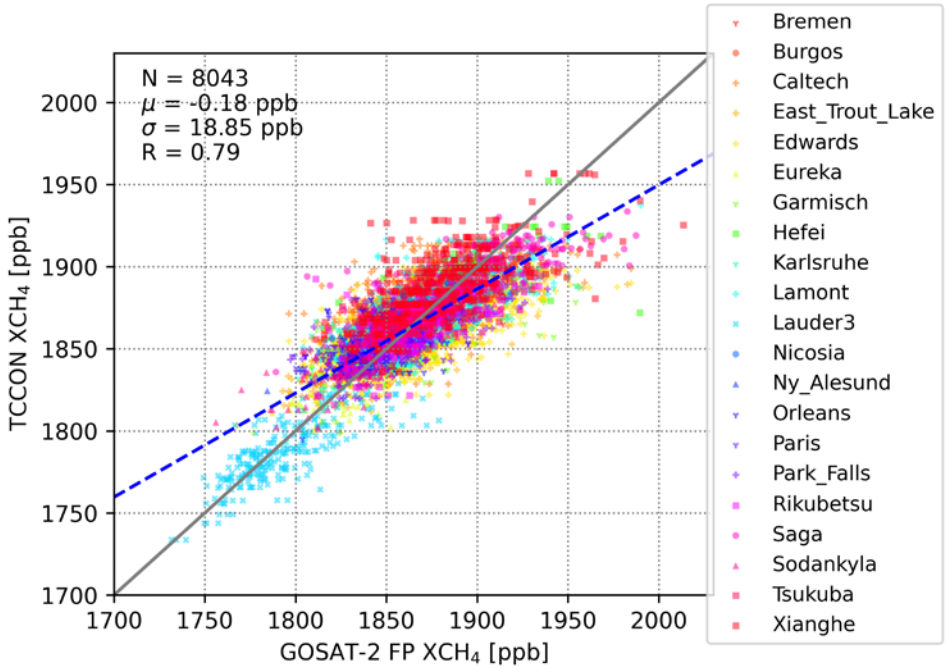


Figure 3: Global $XCO_2$ for the 2019-2022 period for the CO2_GO2_SRFP product on a 1 by 1 degree resolution for both land (top) and ocean/sunglint (bottom) soundings.

