ECMWF COPERNICUS REPORT

Copernicus Climate Change Service

Algorithm Theoretical Basis Document (ATBD) – ANNEX B for products CO2_GOS_SRFP, CH4_GOS_SRFP (v2.3.8, 2009-2018)

C3S_312b_Lot2_DLR – Atmosphere

Issued by: Lianghai Wu, SRON, The Netherlands

Date: 03/11/2019

Ref: C3S_D312b_Lot2.1.3.2-v1.0_ATBD-GHG_ANNEX-B_v3.1

Official reference number service contract: 2018/C3S_312b_Lot2_DLR/SC1

IMPLEMENTED BY ECMWF

Contributors

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

SRON NETHERLANDS INSTITUTE FOR SPACE RESEARCH, UTRECHT, THE NETHERLANDS
(SRON)
I. Aben
L. Wu
O. P. Hasekamp

Table of Contents

History of modifications

Version	Date	Description of modification	Chapters / Sections
1.1	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: Primarily correction of typos and broken links. Some additional explanations added.	All

Related documents

Reference ID	Document
D1	Main ATBD: Buchwitz, M., et al., Algorithm Theoretical Basis Document (ATBD) – Main document for Greenhouse Gas (GHG: CO_2 & CH_4) data set CDR 3 (2003-2018), project C3S_312b_Lot2_DLR – Atmosphere, v3.1, 2019. *(this document is an ANNEX to the Main ATBD)*

Acronyms

Acronym	Definition
ATBD	Algorithm Theoretical Basis Document
CAMS	Copernicus Atmosphere Monitoring Service
C3S	Copernicus Climate Change Service
CDR	Climate Data Record
CDS	(Copernicus) Climate Data Store
CPU	Core Processing Unit
DEM3	Digital Elevation Map 3
DFS	Degrees of Freedom for Signal
DHF	Data Handling Facility
ECMWF	European Centre for Medium Range Weather Forecasting
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
GHG	GreenHouse Gas
GOSAT	Greenhouse Gases Observing Satellite
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
NASA	National Aeronautics and Space Administration
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
RAA	Relative Azimuth Angle
RMS	Root-Mean-Square
RTM	Radiative transfer model
SNR	Signal-to-Noise Ratio
SRON	SRON Netherlands Institute for Space Research
SRTM	Shuttle Radar Topography Mission
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
TBC	To be confirmed
TBD	To be defined / to be determined
TCCON	Total Carbon Column Observing Network
TIR	Thermal Infra Red
TM5	Transport Model 5
TR	Target Requirements
TRD	Target Requirements Document
VZA	Viewing Zenith Angle

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 an Algorithm Theoretical Basis Document (ATBD) 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 are 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 retrieval algorithms to generate the C3S products CO2_GOS_SRFP and CH4_GOS_SRFP.

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

Executive summary

This document describes the RemoTeC algorithm for GHG retrieval from the GOSAT instrument. The algorithm is based on the paper of *Butz et al., 2009*. Tests of the retrieval algorithm have been performed on synthetic GOSAT data (*Butz et al., 2010*), and real GOSAT data (*Butz et al., 2011; Schepers et al., 2012; Guerlet et al., 2012*).

In order to account for the effect of aerosols and cirrus, the developed algorithm retrieves the methane column simultaneously with the aerosol/cirrus amount (column integrated particle number concentration), a parameter related to the particle size distribution, and a parameter describing the height distribution. Here, the particle size distribution is described by a power-law function, which only has two free parameters (related to amount and size). The choice of aerosol/cirrus parameters reflects the information content of the measurements as closely as possible. The retrieval algorithm uses measured radiances in the Short Wave Infra-Red (SWIR) band and additionally in the Near Infra Red (NIR, O_2 A-band). Additional fit parameters are the surface albedo and its 1st order spectral dependence in all bands, and the total column of water vapor, respectively.

In order to obtain a proper characterization of the retrieved XGHG, it is important to first retrieve a vertical profile (layer averaged number density in different layers of the model atmosphere) and use this retrieved vertical profile to calculate the vertical column. Here, we choose to provide the vertical column as a product, and not the full profile, because the Degrees of Freedom for Signal (DFS) of the retrieved CH_4 and CO_2 profile is about 1. The inversion is performed using Phillips-Tikhonov regularization in combination with a reduced step size Gauss-Newton iteration scheme.

The forward model of the retrieval algorithm uses online radiative transfer calculations, fully including multiple scattering. Here, the radiative transfer model developed by *Landgraf et al., 2001*, and *Hasekamp and Landgraf 2002, 2005* is being used. This model uses the Gauss-Seidel iterative method to solve the radiative transfer equation in a plane-parallel, vertically inhomogeneous atmosphere. To avoid time consuming line-by-line calculations we employ the linear-k method developed by *Hasekamp and Butz 2008*. Absorption cross sections of the relevant atmospheric trace gases are tabulated in a lookup table as function of pressure and temperature. Optical properties of aerosols are also calculated from lookup tables as described in *Dubovik et al. 2006.* The linear k-binning method in combination with other speed optimizations allow us to perform the GHG retrievals from GOSAT with online RT calculations within ~45 seconds for a single retrieval. This makes the algorithm feasible for global processing

1. Data product overview

1.1 Column-averaged mixing ratios of CH_4 (XCH_4) and CO_2 (XCO_2)

In this section an overview of the data product (specified in terms of variable, its property, processing level(s) and instrument(s)) is given.

The data products

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

In the following, several satellite instruments are shortly described which are used / can be used to generate the XCH_4 and XCO_2 data products CH4_GOS_SRFP and CO2_GOS_SRFP.

TANSO-FTS is a Fourier-Transform-Spectrometer (FTS) onboard the Japanese GOSAT satellite (*Kuze et al., 2009, 2014, 2016*). The Greenhouse Gases Observing Satellite "IBUKI" (GOSAT) is the world's first spacecraft in orbit dedicated to measure the concentrations of carbon dioxide and methane from space. The spacecraft was launched successfully on January 23, 2009, and has been operating properly since then. GOSAT covers the relevant CO_2, CH_4 and O_2 absorption bands in the NIR and SWIR spectral region as needed for accurate XCO_2 and XCH_4 retrieval (in addition GOSAT also covers a large part of the Thermal Infrared (TIR) spectral region). The spectral resolution of TANSO-FTS is much higher compared to SCIAMACHY and also the ground pixels are smaller (10 km compared to several 10 km for SCIAMACHY). However, in contrast to SCIAMACHY, the GOSAT scan pattern consists of non-consecutive individual ground pixels, i.e., the scan pattern is not gap-free.
For a good general overview about GOSAT see also http://www.gosat.nies.go.jp/en/. As a successor of GOSAT, GOSAT-2 was successfully launched on October 29, 2018, which provides us another instrument to monitor global greenhouse gas concentrations. More information on GOSAT-2 can be found from http://www.gosat-2.nies.go.jp/.

2. Input and auxiliary data

2.1 Satellite instrument

2.1.1 GOSAT TANSO-FTS level 1b

Version V201201 / 201202 of the Level 1b data of the TANSO-FTS (Thermal And Near-infrared Sensor for carbon Observation -Fourier Transform Spectrometer) onboard GOSAT (Greenhouse gas Observing SATtellite) are needed in the project to produce the total column CO2 and CH4 products. They serve as input for the retrieval algorithms to be used in this project.

Originating system:

The GOSAT satellite was launched in January 2009. The GOSAT observational data are routinely processed at the GOSAT Data Handling Facility. The development of the GOSAT Data Handling Facility (GOSAT DHF) was completed in late 2008, and NIES has been maintaining it for the routine processing of the GOSAT data

Data class: Earth Observation

Sensor type and key technical characteristics:

TANSO-FTS is an instrument that utilizes optical interference. Within the instrument the incoming light is split into two beams which propagate in separate optical paths to create an optical path difference between the two. These beams are then recombined to cause interference. FTS measures the intensity of the interference by continuously changing the optical path difference. A spectrum, which is distribution of light intensity over a span of wavelengths, is obtained by performing a mathematical operation called the Fourier transform on that measured data.

FTS observes sunlight reflected from the earth's surface and light emitted from the atmosphere and the surface. The former is observed in the spectral bands 1 through 3 of FTS in the daytime, and the latter is captured in band 4 during both the day and the night. Within this project only level 1 data from the SWIR channels 1-3 is used. Prior to reaching the detectors of the instrument, the light in the bands 1 through 3 is split into two orthogonally-polarized components (P and S components) and measured independently. The light in the band 4, however, is not split. The instrument thereby observes the incoming light in seven different channels.

The TANSO-FTS Level 1b data are radiance spectra that are obtained by performing the Fourier transformation on the signals detected by FTS. A single data file of the FTS Level 1B data contains the radiance spectra obtained during 1/60 of an orbital revolution (defined as "one scene").

Table 1: GOSAT-FTS bands

channel	wavelength range [nm]	resolution [cm^{-1}]
1	758-775	0.2
2	1460-1720	0.2
3	1920-2080	0.2
4	5560-14300	0.2

Data availability & coverage:

TANSO-FTS level-1b data are available since April 2009. Over a three-day period, GOSAT-FTS takes fifty-six thousand measurements, covering the entire globe.

Source data product name & reference to product technical specification documents

GOSAT Level 1 Product Description Document, Japan Aerospace Exploration Agency, March 2010, NEB-080035D, available through http://www.gosat.nies.go.jp/en/.

