ESA Climate Change Initiative “Plus” (CCI+)

Product User Guide (PUG)
XCO₂ via RemoTeC from GOSAT2 (CO₂_GO₂_SRFP)
for the Essential Climate Variable (ECV)
Greenhouse Gases (GHG)

Version 1.1 (PUGv1.1)

for the RemoTeC XCO₂ GOSAT2 Data Product
CO₂_GO₂_SRFP (v1.0.0)

for the Essential Climate Variable (ECV)
Greenhouse Gases (GHG)

Written by:
GHG-CCI+ project team

Lead author:
Trismono Candra Krisna
(SRON Netherlands Institute for Space Research, Utrecht, The Netherlands)

Co-authors:
Ilse Aben, Lianghai Wu, Otto Hasekamp, Jochen Landgraf
(SRON Netherlands Institute for Space Research, Utrecht, The Netherlands)
Change log:

<table>
<thead>
<tr>
<th>VersionNr.</th>
<th>Date</th>
<th>Status</th>
<th>Reason for change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Update purpose of document</td>
</tr>
<tr>
<td>Version 1.1</td>
<td>4. Feb. 2021</td>
<td>As submitted</td>
<td>- Update after ESA reviews</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Remote typos</td>
</tr>
</tbody>
</table>
# Table of Contents

1. Purpose of document........................................................................................................4
2. Greenhouse gases Observing SATellite-2 (GOSAT-2) ..................................................5
3. RemoTeC retrieval algorithm...............................................................................................6
4. XCO₂ data product (RemoTeC-FP, Feb-Oct 2019) .........................................................7
   4.1 Global maps...................................................................................................................7
   4.2 Validation with TCCON ...............................................................................................9
   4.3 Bias correction ............................................................................................................12
5. Description of data format..................................................................................................13
   5.1 Product Content and Format......................................................................................13
   5.2 Quality Flags and Metadata ......................................................................................17
   5.3 Recommended data usage .........................................................................................19
   5.4 Tools for Reading the Data ......................................................................................19
6. References .........................................................................................................................20
1. Purpose of document

This document describes the Product User Guide (PUG) of the CO2_GO2_SRFP product which is a deliverable for the ESA GHG-CCI+ project led by University of Bremen, Germany.

Within the project satellite-derived atmospheric carbon dioxide (CO₂) and methane (CH₄) Essential Climate Variable (ECV) data products are generated and delivered to ESA for inclusion into the ESA-GHG-CCI+ database from which users can access these data products and the corresponding documentation.

The satellite-derived data products are:

- Column-averaged dry-air mixing ratios (mole fractions) of CO₂ and CH₄, denoted XCO₂ (in parts per million, ppm) and XCH₄ (in parts per billion, ppb), respectively.

This document will be focused on the XCO₂ Level-2 product retrieved using the GOSAT-2 Full Physics algorithm developed by SRON Netherlands Institute for Space Research, The Netherlands.
2. Greenhouse gases Observing SATellite-2 (GOSAT-2)

The Japanese Greenhouse gases Observing SATellite-2 (GOSAT2) was launched on 29th October 2018 and started operational observations from February 2019. GOSAT2 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) as well as a dedicated Cloud and Aerosol Imager-2 (TANSO-CAI-2).

The TANSO-FTS2 instrument (Nakajima et al., 2017) has five spectral bands with a high spectral resolution 0.2 cm$^{-1}$. Three operate in the SWIR at 0.75-0.77, 1.56-1.69 and at the extended 1.92-2.33 μm range, providing sensitivity to the near-surface absorbers. The fourth and fifth channels operating in the thermal infrared between 5.5-8.4 and 8.4-14.3 μm providing mid-tropospheric sensitivity.