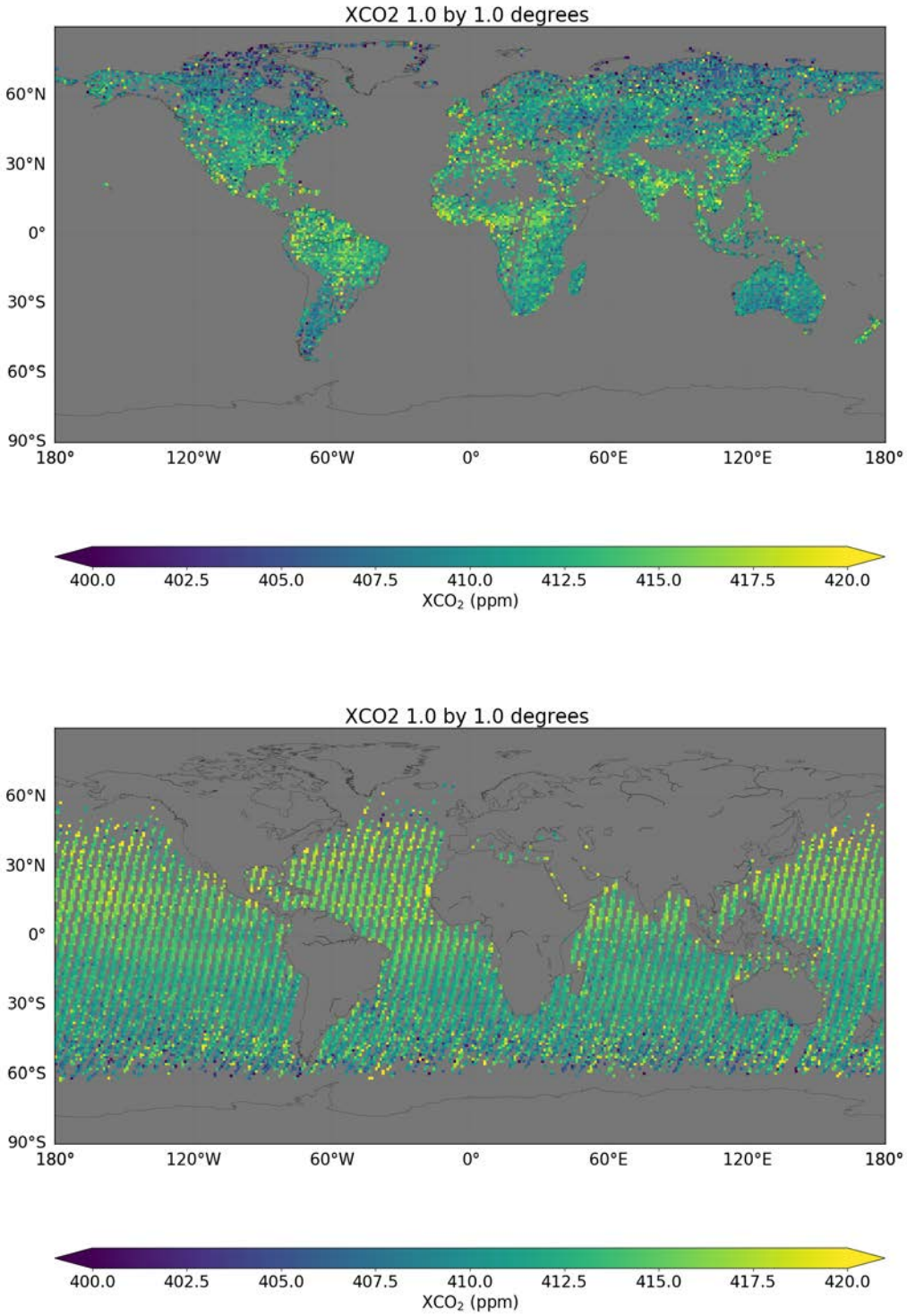
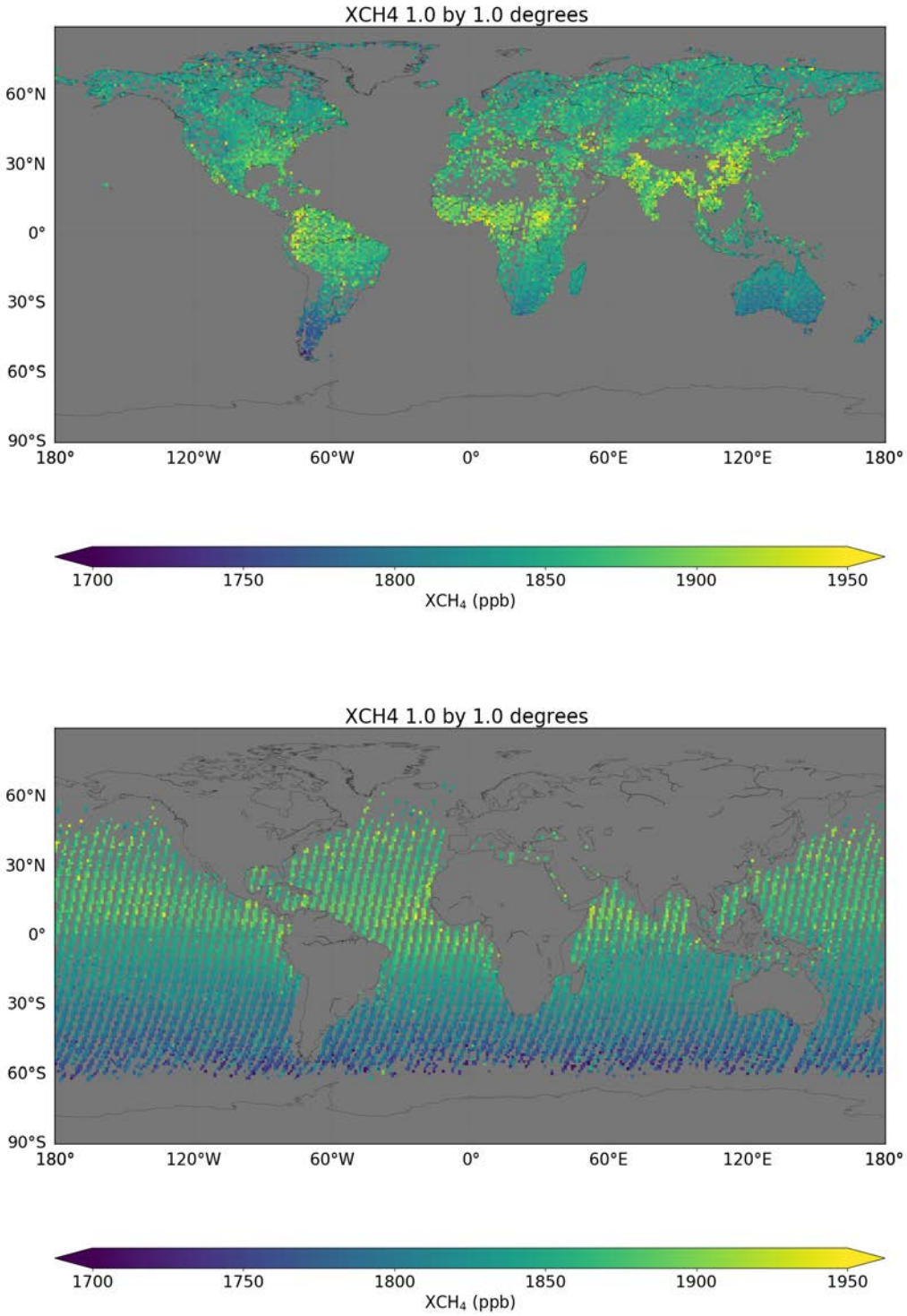