Data quality and reliability

The quality of the retrieved CO2 and CH4 columns has been tested against ground-based observations (i.e. the Total Carbon Column Observing Network (TCCON)network, http://www.tccon.caltech.edu/) and has shown to be of good quality. The L1B data are updated yearly with a new version to improve calibration.

2.2 Other

2.2.1 ECMWF model data

The retrieval algorithms to produce vertical columns of CO_2 and CH_4 need as input for each scene the temperature vertical profile, pressure vertical profile, specific humidity vertical profile, and wind speed. Here, temperature and pressure are needed to calculate absorption cross sections, the specific humidity vertical profile is needed to account for water vapor absorption, and the wind speed is needed to calculate the Fresnel reflection contribution on a rough ocean surface. The meteorological data mentioned above will be taken from the ECMWF model.

Originating system:

ECMWF has developed one of the most comprehensive earth-system models available anywhere. The ECMWF model is uses the '4D-Var' data assimilation approach, which provides a physically consistent best fit to observations. For this project the ERA-interim archive as well as the operational archive is of importance. Here, the ERA interim data will be used to produce the C3S time series as this dataset uses one model version for the entire period. For sensitivity studies it is planned to use one year of data from the operational archive.

Data class: Model

Required ECMWF data:

Class: ERA interim

Stream: Atmospheric model

Type: Analysis

Dates: 01/04/2009 to 01/01/2019

Time: 00:00:00, 06:00:00, 12:00:00, 18:00:00

Spatial grid: N128 Quasi-regular Gaussian grid (~0.7^{o})

Parameters at model levels:

- temperature, specific humidity (all levels)
- logarithm of surface pressure, geopotential (lowest level)

Parameters at surface:

- 10 metre U wind component
- 10 metre V wind component

Data availability & coverage:

All data are required on a global scale, with a typical delay of three months.

Data quality and reliability:

The ECMWF model data sets are considered to be among the best available data sets for meteorological parameters.

2.2.2 Carbon Tracker Data

The retrieval algorithms for CH4 columns from SCIAMACHY and GOSAT that are based on the "proxy approach" retrieve the ratio of the CH4 and CO2 columns, where the CO2 column serves as a proxy for the light path. In order to obtain the CH4 column, the retrieved ratio needs to be multiplied by the best estimate of the CO2 column. It is assumed that data from the Carbon Tracker model can provide the best estimate of the true CO2 column.

Originating system:

Carbon Tracker is a data assimilation system developed at the National Oceanic and Atmospheric Administration's (NOAA's) Earth System Research Laboratory (ESRL). It is based on a state-of-the-art atmospheric transport model coupled to an ensemble Kalman filter. CarbonTracker assimilates atmospheric CO2 mole fractions, using a variety of in situ measurements worldwide.

*Data class***:** Model

*Data availability & coverage***:**

CarbonTracker provides global daily data at a spatial resolution of 3 by 2 degree. Data are available with a delay typically of one or more years.

Source data product name & reference to product technical specification documents

CarbonTracker 3D Mole Fractions

Peters, W. et al. 2007

Data quality and reliability

CarbonTracker has been validated with many independent measurements and is considered to provide be an accurate global CO2 data set.

2.2.3 LMD inversions

In order to obtain the CH_4 column, the retrieved ratio needs to be multiplied by the best estimate of the CO_2 column. Instead of using CO2 column values from CarbonTracker, we can use the data from the LMD CO_2 flask inversions as alternative choice. It is assumed that data from the LMD flask inversion can provide the best up-to-date estimate of the true CO_2 column.

Originating system:

The inversion product used here is the official CAMS v16r1 product that exclusively assimilates about 130 sites of surface air sample measurements from the Global Atmosphere Watch programme

Data class: Model

Data availability & coverage:

The CAMS v16r1 inversion product provides global 3-hourly data at a spatial resolution of 3.75 by 1.89 degree. Data are typically available within one year.

Source data product name & reference to product technical specification documents

Chevalier et al 2010.

Data quality and reliability

The CAMS v16r1 inversion product has been validated with many independent measurements and is considered to provide be an accurate global CO2 data set.

2.2.4 SRTM DEM

The RemoTeC retrieval algorithm for CO_2 and CH_4 columns from GOSAT use information about the surface elevation from an extended SRTM digital elevation map.

Originating system: The original Shuttle Radar Telemetry Mission (SRTM) was provided by the United States National Aeronautics and Space Administration (NASA). The dataset used (DEM3) is based on the SRTM dataset and includes extrapolation and gap filling from various sources.

Data class: Model

Sensor type and key technical characteristics:

n/a

Data availability & coverage:

The original SRTM dataset provides elevation data ranging from 56 degrees south to 60 degrees north at a 90 meter resolution. The adjusted DEM3 dataset extends the coverage, while keeping the 90 meter resolution.

Source data product name & reference to product technical specification documents:

http://www.viewfinderpanoramas.org/dem3.html

2.2.5 TCCON FTS CO_2 and CH_4 data

TCCON data for CO_2 and CH_4 is available publically for all TCCON stations (https://tccon-wiki.caltech.edu/). We use the GGG2014 official release of the data product.

Originating system: Ground based

Data class: Ground based

Sensor type and key technical characteristics:

The measurements are performed using the solar absorption spectroscopy in the near infrared using a Fourier Transform Spectrometer (FTS).

Data availability & coverage:

Coverage is limited to the locations of the TCCON stations themselves. Depending on the instrument setting (gain H, gain M or sunglint), we use different TCCON stations for validaton (see Product User Guide). Data is typically available only after one year, although some stations deliver on a more regular interval (3 months).

Source data product name & reference to product technical specification documents:

Wunch et al., 2011

Data quantity:

Individual measurements can be taken in intervals of about 20 min. The observations can only be taken with the direct sunlight. This limits the amount of data, which is different from site to site.

Data quality and reliability:

For XCO_2 the precision is 0.25% (1ppm) and the systematic error (bias) is 0.2% (0.8 ppm). For XCH_4 the precision is 0.40% (7ppb) and the systematic error (bias) is also 0.40% (7 ppb).

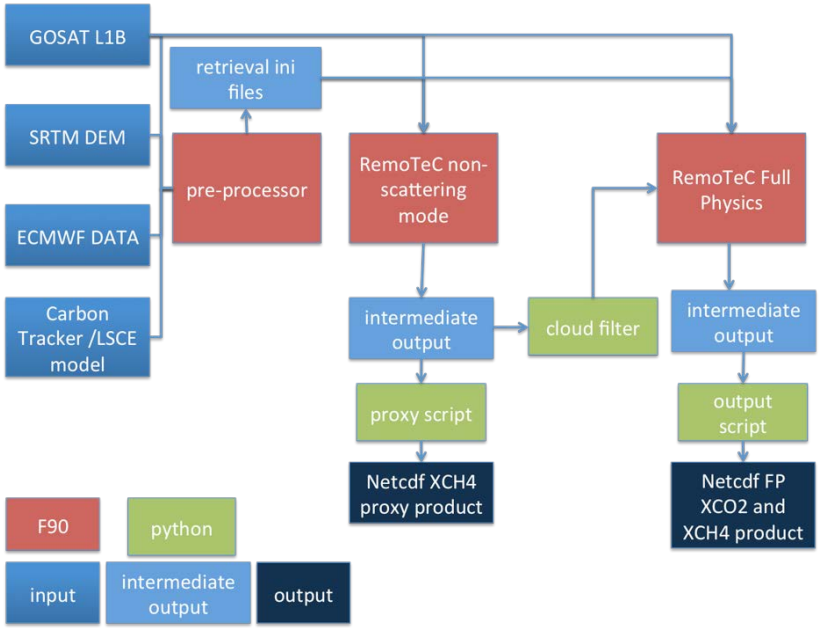
2.2.6 Additional input data

- Absorption cross sections: For the retrieval lookup-tables are used with pre-calculated absorption cross sections of the species of interest (O_2, CO_2, CH_4, H_2O) as a function of wavenumber, temperature, and pressure. One lookup-table per species and per spectral band is being used. At the start of processing at a given CPU (Fig. 3) the cross-section lookup table is read into memory.
- Aerosol optical properties: A lookup table is being used with pre-calculated aerosol optical properties (Mie and t-Matrix theory) as a function of size parameter and refractive index.
- Retrieval settings: A file is read in with retrieval settings such as fit parameters, spectral range, etc.

2.3 Overview of Processing Sub-System

Figure 1 provides a schematic overview of the RemoteC GHG-CCI processing sub system at SRON The first step is to download the required data from the respective data servers to SRON (GOSAT data and ECMWF data are dynamic datasets that are continuously updated, SRTM topography is a static dataset). In the next step a pre-processing program is combining all relevant information per GOSAT ground pixel. This includes interpolation of ECMWF data in space and time to the coordinates of the GOSAT ground pixel, calculating the average height of a GOSAT ground pixel and its standard deviation from the topography database.