The measurement strategy of TANSO-FTS2 is optimized for the characterization of continental-scale sources and sinks. TANSO-FTS2 utilizes a pointing mirror to perform off-nadir measurements at the same location on each 6-day repeat cycle. The pointing mirror allows TANSO-FTS2 to observe up to ±35° across track and ±40° along-track. These measurements nominally consist of 5 across track points spaced ~160km apart with a ground footprint diameter of approximately 9.7 km and a 4 second exposure duration. The satellite has an intelligent pointing monitor camera which makes it possible to adjust the line of sight of the FTS to steer away from cloud contaminated areas. Whilst the majority of data is limited to measurements over land where the surface reflectance is high, TANSO-FTS2 also observes in sunglint mode over the ocean.
3. RemoTeC retrieval algorithm

The CO2_GO2_SRFP product is retrieved from GOSAT-2 TANSO-FTS SWIR spectra using the RemoTeC algorithm that has been jointly developed by SRON and KIT. The algorithm retrieves simultaneously XCH₄ and XCO₂. For the retrieval, we analyze four spectral regions: the 0.77 μm oxygen band, two CO₂ bands at 1.61 and 2.06 μm, as well as a CH₄ band at 1.64 μm. Within the retrieval procedure the sub-columns of CO₂ and CH₄ in different altitude layers are being retrieved. To obtain the column averaged dry air mixing ratios XCO₂ and XCH₄ 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 XCO₂ has been validated with ground based TCCON measurements. To further improve accuracy a bias correction has been developed based on TCCON comparisons. We use the GGG2014 release of the TCCON data (Wunch et al., 2015). More details on the technical aspects of the retrievals can be found in the ATBD GO2-SRFP document (ATBDv1.1, 2020).
4. XCO₂ data product (RemoTeC-FP, Feb-Oct 2019)

In this section, we show examples of the new GOSAT-2 XCO₂ FP data product by showing averaged maps (Sect. 4.1) and by giving a summary of the validation results relative to TCCON (Sect. 4.2).

4.1 Global maps

Figure 1 and Figure 2 show global mean maps of the RemoTeC GOSAT-2 FP-XCO₂ data product. Figure 1 shows the bias-corrected XCO₂ data and Figure 2 the scaled random error, which is described in detail in (E3UBv1.1, 2021). As can be seen, the spatial coverage is limited by cloud-cover (the observations correspond to cloud free scenes), sun illumination conditions, etc.

Figure 1: Global XCO₂ [ppm] for the February-2019 - October 2019 period for the CO2_GOS2_SRFP product on a 2 by 2 degree resolution.
**Figure 2:** As **Figure 1** but for the corresponding uncertainty.
4.2 Validation with TCCON

This section summarizes the main validation results presented in the RemoTeC GOSAT-2 ESA GHG CCI+ End-to-End ECV Uncertainty Budget Version 1.1 (E3UBv1.1, 2021) document.

We used ground based TCCON GGG2014 (Wunch et al., 2015) data obtained from http://tccon.ornl.gov as reference data set. We co-located GOSAT-2 and TCCON measurements with a maximum time difference of 2h, a maximum distance of 2.5 degrees in both longitudinal and latitudinal directions. An altitude correction is applied to account for elevation differences.

In cases with multiple TCCON measurements of the same site co-locating with a GOSAT-2 sounding, we averaged the TCCON measurements. In total, we found about 1587 co-locations with TCCON during the nine-month validation period Feb-Oct 2019. As not all TCCON sites updated their datasets it was only possible to use 9 sites which limits the number of co-locations.

Figure 3 shows the co-locations of 9 sites with co-located non-glint observations. Statistics per site are shown in Error! Reference source not found.. Detailed bias and scatter, i.e., single sounding precision measured by the standard deviation of the difference to TCCON after removing systematic effect are described in E3UBv1.1 (2021).

The standard deviation of the site biases (spatial accuracy) is 0.9 ppm. TCCON observes these gases with a precision on mole fractions of ~0.15% and ~0.2% for CO₂ and CH₄ respectively (Toon et al., 2009). The single measurement precision of GOSAT-2 compared to TCCON amounts to 2.10 ppm.