Figure 4: Global $XCH_4$ for the 2019-2022 period for the CH4_GO2_SRFP product on a 1 by 1 degree resolution for both land (top) and ocean/sunglint (bottom) soundings.



## 1.2 Target requirements

The target requirements for the products in this document are taken from the Target Requirements Document (D2) for greenhouse gas retrievals and presented in Table 2. $CO_2$ and $CH_4$ are important climate-relevant atmospheric gases, so-called greenhouse gases (GHG). Because of their important role for climate, they are classified as Essential Climate Variables (ECVs). The ECV GHG as formulated by GCOS (Global Climate Observing System) is defined as follows: "Retrievals of greenhouse gases, such as $CO_2$ and $CH_4$, of sufficient quality to estimate regional sources and sinks" (*GCOS-154*). This definition contains already the main application of these atmospheric data products; namely to use them (in combination with appropriate modelling) to obtain (improved) information on their (primarily surface) sources and sinks.

Both gases ($CO_2$ and $CH_4$) have lifetimes in the atmosphere of hundreds of years and 10-15 years, respectively. As a consequence of this fact and related human emissions, the atmospheric concentrations of these gases are relatively high compared to other atmospheric trace gases. As a result of this, even a moderate to strong (surface) source or sink typically only results in a relatively small local or regional change (enhancement or depletion relative to the surrounding region) in their vertical columns or their mid/upper tropospheric concentration. The observational requirements are therefore very demanding in particular with respect to random and systematic errors and stability.