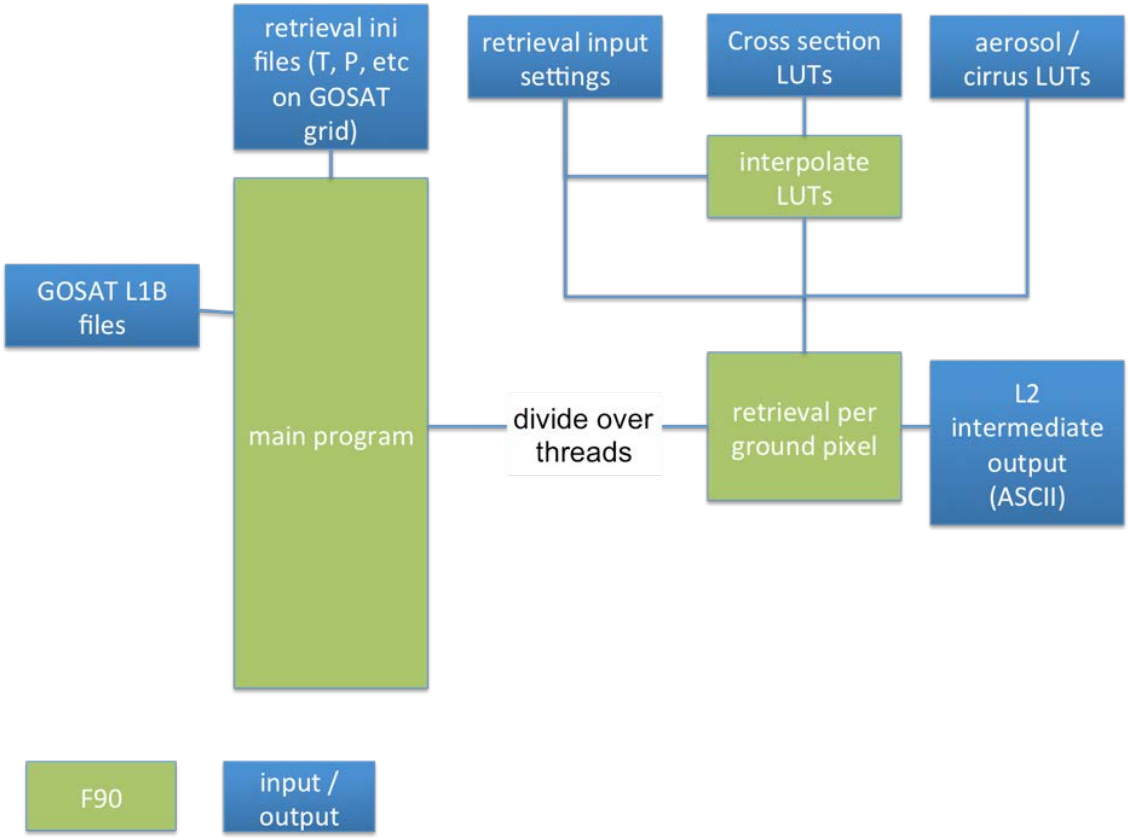
Figure 1: Schematic overview of the RemoteC algorithm processing sub-system.



The pre-processor produces for each GOSAT L1B file an auxiliary input file, hereafter referred to as the 'retrieval ini ' file, that contains this information. In the next step columns of CO_2, CH_4, H_2O, and O_2 are retrieved under the assumption of an atmosphere without aerosol/cirrus/cloud scattering ('RemoTeC non scattering mode' in Figure 1). The outcome of these retrievals (in an intermediate ACII output file) is used to create the RemoTeC XCH_4 proxy product, but also for cloud filtering to select cloud free scenes to be processed by the RemoTeC Full Physics algorithm. The Full Physics retrieval produces intermediate (ASCII) output files which go into an *a posterior* filtering procedure, quality check (based on non-convergence, parameter boundary hits, retrieved aerosol parameters), and bias correction and finally a Netcdf output file is created.

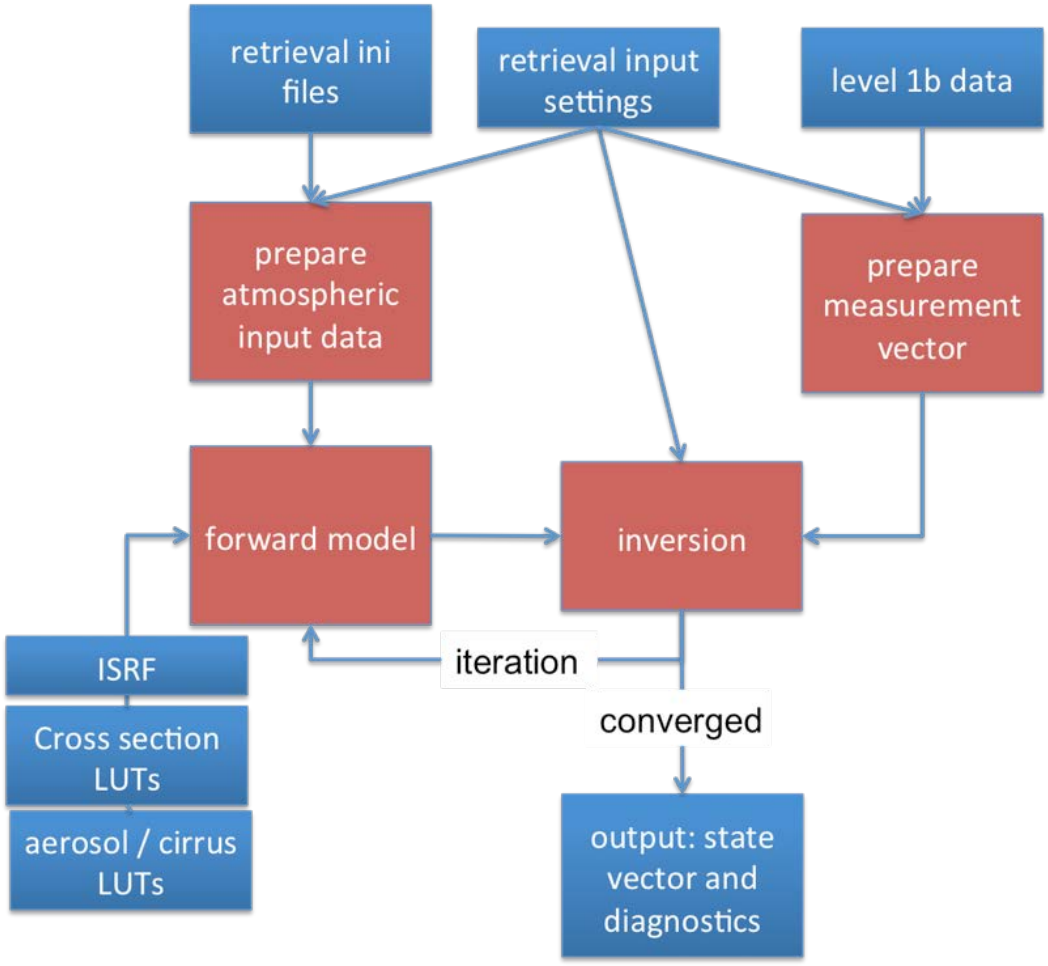
Figure 2 gives an schematic overview of the core RemoTeC retrieval algorithm (same for non-scattering and Full Physics). Here, multi-threading capability is implemented using *openMP,* where different ground pixels are divided over multiple threads. Figure 2 shows the processing per ground pixel (i.e. for a single thread) in more detail.

Figure 2: Schematic overview of the RemoTeC retrieval procedure including multi-threading.



The static input that is required is a lookup table with the relevant absorption cross sections (read into memory at beginning of processing), a lookup table with aerosol optical properties, and a file indicating the retrieval settings (i.e. fit parameters, spectral range, etc.). Further, the auxiliary retrieval input files that are produced by the pre-processor are needed for each GOSAT ground pixel to be processed, together with the GOSAT-FTS level 1 data. The retrieval per pixel is then run (iterative scheme with forward model and inversion module) and after convergence an intermediate (ASCII) output file is created that is used in the a posteriori filtering and quality check (see Figure 3) and the processing of the next ground pixel starts.

Figure 3: Overview of RemoTeC processing per ground pixel



3. Algorithms

Any retrieval algorithm aims at inferring an atmospheric state vector **x** from a measurement vector **y**. The state vector is linked to the measurement vector through the true forward model **f**(**x**,**b**) that depends on the state vector **x** and the vector **b** containing ancillary parameters that are not retrieved,

$$\mathbf{y} = \mathbf{f}(\mathbf{x}, \mathbf{b}) + \mathbf{e}_y \tag{1}$$

where $\mathbf{e}_y$ represents the measurement noise vector. A retrieval method approximates the true forward model **f** by a retrieval forward model **F**, with a forward model error vector $\mathbf{e}_F$,

$$\mathbf{y} = \mathbf{F}(x, b) + \mathbf{e}_\mathbf{y} + \mathbf{e}_\mathbf{F} \tag{2}$$

For full physics retrieval from the GOSAT FTS instrument the measurement vector contains the measured intensities in the NIR and SWIR (see Table 2).
Table 2: Spectral ranges from the NIR and SWIR band included in the measurement vector

band	used spectral range
1 (O_2 a)	12920 – 13195 cm^{-1}
2 (CO_2)	6170 – 6277 cm^{-1}
3 (CH_4)	6045 – 6138 cm^{-1}
4 (CO_2)	4806 – 4896 cm^{-1}

For the retrieval procedure it is needed that the non-linear forward model is linearized so that the retrieval problem can be solved iteratively. For iteration step n the forward model is approximated by

$$\mathbf{F}(\mathbf{x}, \mathbf{b}) \approx \mathbf{F}(\mathbf{x}_n, \boldsymbol{b}) + \mathbf{K}(\mathbf{x}_n - \mathbf{x}) \tag{3}$$

where x_n is the state vector for the n-th iteration step and K is the Jacobian matrix

$$\mathbf{K} = \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \tag{4}$$

Below, we will describe the retrieval forward model, state vector, ancillary parameter vector, and the inversion method in more detail.