The validation results can be summarized as follows:

- Single measurement precision (“precision”, 1-sigma): 2.10 ppm
- Mean bias (all observations, global offset): 0.01 ppm
- Spatial accuracy (standard deviation site biases): 0.9 ppm
**Figure 3**: Validation of non-glint single soundings of FP-CO₂ with co-located TCCON measurements at all TCCON sites for the period Feb-Oct 2019. Numbers in the figures: \( \mu = \text{bias}, \) i.e., average of the difference; \( \sigma = \text{single measurement precision}, \) i.e., standard deviation of the difference; \( N = \text{number of co-locations}. \)
Figure 4: Validation statistics bias (top) and scatter (bottom) per TCCON site for RemoTeC FP CO2 (bias corrected). The summarizing values (“overall”) represent the standard deviation of the site biases and the average scatter relative to TCCON.
4.3 Bias correction

From comparison with TCCON it was found that the error in XCO₂ correlates with the retrieved albedo \( \alpha \) at 1.6 um in band 2. Based on this correlation the following bias correction has been developed for XCO₂:

\[
X_{\text{CO}_2}^{\text{corr}} = X_{\text{CO}_2} \times (a + b \times \alpha)
\]

with \( a = 0.98997 \), \( b = 0.04581 \).

The bias correction parameters are obtained from fits to the GOSAT-2-TCCON differences.

For sunglint observations there were not enough co-located TCCON observations, therefor the parameter was obtained from comparison to GOSAT-1 data:

\[
X_{\text{CO}_2}^{\text{corr}} = X_{\text{CO}_2} \times (a + b \times RO_2)
\]

With \( a = 1.3822 \), \( b = 0.3912 \) and \( RO_2 = \) retrieved \( O_2 \) ratio.
5. Description of data format

5.1 Product Content and Format

The CH4_GO2_SRFP and CO2_GO2_SRFP data products are stored per day in a single NetCDF file. Retrieval results are provided for the individual GOSAT-2 spatial footprints, i.e. no averaging has been applied. The product file contains the key standard products, 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 products specific for the RemoTeC algorithm.

Table 1: Common dimensions for the CO2_GO2_SRFP product

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Type</th>
<th>Unlimited</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>sounding_dim</td>
<td>int</td>
<td>no</td>
<td></td>
<td>Number of sounding</td>
</tr>
<tr>
<td>polarization_dim</td>
<td>int</td>
<td>no</td>
<td></td>
<td>Number of polarization = 2</td>
</tr>
<tr>
<td>level_dim</td>
<td>int</td>
<td>no</td>
<td></td>
<td>Number of level = 13</td>
</tr>
<tr>
<td>layer_dim</td>
<td>int</td>
<td>no</td>
<td></td>
<td>Number of layer = 12</td>
</tr>
<tr>
<td>window_dim</td>
<td>int</td>
<td>no</td>
<td></td>
<td>Number of retrieval window = 4</td>
</tr>
<tr>
<td>char_l1bname</td>
<td>int</td>
<td>no</td>
<td></td>
<td>Number of character of L1B name = 44</td>
</tr>
</tbody>
</table>

Table 2: Common variables for the CO2_GO2_SRFP product

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Dim.</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>solar_zenith_angle</td>
<td>float</td>
<td>sounding_dim</td>
<td>degrees</td>
<td>Angle between line of sight to the sun and local vertical</td>
</tr>
<tr>
<td>sensor_zenith_angle</td>
<td>float</td>
<td>sounding_dim</td>
<td>degrees</td>
<td>Angle between the line of sight to the sensor and the local vertical</td>
</tr>
</tbody>
</table>
### Table 3: Product specific (additional) variables for the CO2_GO2_SRFP product