Table 2: Goal (G), breakthrough (B) and threshold (T) target requirements for $XCH_4$ and $XCO_2$.

| **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) |

## 1.3 Data usage information

### 1.3.1 Product Content and Format

The CH4_GO2_SRFP and CO2_GO2_SRFP v2.0.0 data products are stored per day in a single NetCDF (version 4) file. Retrieval results are provided for the individual GOSAT-2 spatial footprints, i.e., no averaging has been applied. The product file contains a set of common variables, i.e., the retrieved column averaged dry air mixing ratios $XCO_2$ and $XCH_4$ with bias correction, averaging kernels and quality flags, as well as secondary variables specific for the RemoTeC algorithm. Common and retrieval-specific variables for the $XCO_2$ product are given in Tables 3 and 4, respectively. Common and retrieval-specific variables for the $XCH_4$ product are given in Tables 5 and 6, respectively.

Table 3: Common variables for the CO2_GO2_SRFP product

| Name | Type | Dim. | Units | Description |
|---|---|---|---|---|
| solar_zenith_angle | float | n | degrees | Angle between line of sight to the sun and local vertical |
| sensor_zenith_angle | float | n | degrees | Angle between the line of sight to the sensor and the local vertical |
| time | float | n | seconds | Seconds since 1970-01-01 00:00:00 |
| longitude | float | n | degrees_east | Center longitude |
| latitude | float | n | degrees_north | Center latitude |
| pressure_levels | float | n, 13 | hPa | Pressure levels |
| pressure_weight | float | n, 12 | | Layer dependent weights needed to apply the averaging kernels |
| xco2 | float | n | 1e-6 | Retrieved column dry-air mole fraction of atmospheric carbon dioxide (XCO2) in ppm |
| xco2_uncertainty | float | n | 1e-6 | 1-sigma uncertainty of the retrieved column-average dry-air mole fraction of atmospheric carbon dioxide |
| xco2_averaging_kernel | float | n, 12 | | Normalized column averaging kernel |
| co2_profile_apriori | float | n, 12 | 1e-6 | A priori dry-air mole fraction profile of atmospheric carbon dioxide |
| xco2_quality_flag | int | n | | Quality flag for XCO2 retrieval, 0 = good, 1 = bad |

Table 4: Product specific (additional) variables for the CO2_GO2_SRFP product

| Name | Type | Dim. | Units | Description |
|---|---|---|---|---|
| flag_landtype | int | n | | 0 = land, 1 = ocean |
| flag_sunglint | int | n | | 0 = no sunglint, 1 = sunglint |
| gain | char | n | | Number of gain coefficient calculated from solar calibration mode data. [1P 1S 2P 2S 3P 3S] |
| exposure_id | int | n | | Exposure identification number of the sounding |
| l1b_name | char | n | | Name of the Level 1B file of the sounding |
| signal_to_noise_window | float | n, 4, 2 | | Signal to noise ratio per retrieval window and for both polarization directions |
| dry_airmass_layer | float | n, 12 | m-2 | Dry airmass per layer |
| altitude | float | n | m | Vertical distance above the surface |
| air_temperature | float | n, 13 | K | The bulk temperature of the air at each level |
| surface_elevation_stdev | float | n | m | Standard deviation of the surface elevation within the sounding |
| x_wind | float | n, 13 | m s-1 | Eastward wind velocity |
| y_wind | float | n, 13 | m s-1 | Northward wind velocity |
| chi2 | float | n | | Chi-squared value of the sounding |
| optical_thickness_of_atmosphere_layer_due_to_ambient_aerosol | float | n, 4 | | Scattering optical thickness per retrieval window |
| raw_xco2 | float | n | 1e-6 | Retrieved column dry-air mole fraction of atmospheric carbon dioxide (XCO2) in ppm before bias correction |
| raw_xco2_err | float | n | 1e-6 | 1-sigma statistical uncertainty of the retrieved column-average dry-air mole fraction of atmospheric carbon dioxide |
| h2o_column | float | n | m-2 | Retrieved total water column |
| surface_albedo_758 | float | n | | The retrieved albedo at 758 nm |
| surface_albedo_1593 | float | n | | The retrieved albedo at 1593 nm |
| surface_albedo_1629 | float | n | | The retrieved albedo at 1629 nm |
| surface_albedo_2042 | float | n | | The retrieved albedo at 2042 nm |
| intensity_offset_o2a | float | n | W cm-2 | The retrieved intensity offset in the O2A band |
| aerosol_size | float | n | | Retrieved size parameter of the aerosol distribution |
| aerosol_central_height | float | n | m | Peak height of the aerosol Gaussian height distribution |
| aerosol_total_column | float | n | m-2 | Retrieved total aerosol column |