3.1 Forward Model

The retrieval forward model **F** simulates the measurement vector y for a given model atmosphere defined by the state vector **x** and the ancillary parameter vector **b**. The simulated intensity for a given spectral pixel *i* is given by

$$I_i = \int_{\lambda_{min}}^{\lambda_{max}} I(\lambda) S_i(\lambda) d\lambda \quad (5)$$

where $S_i(\lambda)$ is the Instrument Spectral Response Function (ISRF) for spectral pixel *I* and $I(\lambda)$ is the modeled intensity at high spectral resolution. In the NIR and SWIR channel $I(\lambda)$ contains many fine spectral structures due to molecular absorption, so it has to be calculated at fine spectral resolution (0.1 cm^{-1} in the NIR band and 0.02 cm^{-1} in the SWIR).

3.1.1 Model Atmosphere and Optical Properties

For the RemoTeC algorithm described here the model atmosphere is defined for NLAY=36 homogeneous vertical layers that are equidistant in pressure, the lowest pressure level being defined by the surface pressure. The absorbing trace gases of interest are O_2 (in the NIR band) and CH_4, H_2O, and CO in the SWIR band. The layer sub-columns of these gases are for the first iteration step of each retrieval calculated from the input profiles of CH_4, CO (TM5) and H_2O (ECMWF) and the temperature and pressure profiles (ECMWF). They are obtained on the grid of the model atmosphere by linear interpolation. Here, first the surface pressure p_{surf} is obtained by interpolation of the input pressure profile as function of height to the surface height (input) for the corresponding ground pixel. Next the pressure values at the layer boundaries are calculated, with the pressure p_k at the lower boundary of layer k (counting from top to bottom) is given by:

$$p_{lev,k} = p_{min} + \Delta p * k \quad (6)$$

$$\Delta p = (p_{surf} - p_{min})/NLAY$$

where p_{min} is the pressure value of the upper boundary of the input (ECMWF) atmosphere. The different atmospheric profiles are constructed on this pressure grid. For example, the methane sub-column DV_CH4_k for the layer bounded by pressure levels $p_{lev,k}$ and $p_{lev,k+1}$ is given by:

$$DV\_CH4_k = XCH4_k \times DV\_AIR_k \quad (7)$$

where $XCH4_k$ is the methane dry air mixing ratio linearly interpolated from the input pressure grid to the pressure at the 'middle' of layer k defined by $(p_k+p_{k+1})/2$. DV_AIR_k is the sub-column of air in layer *k*, given by

$$DV\_AIR_k = (p_{lev,k+1} - p_{lev,k}) \times R/(M \times g_k \times \left(1 + \frac{XH_2O_k}{1.60855}\right)) \quad (8)$$

where R is Avogadro's number, M is the molecular mass of air, g_k is the gravity constant in altitude layer k, and 1.60855 is the mass of air relative to the mass of water (*Wunch et al., 2010*). The sub columns of CO and H_2O are calculated in the same manner as for CH_4, and the O_2 sub-column is obtained by multiplying the air sub-column by the O_2 mixing ratio (=0.2095).

For a radiative transfer calculation at a given wavelength the layer absorption optical thickness, scattering optical thickness, and scattering phase function for each model layer are needed. For layer *k* of the model atmosphere the CH_4 absorption optical thickness at wavelength λ_j is calculated by:

$$\tau_{abs,ch4}(\lambda_j) = \sum_{i=1}^{N} \sigma(p_i, T_i, \lambda_j) * DV\_CH4_k / N \quad (9)$$

where N is the number of sub-layers in which the model atmosphere layers are divided (set at N=2) and σ is the absorption cross-section of CH_4 at wavelength λ_j and the pressure p_i and temperature T_i for the middle of model sub-layer *i*. The absorption optical thickness for the other trace gases is calculated in the same way. Pre-calculated absorption cross-sections for CH_4, CO, H_2O, and O_2 are stored as lookup-table as a function of pressure, temperature, and wavenumber. These cross-section lookup-tables are calculated from the latest spectroscopic databases (*Jenouvrier et al., 2007; Rothman et al., 2009*) assuming Voigt line shapes. For water vapor, an updated spectroscopic line list has been developed by *Scheepmaker et al. 2012*. In the NIR spectral range the absorption cross sections of O_2 are calculated according to *Tran et al. 2006* taking into account line mixing and collision induced absorption. The cross-section for pressure p_i, temperature T_i and wavelength λ_j are obtained by linear interpolation from the tabulated values. The reason that each of the NLAY layers of the model atmosphere is further divided into N sub-layers, is to properly account for the strong dependence of temperature and pressure of the absorption cross-sections. So, the absorption cross-sections are needed for NLAY X N = 36 X 2=72 vertical layers equidistant in pressure.

The total molecular absorption optical thickness is obtained by summing the contribution of the different trace gases. The Rayleigh scattering optical thickness for layer k and wavelength λ_j is given by

$$\tau_{ray,k}(\lambda_j) = \sigma_{ray}(\lambda_j) * DV\_AIR_k \quad (10)$$

where σ_{ray} is the Rayleigh scattering cross section given by (*Buchholtz,1995*)

$$\sigma_{ray}(\lambda) = A\lambda^{-(4+X)} \quad (11)$$
$$X = B\lambda + C/\lambda - D$$

with A= 4.02E-28, B=0.389, C=0.04926, D=0.3228.

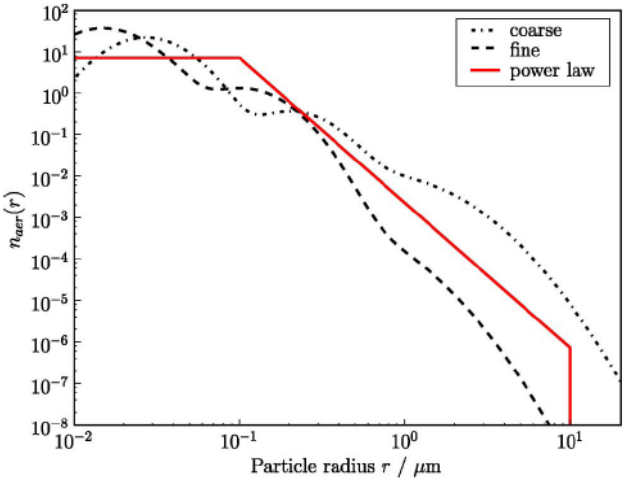
The Rayleigh scattering phase function is given by (e.g. *Hansen and Travis, 1974*)

$$P(\Theta) = \frac{3}{4}(1+\cos^2\Theta)\frac{(1-\delta)}{(1+\delta/2)} \quad (12)$$

where δ is the depolarization ratio and Θ is the scattering angle defined by

$$\cos(\Theta) = -u_0 u_v + \sqrt{(1-u_v)^2}\sqrt{(1-u_{v0})^2}\cos(\varphi_0 - \varphi_v) \quad (13)$$

Figure 4: Aerosol size distribution $n_{aer}(r)$ as a function of particle radius r. The retrieval method relies on a power law (red solid) size distribution. Also shown are more realistic Multi-modal lognormal size distributions for a fine mode (black dashed) and a coarse mode (black dotted) dominated aerosol type.



where u_0 and u_v are the cosines of the solar and viewing zenith angle, respectively (absolute values) and φ_0 and φ_v are the solar and viewing azimuth angles.

In addition to trace gases, also aerosols are present in the model atmosphere. In our algorithm, aerosols are described by the following parameters (following *Butz et al. (2009;2010)*):

1. Number of particles in each layer of the model atmosphere. This is provided by the total amount of particles, N_{aer}, and a normalized altitude distribution described by a Gaussian function of center height z_{aer} and width w_0. Hence, in model layer k with central height z_k, the number of particles is given by $h(z_k) = N_{aer} \exp[-\frac{4\ln(z_k - z_{aer})^2}{(2w_o)^2}]$
2. A size distribution of $n_{aer}(r)$, described by a power law size function, characterized by power law exponent and upper and lower cut-off (e.g. *Mishchenko et al., 1999*).

$$n(r) = \begin{cases} A, & r \le r_1 \\ A(r/r_1)^{-\alpha} & r_1 < r \le r_2 \\ 0 & r > r_2 \end{cases} \qquad (14)$$

The cut-offs are r_1,=0.1μm, r_2 = 10 μm and the constant A is determined from normalization of the size distribution. Figure 4 illustrates $n_{aer}(r)$ and compares it to a more realistic multimodal lognormal size distribution (*Stier et al., 2005*). Through its parameter α, the aerosol size distribution controls the spectral dependence of aerosol optical properties among the considered retrieval windows.