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Dim.</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>flag_landtype</td>
<td>int</td>
<td>sounding_dim</td>
<td></td>
<td>0 = land, 1 = ocean</td>
</tr>
<tr>
<td>flag_sunglint</td>
<td>int</td>
<td>sounding_dim</td>
<td></td>
<td>0 = no sunglint, 1 = sunglint</td>
</tr>
<tr>
<td>gain</td>
<td>int</td>
<td>sounding_dim</td>
<td></td>
<td>gain setting of sensor</td>
</tr>
<tr>
<td>exposure_id</td>
<td>int</td>
<td>sounding_dim</td>
<td></td>
<td>Exposure identification number of the sounding</td>
</tr>
<tr>
<td>l1b_name</td>
<td>char</td>
<td>sounding_dim, char_l1bname</td>
<td>Name of the Level 1B file of the sounding</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>-----------------------------</td>
<td>------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>signal_to_noise_window</td>
<td>float</td>
<td>sounding_dim, window_dim, polarization_dim</td>
<td>Signal to noise ratio per retrieval window and for both polarization directions</td>
<td></td>
</tr>
<tr>
<td>dry_airmass_layer</td>
<td>float</td>
<td>sounding_dim, layer_dim</td>
<td>m-2</td>
<td>Dry airmass per layer</td>
</tr>
<tr>
<td>altitude</td>
<td>float</td>
<td>sounding_dim</td>
<td>m</td>
<td>Vertical distance above the surface</td>
</tr>
<tr>
<td>air_temperature</td>
<td>float</td>
<td>sounding_dim, level_dim</td>
<td>K</td>
<td>The bulk temperature of the air at each level</td>
</tr>
<tr>
<td>surface_elevation_stdev</td>
<td>float</td>
<td>sounding_dim</td>
<td>m</td>
<td>Standard deviation of the surface elevation within the sounding</td>
</tr>
<tr>
<td>x_wind</td>
<td>float</td>
<td>sounding_dim, level_dim</td>
<td>m s-1</td>
<td>Eastward wind velocity</td>
</tr>
<tr>
<td>y_wind</td>
<td>float</td>
<td>sounding_dim, level_dim</td>
<td>m s-1</td>
<td>Northward wind velocity</td>
</tr>
<tr>
<td>chi2</td>
<td>float</td>
<td>sounding_dim</td>
<td></td>
<td>Chi-squared value of the sounding</td>
</tr>
<tr>
<td>optical_thickness_of_atmosphere_layer_due_to_ambient_aerosol</td>
<td>float</td>
<td>sounding_dim, window_dim</td>
<td></td>
<td>Scattering optical thickness per retrieval window</td>
</tr>
<tr>
<td>raw_xco2</td>
<td>float</td>
<td>sounding_dim</td>
<td>1e-6</td>
<td>Retrieved column dry-air mole fraction of atmospheric carbon dioxide (XCO2) in ppm before bias correction</td>
</tr>
<tr>
<td>raw_xco2_err</td>
<td>float</td>
<td>sounding_dim</td>
<td>1e-6</td>
<td>1-sigma statistical uncertainty (unscaled) of the retrieved column-average dry-air mole fraction of atmospheric carbon dioxide</td>
</tr>
<tr>
<td>h2o_column</td>
<td>float</td>
<td>sounding_dim</td>
<td>m-2</td>
<td>Retrieved total water column</td>
</tr>
<tr>
<td>surface_albedo_758</td>
<td>float</td>
<td>sounding_dim</td>
<td></td>
<td>The retrieved albedo at 758 nm</td>
</tr>
<tr>
<td>surface_albedo_1593</td>
<td>float</td>
<td>sounding_dim</td>
<td></td>
<td>The retrieved albedo at 1593 nm</td>
</tr>
<tr>
<td>surface_albedo_1629</td>
<td>float</td>
<td>sounding_dim</td>
<td></td>
<td>The retrieved albedo at 1629 nm</td>
</tr>
<tr>
<td>surface_albedo_2042</td>
<td>float</td>
<td>sounding_dim</td>
<td></td>
<td>The retrieved albedo at 2042 nm</td>
</tr>
<tr>
<td>intensity_offset_o2a</td>
<td>float</td>
<td>sounding_dim</td>
<td>W cm-2</td>
<td>The retrieved intensity offset in the O2A band</td>
</tr>
<tr>
<td>Variable</td>
<td>Type</td>
<td>Dimension</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------</td>
<td>-------------</td>
<td>-------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>intensity_offset_band_2</td>
<td>float</td>
<td>sounding_dim</td>
<td>The retrieved intensity offset in spectral window 2</td>
<td></td>
</tr>
<tr>
<td>intensity_offset_band_3</td>
<td>float</td>
<td>sounding_dim</td>
<td>The retrieved intensity offset in spectral window 3</td>
<td></td>
</tr>
<tr>
<td>intensity_offset_band_4</td>
<td>float</td>
<td>sounding_dim</td>
<td>The retrieved intensity offset in spectral window 4</td>
<td></td>
</tr>
<tr>
<td>aerosol_size</td>
<td>float</td>
<td>sounding_dim</td>
<td>Retrieved size parameter of the aerosol distribution</td>
<td></td>
</tr>
<tr>
<td>aerosol_central_height</td>
<td>float</td>
<td>sounding_dim</td>
<td>Peak height of the aerosol Gaussian height distribution</td>
<td></td>
</tr>
<tr>
<td>aerosol_total_column</td>
<td>float</td>
<td>sounding_dim</td>
<td>Retrieved total aerosol column</td>
<td></td>
</tr>
</tbody>
</table>
5.2 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: retrievals for normal or sunglint data, 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':