Table 5: Common variables for the CH4_GO2_SRFP product

| Name | Type | Dim. | Units | Description |
|---|---|---|---|---|
| solar_zenith_angle | float | n | degrees | Angle between line of sight to the sun and local vertical |
| sensor_zenith_angle | float | n | degrees | Angle between the line of sight to the sensor and the local vertical |
| time | float | n | seconds | Seconds since 1970-01-01 00:00:00 |
| longitude | float | n | degrees_east | Center longitude |
| latitude | float | n | degrees_north | Center latitude |
| pressure_levels | float | n, 13 | hPa | Pressure levels |
| pressure_weight | float | n, 12 | | Layer dependent weights needed to apply the averaging kernels |
| xch4 | float | n | 1e-9 | Retrieved column dry-air mole fraction of atmospheric methane (XCH4) in ppb |
| xch4_uncertainty | float | n | 1e-9 | 1-sigma uncertainty of the retrieved column-average dry-air mole fraction of atmospheric methane |
| xch4_averaging_kernel | float | n, 12 | | Normalized column averaging kernel |
| ch4_profile_apriori | float | n, 12 | 1e-9 | A priori dry-air mole fraction profile of atmospheric methane |
| xch4_quality_flag | int | n | | Quality flag for XCH4 retrieval, 0 = good, 1 = bad |

Table 6: Product specific (additional) variables for the CH4_GO2_SRFP product

| Name | Type | Dim. | Units | Description |
|---|---|---|---|---|
| flag_landtype | int | n | | 0 = land, 1 = ocean |
| flag_sunglint | int | n | | 0 = no sunglint, 1 = sunglint |
| gain | char | n | | Number of gain coefficient calculated from solar calibration mode data. [1P 1S 2P 2S 3P 3S] |
| exposure_id | int | n | | Exposure identification number of the sounding |
| l1b_name | char | n | | Name of the Level 1B file of the sounding |
| signal_to_noise_window | float | n, 4, 2 | | Signal to noise ratio per retrieval window and for both polarization directions |
| dry_airmass_layer | float | n, 12 | m-2 | Dry airmass per layer |
| altitude | float | n | m | Vertical altitude above the surface |
| air_temperature | float | n, 13 | K | The bulk temperature of the air at each level |
| surface_altitude_stdev | float | n | m | Standard deviation of the surface elevation within the sounding |
| x_wind | float | n, 13 | m s-1 | Eastward wind velocity |
| y_wind | float | n, 13 | m s-1 | Northward wind velocity |
| chi2 | float | n | | Chi-squared value of the sounding |
| optical_thickness_of_atmosphere_layer_due_to_ambient_aerosol | float | n, 4 | | Scattering optical thickness per retrieval window |
| raw_xch4 | float | n | 1e-9 | Retrieved column dry-air mole fraction of atmospheric methane (XCH4) in ppb before bias correction |
| raw_xch4_err | float | n | 1e-9 | 1-sigma statistical uncertainty of the retrieved column-average dry-air mole fraction of atmospheric methane |
| h2o_column | float | n | m-2 | Retrieved total water column |
| surface_albedo_758 | float | n | | The retrieved albedo at 758 nm |
| surface_albedo_1593 | float | n | | The retrieved albedo at 1593 nm |
| surface_albedo_1629 | float | n | | The retrieved albedo at 1629 nm |
| surface_albedo_2042 | float | n | | The retrieved albedo at 2042 nm |
| intensity_offset_o2a | float | n | W cm-2 | The retrieved intensity offset in the O2A band |
| aerosol_size | float | n | | Retrieved size parameter of the aerosol distribution |
| aerosol_central_height | float | n | m | Peak height of the aerosol Gaussian height distribution |
| aerosol_total_column | float | n | m-2 | Retrieved total aerosol column |