3. The complex refractive index m =m_r+*i*m_i, which is assumed independent of wavelength within a retrieval window.

From the aerosol size distribution, refractive index, and number of particles of each layer the aerosol scattering optical thickness $\tau_{scat,aer}$ and aerosol absorption optical thickness $\tau_{abs,aer}$ are calculated

$$\tau_{scat,aer}(z_k) = \sigma_{scat,aer} h(z_k)$$
$$\tau_{abs,aer}(z_k) = \sigma_{abs,aer} h(z_k) \quad (15)$$

where $\sigma_{scat,aer}$ and $\sigma_{abs,aer}$ are the aerosol scattering and absorption cross-section, respectively. They are obtained by:

$$\sigma_{scat,aer} = \sum_{i=1}^{M} K_{scat,i}(m) r_i n_{aer}(r_i) v(r_i)$$
$$K_{scat,i} = \int_{\Delta \ln r_i} \frac{\sigma_{scat}(r,m)}{v(r)} d\ln r \quad (16)$$

where *v* denotes particle volume and $K_{scat,i}$,representative for particle radius r_i, is pre-calculated for M size bins according *to Dubovik et al., 2006* and stored in a lookup table as function of aerosol size parameter x=2πr/λ, real refractive index, and imaginary refractive index. The values for the actual aerosol refractive index are obtained by linear interpolation from the tabulated values. The lookup table contains values for spheres (Mie calculations) as well as for spheroids with a pre-defined axis ratio distribution (*Dubovik et al, 2006*) but in our algorithm baseline we only consider spherical particles. Similar expressions hold for the absorption cross-section and the aerosol scattering phase function.

Finally, the total optical properties per layer in the model atmosphere are obtained by combining the contribution of gases and aerosols:

$$\tau_{abs} = \tau_{abs,mol} + \tau_{abs,aer}$$
$$\tau_{sca} = \tau_{sca,mol} + \tau_{sca,aer}$$
$$P(\Theta) = \frac{\tau_{sca,mol} P_{mol}(\Theta) + \tau_{sca,aer} P_{aer}(\Theta)}{\tau_{sca,mol} + \tau_{sca,aer}} \quad (17)$$

For multiple scattering calculations the scattering phase function is needed in the form of expansion coefficients for generalized spherical functions, where expansion coefficient α_l with index *l* is given by:

$$\alpha^l = \frac{2l+1}{2}\int_{-1}^{1} P_{0,0}^l(cos\Theta)P(\Theta)dcos\Theta \quad (18)$$

Where $P_{0,0}^l$ is element (1,1) of the Generalized Spherical Function matrix (e.g. *de Haan et al., 1987*).

To summarize, the forward model needs the following inputs:

- Surface pressure to define the equidistant pressure grid
- Sub-columns of CH_4, CO, H_2O, O_2, and air for the vertical layers of the model atmosphere.
- Pressure and temperature at the middle of the model sub-layers for absorption cross-sections.
- The aerosol column N_{aer}.
- The aerosol size parameter α (power of the power law size distribution function).
- The central height z_{aer} and width w_0 of the Gaussian aerosol altitude distribution.
- Solar Zenith Angle (SZA).
- Viewing Zenith Angle (VZA).
- Relative Azimuth Angle (RAA).
- The aerosol complex refractive index $m = m_r + im_i$
- A high spectral resolution solar reference spectrum.
- Lookup tables with absorption cross-sections of CH_4, CO, H_2O, and O_2 as function of pressure, temperature, and wavenumber.
- Lookup tables with pre-calculated aerosol properties as function of aerosol size parameter, real refractive index, and imaginary refractive index (cross sections and phase functions integrated over each size parameter bin).

Based on these inputs the optical properties can be calculated for each layer of the model atmosphere.

3.1.2 Modeling the top-of-atmosphere radiances

Based on the optical properties (τ_{abs}, τ_{sca}, $P(\Theta)$) defined for each wavelength and layer of the model atmosphere we can compute the top-of-atmosphere radiance as measured by the instrument. The first step is to separate the radiation field in a singly scattered component $\mathbf{I}_{ss}$ and a multiply scattered component $\mathbf{I}_{ms}$, respectively:

$$I = I_{ss} + I_{ms} \tag{19}$$

The computation of I_{ss} for a given wavelength is straightforward:

$$\begin{aligned} I_{ss} = F_0 \sum_{k=1}^{NLAY} \omega_k P_k(\Theta) \left[1 - e^{-(\tau_{tot,k}\left(\frac{1}{u_0}+\frac{1}{u_v}\right)}\right] u_0 \\ /4\pi(u_o + u_v) e^{-(\sum_{i=1}^{k} \tau_{tot,i}/u_0)} \, e^{-(\sum_{i=1}^{k} \tau_{tot,i}/u_v)} \\ + e^{-(\sum_{k=1}^{NLAY} \tau_{tot,i}/u_0)} e^{-(\sum_{k=1}^{NLAY} \tau_{tot,i}/u_{0v})} R_{surf} \end{aligned} \tag{20}$$

where F_0 is the incoming total flux, $\tau_{tot} = \tau_{abs}+\tau_{sca}$, $\omega = \tau_{sca}/\tau_{tot}$, u_0 is the cosine of the solar zenith angle, u_v is the cosine of the viewing zenith angle, and R_{surf} is the surface reflection for the specific solar and viewing geometry under consideration. Besides the I_{ss} itself, also the derivatives with respect to $\tau_{sca,k}$, $\tau_{abs,k}$ ω_k and P_k are needed:

$$\begin{aligned} \frac{\partial I_{ss}}{\partial \tau_{tot,k}} = F_0 \omega_k P_k(\Theta) \left[e^{-(\tau_{tot,k}\left(\frac{1}{u_0}+\frac{1}{u_v}\right)}\right] u_0 / 4\pi(u_o + u_v) e^{-(\sum_{i=1}^{k} \tau_{tot,i}/u_0)} e^{-(\sum_{i=1}^{k} \tau_{tot,i}/u_v)} \\ - \sum_{i=1}^{K} \omega_k P_k(\Theta) \left[1 - e^{-(\tau_{tot,k}\left(\frac{1}{u_0}+\frac{1}{u_v}\right)}\right] u_0 \\ /4\pi(u_o + u_v) e^{-(\sum_{i=1}^{k} \tau_{tot,i}/u_0)} \, e^{-(\sum_{i=1}^{k} \tau_{tot,i}/u_v)} \end{aligned} \tag{21}$$

$$\begin{aligned} \frac{\partial I_{ss}}{\partial P_k} = F_0 u_0 / 4\pi(u_o + u_v) e^{-(\sum_{i=1}^{k} \tau_{tot,i}/u_0)} e^{-(\sum_{i=1}^{k} \tau_{tot,i}/u_v)} \left[1 \right. \\ \left. - e^{-(\tau_{tot,k}\left(\frac{1}{u_0}+\frac{1}{u_v}\right)}\right] \omega_k \end{aligned} \tag{22}$$

$$\frac{\partial I_{ss}}{\partial \omega_k} = F_0 u_0 / 4\pi (u_o + u_v) e^{-(\sum_{i=1}^{k} \tau_{tot,i}/u_0)} e^{-(\sum_{i=1}^{k} \tau_{tot,i}/u_v)} P_k(\Theta) \left[e^{-(\tau_{tot,k}(\frac{1}{u_0}+\frac{1}{u_v})} \right] \quad (23)$$

$$\frac{\partial I_{ss}}{\partial \tau_{sca,k}} = \frac{\partial I_{ss}}{\partial \omega_k} \frac{1}{\omega_k} - \frac{\partial I_{ss}}{\partial \tau_{tot,k}} \quad (24)$$

$$\frac{\partial I_{ss}}{\partial \tau_{abs,k}} = \frac{\partial I_{ss}}{\partial \tau_{tot,k}} - \frac{\partial I_{ss}}{\partial \omega_k} \omega_k / \tau_{tot,k} \quad (25)$$

The computation of the multiply scattered radiation for a given wavelength involves the solution of the plane-parallel radiative transfer equation. We solve this equation using the Gauss-Seidel iterative method. The solution is described in detail by *Landgraf et al., 2001* and will not be repeated here. The calculation of the derivatives with respect to optical properties per layer *k* of the model atmosphere is performed using the forward adjoint perturbation theory and is described in detail by *Hasekamp and Landgraf 2005a*.

The radiative transfer calculations yield the derivatives of the radiance with respect to the optical parameters τ_{sca}, τ_{abs}, and P or its expansion coefficients α_l. From these derivatives, the derivatives with respect to a physical parameter x can be calculated in a straightforward manner by the derivative chain rule

$$\frac{\partial I_{ss}}{\partial x} = \sum_{k=1}^{NLAY} \frac{\partial I_{ss}}{\partial \tau_{sca,k}} \frac{\partial \tau_{sca,k}}{\partial x} + \frac{\partial I_{ss}}{\partial \tau_{abs,k}} \frac{\partial \tau_{abs,k}}{\partial x} + \frac{\partial I_{ss}}{\partial P_k} \frac{\partial P_k}{\partial x} \quad (26)$$

$$\frac{\partial I_{ms}}{\partial x} = \sum_{k=1}^{NLAY} \frac{\partial I_{ms}}{\partial \tau_{sca,k}} \frac{\partial \tau_{sca,k}}{\partial x} + \frac{\partial I_{ms}}{\partial \tau_{abs,k}} \frac{\partial \tau_{abs,k}}{\partial x} + \sum_{l=0}^{M} \frac{\partial I_{ms}}{\partial \alpha_k^l} \frac{\partial \alpha_k^l}{\partial x} \quad (27)$$