- Error on retrieved XCO2 < 2.0 ppm*
- \( \chi^2 \) of fit < 4.5
- \( \chi^2 \) of fit in O2A-band < 8**
- SNR > 50
- Standard deviation of surface elevation within GOSAT ground pixel should be < 80 m
- 0 < aerosol_filter < 300
- Aerosol Optical Thickness < 0.6***
- 3 < Aero_size < 5
- SZA < 70°
- 2e-9 < Intensity Offset O2A-band < 5e-9****
- Blended Albedo < 0.9
- 0.99 < CO2 (1.6 micron) / CO2 (2.0 micron) < 1.015
- 0.95 < O2 (retrieved) / O2 (prior) < 1.02
- 0.95 < H2O (1.6 micron) / H2O (2.0 micron) < 1.08

* For the GOSAT-1 CO2_GOS_SRFP product a value of 1.2 is used.
** For the GOSAT-1 CO2_GOS_SRFP product a value of 4 is used
*** For the GOSAT-1 CO2_GOS_SRFP product a value of 0.3 is used.
**** For the GOSAT-1 CO2_GOS_SRFP product values of -1e9 and 3e-9 are used, values are adjusted here to match the distribution found in GOSAT-1.

For **sunglint**:

- Number of iteration steps in retrieval < 31
- Cirrus Signal < 8e-10
- $\chi^2$ of fit < 4
- $\chi^2$ of fit in O2A-band < 4
- $\chi^2$ of fit in window 4 < 10
- -13e-5 < Retrieved slope of albedo in window 4 < -4.5e-5 ^
- 1.5e-9 < Intensity Offset O2A-band < 3.75e-9 ^^
- 0.99 < CO2 (1.6 micron) / CO2 (2.0 micron) < 1.01
- 0.965 < O2 (retrieved) / O2 (prior) < 1.00
- 0.95 < H2O (1.6 micron) / H2O (2.0 micron) < 1.05

*for the GOSAT-1 CO2_GOS_SRFP product values of -1e-5 and 5e-5 are used, values are adjusted here to match the distribution found in GOSAT-1.

**for the GOSAT-1 CO2_GOS_SRFP product values of -1e9 and 3e-9 are used, values are adjusted here to match the distribution found in GOSAT-1.
5.3 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-2-FTS instrument settings (normal, sunglint).

Also, use only data over land (land type=0) except for sunglint cases.

If the data are to be compared with other XCO₂ and/or XCH₄ 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 XCO₂ should be compared to [VCO₂]₀₉₅/VAIR₀₉₅ where [VAIR]₀₉₅ is the total dry air column provided by the model and [VCO₂]₀₉₅ is the model total CO₂ column after applying the column averaging kernel, viz.:

\[ [VCO₂]₀₉₅ = [VCO₂]₉₅ + aᵀ (x_{model} - x_{prior}) \]

where [VCO₂]₀₉₅ is the prior CO₂ total column used in the retrieval, x_{model} is the vertical CO₂ 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 XCH₄ product, but then replacing all instances of CO₂ with CH₄.

5.4 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).
6. References


*** end of document