### 1.3.2 Quality Flags and Metadata

There are quality flags "xco2_quality_flag" and “xch4_quality_flag” included in the data files. The quality flag can have 2 values:

- 0: retrieval quality has been checked
- 1: data should not be used (e.g., bad fit to data, residual cloud contamination)

For a GOSAT-2 ground pixel to be processed by the RemoTeC Full Physics algorithm it has to fulfill the following criteria: GOSAT-2 nominal quality flags should be good and the standard deviation of the elevation in the pixel should be less than 1000 meters (to filter out the most extreme terrains). After the retrieval step the data that fulfill the following criteria are flagged as '0' for land:

- Number of iteration steps in retrieval < 31.
- $\chi 2$ of fit < 8
- SNR > 50
- Standard deviation of surface elevation within GOSAT-2 ground pixel should be < 100 m
- Aerosol Optical Thickness < 0.8
- 3 < Aero_size < 6
- SZA < 75°
- 0 < (2.4*albedo[0.76 micron]) – (1.13*albedo[2.0 micron]) < 0.8
- 0 < peak height of aerosol Gaussian height distribution < 8000
- 0 < cirrus signal (2 micron) < 2e-9
- 0.99 < CO2 (1.6 micron) / CO2 (2.0 micron) < 1.018
- 0.96 < O2 (retrieved) / O2 (prior) < 1.04
- 0.95 < H2O (1.6 micron) / H2O (2.0 micron) < 1.08

For sunglint:

- Number of iteration steps in retrieval < 31.
- $\chi 2$ of fit < 8
- SNR > 50

- Standard deviation of surface elevation within GOSAT-2 ground pixel should be < 100 m
- SZA < 75°
- 0 < (2.4*albedo[0.76 micron]) – (1.13*albedo[2.0 micron]) < 0.3
- 0 < cirrus signal (2 micron) < 2e-9
- 0.99 < CO2 (1.6 micron) / CO2 (2.0 micron) < 1.003
- 0.96 < O2 (retrieved) / O2 (prior) < 1.04
- 0.95 < H2O (1.6 micron) / H2O (2.0 micron) < 1.08

### 1.3.3 Recommended data usage

It is strongly recommended to only use the bias-corrected data in: "xco2" and “xch4” except if users explicitly correct for biases themselves (e.g., in an inverse modeling framework). Here, it should be noted that the bias correction has been developed independently for the different GOSAT-FTS-2 instrument settings (land & sunglint).

Also, use only data over land (land type=0) except for sunglint cases.
If the data are to be compared with other $XCO_2$ and/or $XCH_4$ data for which vertical profile information is available (e.g., inverse modeling, comparison to models, comparison to measured profiles), the column averaging kernels should be used. Here it should be noted that the column averaging kernels are to be applied to layer sub-columns (m-2), as these are the quantities directly retrieved in the RemoTeC algorithm.