In order to avoid time consuming multiple scattering calculations on the high spectral resolution line-by-line grid we aim at reducing the number of spectral calculations, following the approach of *Hasekamp and Butz 2008*. For this purpose, we consider the intensity $\mathbf{I}_{ms}$ as a function of total absorption optical thickness τ_{abs} and its normalized vertical distribution **n** (similar to k-distribution and spectral mapping methods),

$$I_{ms}(\lambda) = I_{ms}(\tau_{abs}(\lambda), \boldsymbol{n}(\lambda)) \tag{28}$$

where we assume that the atmospheric scattering properties and surface reflection properties are constant over the spectral range under consideration. Here, the explicit separation between total absorption optical thickness and its vertical distribution is chosen for later convenience. Element n_k of the vector **n** represents the relative contribution of the absorption optical thickness of altitude layer *k* of the model atmosphere to the total absorption optical thickness of the atmosphere, such that

$$\tau_{abs}^{k}(\lambda) = n_k(\lambda)\tau_{abs}(\lambda) \tag{29}$$

where τ^k_{abs} is the absorption optical thickness of layer *k* of the model atmosphere.
For a vertically homogeneous atmosphere the advantage of the description as function of absorption optical thickness is obvious. Namely, the intensity depends smoothly on absorption optical thickness, which means that only calculations at a limited number of values of τ_{abs} are needed from which the Stokes parameters at other values of τ_{abs} can be obtained by interpolation (see e.g. *van Diedenhoven et al., 2006*). Finally, the Stokes parameters can be mapped back into wavelength space. To apply this procedure for a non-homogeneous atmosphere, one has to assume that the vertical distribution **n**(*z*) of τ_{abs} can be approximated by a vertical distribution that is independent of wavelength in the spectral interval under consideration. This is the underlying assumption of the correlated k method. For the simulation of moderate- or high resolution spectra in the near- and shortwave infrared spectral ranges, this assumption may cause errors of several percent in reflectance spectra for realistic inhomogeneous terrestrial atmospheres (see e.g. *Duan et al., 2005*).

Clearly, an efficient radiative transfer model is needed that accounts for the vertical distribution of absorption optical thickness at different wavelengths. For this purpose we use the transformation into absorption optical thickness space, and perform calculations for a limited number *N* values of τ^k_{abs} of the absorption optical thickness and corresponding vertical distributions $\mathbf{n}^k$, with *k*=1,…,*N*. From the reference calculations $\mathbf{I}_{ms}(\tau^k_{abs},\mathbf{n}^k)$ we want to obtain the multiply scattered intensity vector $\mathbf{I}_{ms}(\lambda_j)$ at any wavelength λ_j in the spectral range under consideration with absorption optical thickness $\tau_{abs}(\lambda_j)$ and its vertical distribution $\boldsymbol{n}(\lambda_j)$.

The basic principle of our linear-k method is to account for differences between the actual vertical distribution $\boldsymbol{n}(\lambda_j)$ and the vertical distributions $\mathbf{n}^k$ used in the reference calculations, by employing the linear approximation:

$$I_{ms}\left(\tau_{abs}^{k}, \boldsymbol{n}(\lambda_j)\right) \approx I_{ms}(\tau_{abs}^{k}, \mathbf{n}^k) + \frac{\partial I_{ms}}{\partial \boldsymbol{n}}\left[\boldsymbol{n}(\lambda_j) - \mathbf{n}^k\right], \tag{30}$$

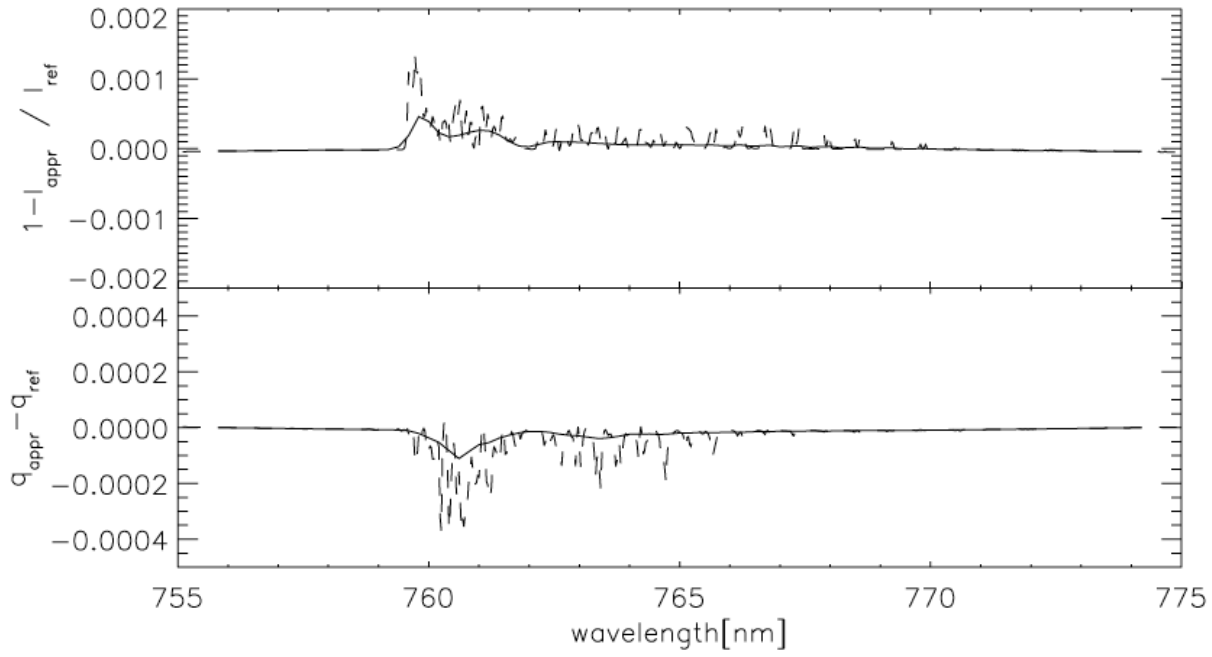
where the derivatives with respect to **n** follow from the forward adjoint perturbation theory (*Hasekamp and Landgraf, 2005*). Applying this equation yields the intensity vector of the multiply scattered radiation at the grid points τ^k_{abs}, corrected for the actual vertical distribution. In order to obtain the intensity $\mathbf{I}_{ms}$ at $\tau_{abs}(\lambda_j)$ we fit a second order polynomial to the logarithm of the (absorption profile corrected) intensities at the grid points, using the grid points closest to $\tau_{abs}(\lambda_j)$ and the two neighbouring points. In order to correct for variation of scattering properties and surface albedo within the spectral range under consideration, also a linear correction is used.

The grid points are chosen equidistant on a logarithmic scale. For grid point *k* the total absorption optical thickness is given by:

$$\tau_{grid}^{k} = e^{\log(\tau_{\min})+(k-1)[[\log(\tau_{\max})-\log(\tau_{\min})]/N]} \tag{31}$$

where τ_{min} and τ_{max} are respectively the minimum and maximum absorption optical thickness in the spectral window under consideration. If τ_{max} is larger than 15 its value is set to 15, because for larger values of the absorption optical thickness the radiation field is dominated by single scattering (being calculated exactly) and hence interpolation errors are of minor importance. The rationale of choosing a logarithmic scale is to obtain more grid points at small values of absorption optical thickness, where multiple scattering effects are most important. If a certain spectral range is influenced by considerable absorption by two or more species we use 2 grids: one for the target absorber and one for the total of other absorbers. For the radiative transfer calculations in the methane retrieval algorithm we use 10 grid points in the NIR band and 5 X 4 = 20 grid points (5 for CH_4 and 4 for H_2O and CO combined) in the SWIR band. Figure 5 illustrates the accuracy of the linear-k method for the NIR spectral range. For more information on the linear-k method we refer to the paper of *Hasekamp and Butz 2008*.

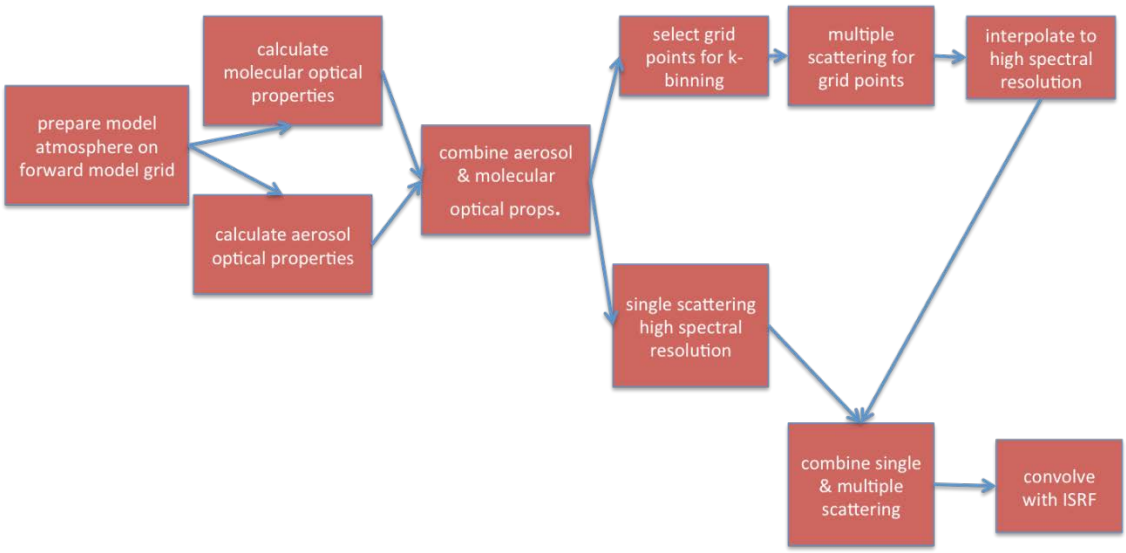
Figure 5: Relative difference between a spectrum calculated using the linear-k method and a spectrum obtained using line-by-line calculations. Upper panel shows the intensity and the lower panel the degree of linear polarization. The solid and dashed lines correspond to spectral resolutions of S5P and GOSAT, respectively. For the calculations a boundary layer aerosol was used with an optical thickness of 0.3 at 765 nm. Furthermore, we used a solar zenith angle (SZA) of 50° and a viewing zenith angle of 0°.



3.1.3 Summary of Forward Model

The different steps of the forward model calculation are summarized in Figure 6.

Figure 6: Overview of forward model



3.1.4 Inverse algorithm

Definition of state vector and ancillary parameters

The **state vector x** contains the following elements (between brackets are optional elements):

- CO_2 sub-columns in 12 vertical layers (layer interfaces coincide with NLAY layers of forward model grid).
- CH_4 sub-columns in 12 vertical layers (layer interfaces coincide with NLAY layers of forward model grid).
- H_2O total column.
- Aerosol parameter N_{aer} (number column).
- Aerosol parameter α (size parameter).
- Aerosol parameter z_{aer} (central height of Gaussian height distribution).
- Lambertian surface albedo in all bands band.
- First order spectral dependence of surface albedo in all bands.
- Spectral shift of Earth radiances in all bands (higher orders optional).
- Spectral shift of Earth radiances in all bands (higher orders optional).
- Intensity offset in the NIR band.
- (offset in input temperature profile).
- (surface pressure).

The **ancillary parameter vector b** contains the following parameters:

- H_2O sub-columns in 36 vertical layers of forward model grid.
- Temperature vertical profile at 72 layers of cross-section vertical grid.
- Pressure vertical profile at 72 layers of cross-section vertical grid.
- The aerosol complex refractive index (fixed at 1.4-0.01i for NIR and 1.47-0.008i for SWIR).
- The width w_o of the Gaussian aerosol height distribution (fixed at 2000 meter).

Table 3: A priori values for the different state vector elements.

State vector element	**A priori value**
CH_4 sub-columns	TM4
CO2 sub-columns	Carbontracker
H_2O total column	ECMWF
N_{aer}	AOT=0.1 @760 nm
α	fixed at 3.5
z_{aer}	fixed at 5000 meter
surface albedo (NIR + SWIR)	no prior value needed (first guess at maximum of measured reflectance)
spectral shifts	no prior needed (first guess = 0)
temperature offset	no prior needed (first guess = 0)
surface pressure	ECMWF + SRTM DEM

3.1.5 Inversion Procedure

The inverse method optimizes the state vector ***x*** with respect to the measurements ***y*** after applying the forward model ***F*** to ***x***. The inverse method is based by default on a Phillips-Tikhonov regularization scheme (*Phillips, 1962; Tikhonov, 1963; Hasekamp and Landgraf, 2005a*). Regularization is required because the inverse problem is ill-posed, i.e., the measurements ***y*** typically contain insufficient information to retrieve all state vector elements independently. The inverse algorithm finds ***x*** by minimizing the cost function that is the sum of the least-squares cost function and a side constraint weighted by the regularization parameter γ according to

$$\hat{\boldsymbol{x}} = \min_{x}(||\mathbf{S}_y^{-\frac{1}{2}}(\mathbf{F}(\mathbf{x}) - \mathbf{y})||^2 + \gamma||\mathbf{W}(\mathbf{x} - \mathbf{x}_a)||^2) \quad (32)$$

where $\boldsymbol{S_y}$ is the diagonal measurement error covariance matrix, which contains the noise estimate. $\mathbf{x}_a$ is an a priori state vector (see Table 3), and ***W*** is a weighting matrix (see below).

For the linearized forward model for iteration step n, the equation for the updated state vector x_{n+1} reduces to

$$\mathbf{x}_{n+1} = \min(||\mathbf{K}'(\mathbf{x}'-\mathbf{x}_n') - \mathbf{y}'||^2 + \gamma ||\mathbf{x}'-\mathbf{x}_a'||^2) \quad (33)$$

with the weighted quantities $\mathbf{x}' = \mathbf{W}\mathbf{x}$, $\mathbf{y}' = \mathbf{S}_y^{-1/2}(\mathbf{y} - \mathbf{F}(\mathbf{x}_n))$ and $\mathbf{K}' = \mathbf{S}_y^{-1/2}\mathbf{K}\mathbf{W}^{-1}$.

The solution reads

$$\mathbf{x}_{n+1} = \mathbf{G}'\mathbf{y}' + \mathbf{A}'\mathbf{x}_n' + (I - \mathbf{A}')\mathbf{x}_{apr}' \quad (34)$$

with ***A'*** the averaging kernel matrix and ***G'*** the contribution function matrix given by $\mathbf{A}'=\mathbf{G}'\mathbf{K}'$ and $\mathbf{G}'=(\mathbf{K}'^{T}\mathbf{K}'+\gamma\mathbf{I})^{-1}\,\mathbf{K}'^{T}$.

If the retrieval converges after a given number of steps *N* (typically 7-8), the final state vector $\boldsymbol{x}_{retr} = \boldsymbol{x}_N$ is related to the true state vector and to the prior via

$$\mathbf{x}_{retr} = \mathbf{A}\mathbf{x}_{true} + (\mathbf{I} - \mathbf{A})\mathbf{x}_a + \mathbf{G}\mathbf{e}_y + \mathbf{G}\mathbf{e}_F \quad (35)$$

The covariance matrix $\mathbf{S}_x$ describing the retrieval noise ($\mathbf{Ge}_y$) is given by

$$\mathbf{S}_x = \mathbf{G}\mathbf{S}_y\boldsymbol{G}^T \quad (36)$$

The target retrieval quantity is the column averaged dry air greenhouse gas mixing ratio, XGHG, where GHG denotes either CO_2 or CH_4. This quantity is obtained from the CH_4 and CO_2 entries of the retrieved state vector through

$$XGHG = \boldsymbol{h}^T x_{retr}/V_{air,dry} \quad (37)$$

where h is the total column operator for methane (summing up the partial columns in the state vector) and $V_{air,dry}$ is the dry air column calculated from the surface pressure and water vapor profile, both obtained from a meteorological model (required as input). The retrieval noise ΔXCH_4 on XCH_4 is given by

$$\Delta XCH_4 = \frac{\sum_{i=1}^{12}\sum_{j=1}^{12} S_{x,i,j}}{V_{air,dry}} \quad (38)$$

And a similar expression for CO_2.

For validation and application purposes it is important to realize that the retrieved XCH_4 or XCO_2 is in fact a representation of $\mathbf{ax}_{true} / V_{air,dry}$, where the quantity

$$\mathbf{a} = \mathbf{h}^T\mathbf{A} \quad (39)$$

is referred to as the column averaging kernel (*Rodgers and Connor, 2003*).

3.1.6 Regularization of state vector and iteration strategy

To retrieve a meaningful state vector ***x***, the side-constraint in the minimization equation should be chosen in a way that contributions from measurement noise are minimized while retaining all valuable information in the first part of the merit-function. The inverse algorithm relies on a regularized Phillips-Tikhonov scheme. The diagonal weighting matrix ***W*** is given by

$$\mathbf{W} = \mathbf{L}\mathbf{W}' \tag{40}$$

with

$$\mathbf{L} = \begin{pmatrix} 1 & -1 & \cdots & & & & \\ -1 & 2 & -1 & & & & \\ \vdots & & \ddots & & \vdots & & \\ & & -1 & 1 & & & \\ & & \cdots & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \tag{41}$$

The upper left 12 by 12 sub matrix works on the state vector elements that contain the CH_4 sub-columns in the 12 altitude layers of the retrieval vertical grid. This sub-matrix corresponds to the matrix product

$$\mathbf{L}_1^T\mathbf{L}_1 \tag{42}$$

where $\mathbf{L}_1$ is the first derivative matrix. The W'_{jj} is given by 1/MAX[K(1:12,1:12)] for the state vector elements corresponding to the 12 sub-columns of methane,1/MAX[K(:,j)] for the aerosol parameters, and 0 for all other parameters (which means they are not constrained by the side constraint and are retrieved in a least-squares sense). The value for γ is fixed such that the Degrees of Freedom for Signal (DFS) for the methane profile is in the range 1.0-1.5. This value is found empirically.

If the true solution is far from the current iteration $\boldsymbol{x_n}$ and the forward model $\boldsymbol{F(x_n)}$ far from being linear, the linear approximation may fail and yield a solution $\boldsymbol{x_{n+1}}$ that is further away from the true solution than $\boldsymbol{x_n}$. To avoid diverging retrievals, we use a Gauss-Newton scheme (see for instance *Rodgers 2000*) with reduced step-size by introducing a filter factor that limits the update per iteration step of the state vector. The updated state vector is then given by:

$$\mathbf{x}_{n+1} = \Lambda\mathbf{G}'\mathbf{y}' + \mathbf{A}'\mathbf{x}_n' + (I - \mathbf{A}')\mathbf{x}_{apr}' \quad (43)$$

with the filter factor Λ given by

$$\Lambda = \frac{1}{1+\xi}, \xi > 0 \quad (44)$$

If ξ is large the update of the state vector is small. If ξ = 0, the equation for the updated state vector is equivalent to the pure Phillips-Tikhonov equation. The iteration is started with a large ξ, typically on the order of 10. It is then reduced or increased in the following iteration steps according to an empirically found scheme similar to Levenberg-Marquardt strategies (*Rodgers, 2000*). We accept solution $\boldsymbol{x'_{n+1}}$ and decrease ξ by a factor of 2.5 if the least squares norm of iteration n+1 is smaller than 1.1 times the least squares norm of iteration n. Otherwise, we discard the solution of iteration n+1, increase ξ by a factor of 2.5 and solve again for $\boldsymbol{x'_{n+1}}$. If ξ is smaller than a threshold value of 0.05, it is set to zero and the iteration is continued without a reduction in step size, assuming that the current state vector is sufficiently close to the true solution to finally converge.