For model comparisons the retrieved $XCO_2$ should be compared to $[VCO2]'_{model}/[VAIR]_{model}$ where $[VAIR]_{model}$ is the total dry air column provided by the model and $[VCO2]'_{model}$ is the model total $CO_2$ column after applying the column averaging kernel, viz.:

$$[VCO2]'_{model} = [VCO2]_{prior} + a^T(x_{model} - x_{prior}) \quad (3)$$

Where $[VCO2]_{prior}$ is the prior $CO_2$ total column used in the retrieval, $x_{model}$ is the vertical $CO_2$ profile from the model (as sub-columns) and $x_{prior}$ is the prior vertical profile from the retrieval. For application of the column averaging kernel the model vertical profile should be re-calculated on the vertical grid of the retrieval (preferred) or the averaging kernel has to be interpolated to the vertical grid of the model. This procedure holds in the same way for the SRFP $XCH_4$ product, but then replacing all instances of $CO_2$ with $CH_4$.

### 1.3.4 Tools for Reading the Data

The data are stored in Netcdf format which can be read with standard tools in the common programming languages (IDL, Matlab, Python, Fortran90, C++, etc). In python, several modules can be used to manipulate Netcdf files, such as netCDF4 (https://pypi.org/project/netCDF4/) , h5py (https://docs.h5py.org/en/stable/) and xarray (https://docs.xarray.dev/en/stable/). Netcdf functionality can also be implemented in compiled languages Fortran (https://docs.unidata.ucar.edu/netcdf-fortran/current/) and C++ (http://unidata.github.io/netcdf-cxx4/index.html).

### 1.3.5 Known Limitations and Issues

- The data retrieved for the land observations are considered highest quality and are well validated. In the "raw" retrievals (i.e., before bias correction) there is a bias between land and sunglint retrievals. Although these biases have been corrected in the bias-corrected products, there may still be a small residual bias left, especially due to the limited number of validation sites for sunglint retrievals.

## 2. Data access information

The data products and corresponding documentation are / will be made available via the Copernicus Climate Data Store (CDS):
https://cds.climate.copernicus.eu/#!/home

Direct link to $CO_2$ products:
https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-carbon-dioxide?tab=overview

Direct link to $CH_4$ products:
https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-methane?tab=overview

Tabs / riders lead to the following items:

- Overview
  - Short overview of all products
- Download data
  - Data access information
- Quality assessment
  - The CDS datasets are assessed by the Evaluation and Quality Control (EQC) function of C3S independently of the data supplier and the EQC information are available on this site.
- Documentation
  - Links to the following documents:
    - Algorithm Theoretical Basis Document (ATBD)
    - Product User Guide (PUG)
    - Product Quality Assurance Document (PQAD)
    - Product Quality Assessment Report (PQAR)
    - System Quality Assurance Document (SQAD)
    - Target Requirements and Gap Analysis (TRDGAD)
  - Note that pdf versions of all documents (including previous versions) are (also) available from here: https://www.iup.uni-bremen.de/carbon_ghg/cg_data.html#C3S_GHG
- View
  - Visualization of selected data products in terms of global maps

# References

**ESA-CCI-GHG-URDv2.1:** Chevallier, F., et al., User Requirements Document (URD), ESA Climate Change Initiative (CCI) GHG-CCI project, Version 2.1, 19 Oct 2016, link: http://www.esa-ghg-cci.org/?q=webfm_send/344, 2016.

**TRD GAD GHG, 2020:** Buchwitz, M., Aben, I., Armante, R., Boesch, H., Crevoisier, C., Hasekamp, O. P., Wu, L.., Reuter, M., Schneising-Weigel, O., Target Requirement and Gap Analysis 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_312b_Lot2), Version 2.11, 9-April-2020, pp. 80, 2020.

**Wunch et al. 2015:** Wunch, D., Toon, G.C., Sherlock, V., Deutscher, N.M., Liu, X., Feist, D.G., Wennberg, P.O., The Total Carbon Column Observing Network's GGG2014 Data Version. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA (available at: doi:10.14291/tccon.ggg2014.documentation.R0/1221662), 2015.

Copernicus Climate Change Service

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: https://support.ecmwf.int/