3.1.7 Convergence criteria

The iteration is terminated and the retrieval is considered to have converged to a valid solution $\boldsymbol{x_{retr}}$ if the following four conditions are all met:

1. The update of the state vector has become smaller than its theoretical uncertainty.
2. The state vector corresponding to CO_2 and/or CH_4 entries have never reached unrealistic values during the retrieval (negative methane densities for instance).
3. The merit-function has not increased for the current iteration step, and the step-size factor has reached 0.
4. The merit-function divided by the degrees of freedom is smaller than 2.0.

3.1.8 Cloud Filtering

Cloud filtering in RemoTeC is based on retrieved columns of oxygen (VO2), carbon dioxide (VCO2) and water vapor (VH2O) retrieved independently from the 0.75 micron, 1.6 micron, and 2.0 micron bands, respectively, under the assumption of a non-scattering atmosphere: 0.95<VO2_retrieved/VO2_ecmwf<1.02, 0.99<VCO2_1.6micron/VCO2_2.0micron<1.01, 0.95<VH2O_1.6micron / VH2O_2.0micron < 1.07. The rationale for these cloud filters is that scenes with a large light path deviation with respect to a non scattering atmosphere will result in different CO2 and H2O columns retrieved (without scattering) from the 1.6 and 2.0 micron band due to different light path sensitivities in the two bands. Also, the retrieved O2 column will deviate more from the ECMWF O2 column for large light path differences with a non scattering atmosphere. In

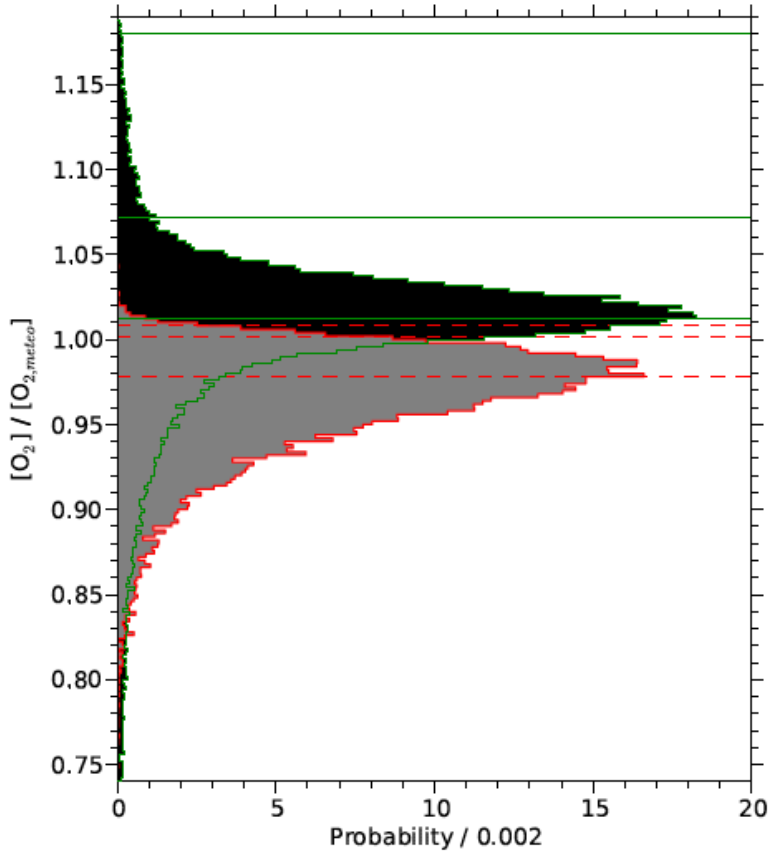
addition to these filters based on non scattering retrievals we also filter out cases for which the aerosol filter $f = AOT_{SWIR} \times \frac{h}{\alpha} < 110$. This filters out scenes with high AOT, large particles, high in the atmosphere, which are the most difficult cases in terms of light path adjustment. In this way we filter out difficult aerosol cases and scenes contaminated with thin cirrus.

3.1.9 Scaling of O_2 Cross Sections

Our full physics retrieval method relies on accurate modeling ofthe O2 A-band which provides a large part of the information content on atmospheric scattering properties. Therefore, we first investigate whether our model of the O2 A-band is consistent with the measurements before aiming at XCO2 and XCH4 retrievals. Butz et al. 2011 proposed a method to discriminate between errors related to unaccounted light scattering effects and other sources of error using ocean-glint observations. The method has been demonstrated using measurements of the Thermal And Near infrared Sensor for carbon Observation (TANSO) Fourier Transform Spectrometer (FTS) onboard GOSAT. Figure 7 (adapted from *Butz et al., 2013*) motivates the general concept by comparing the total column concentration of molecular oxygen ([O2]) retrieved from TANSO-FTS O2 A-band (~13 100 cm−1) spectra to [O2] calculated with high accuracy from meteorological support data. The retrievals neglect any (particulate as well as molecular) light scattering effects and, thus, deviations between meteorological and retrieved [O2] are to be expected. Figure 7, however, shows that these deviations exhibit a quite different pattern for land-nadir and ocean-glint observations. For land nadir, the distribution scatters rather symmetrically about the occurrence peak with a long tail to strong overestimation as well as underestimation. For ocean glint, in contrast, the distribution exhibits a sharp “upper edge” and wide lower tail. The occurrence peak for ocean glint is offset to lower retrieved [O2] with respect to the land-nadir retrievals. The rationale to explain the observed pattern assumes that, for ocean glint, virtually all scattering effects result in a shortening of the lightpath while, for land nadir, scattering effects can have a lightpath enhancing as well as a lightpath shortening effect. Lightpath enhancement typically requires a scattering event at the Earth’s surface, which, however, is dark for the off-glint ocean. Retrievals of [O2] from oceanglint observations therefore virtually always suffer from unaccounted net lightpath shortening and, thus, yield an underestimation of the true [O2] if the retrieval algorithm neglects light scattering effects. For land nadir, the trade-off between lightpath shortening and lightpath enhancement is largely controlled by surface albedo, which typically varies substantially for land surfaces. This rationale fits the pattern observed in Figure 7, explaining the sharp upper edge and wide lower tail for ocean glint as well as the rather symmetric distribution for land nadir.

If this argument is true, simple “non-scattering” [O2] retrievals from the O2 A-band can be used to find an ensemble of clean ocean-glint scenes where lightpath modification due to scattering effects is negligible. Following our rationale above, the upper edge of ocean-glint retrievals, i.e. the retrievals which show the least underestimation of the expected meteorological O2 concentration, are the scenes with the least contamination by scattering effects. For the upper edge ensemble, light scattering as a source of error is negligible, making it a good candidate ensemble to examine other sources of error such as spectroscopic uncertainties or instrumental deficiencies. Using this method we found the need to scale the O_2 absorption cross sections by a factor 1.03 in order to find agreement with the O_2 column inferred from ECMWF data.

Figure 7: Histogram of the ratio between the total column O2 concentration retrieved from TANSO-FTS O2 A-band spectra ([O2]) and the expected meteorological O2 total column concentration ([O2,meteo]). Histograms are shown separately for retrievals from nadir soundings over land (green lines, black fill) and from glintspot soundings over the ocean (red lines, grey fill). The horizontal lines depict the 99th and 95th percentiles and the occurrence peak (from top to bottom). Retrievals neglect molecular as well as particulate light scattering effects and cover more than 3 yr of GOSAT operations from June 2009 to September 2012 (_1 million nadir soundings, _0.3 million glint soundings). Basic quality screening is based on convergence of the iterative algorithm, instrument anomaly flagging, goodness of fit, and cirrus contamination detectable at the 5160 cm−1 water vapour absorption band.



4. Output data

The output data are stored in one netcdf file per day. The file size varies between 3 and 5 Mb. Separate output files are being generated for the XCH_4 Full Physics and the XCO_2 Full Physics products.

Note that the format of the main output data, which are the Level 2 data products, is described in the separate Product User Guide (PUG) document.

References

Bucholz et al., 1995: Bucholz, A., Rayleigh-scattering calculations for the terrestrial atmosphere," Appl. Opt. 34, 2765-2773, 1995.

Buchwitz et al., 2013a: Buchwitz, M., M. Reuter, O. Schneising, H. Boesch, S. Guerlet, B. Dils, I. Aben, R. Armante, P. Bergamaschi, T. Blumenstock, H. Bovensmann, D. Brunner, B. Buchmann, J. P. Burrows, A. Butz, A. Chédin, F. Chevallier, C. D. Crevoisier, N. M. Deutscher, C. Frankenberg, F. Hase, O. P. Hasekamp, J. Heymann, T. Kaminski, A. Laeng, G. Lichtenberg, M. De Mazière, S. Noël, J. Notholt, J. Orphal, C. Popp, R. Parker, M. Scholze, R. Sussmann, G. P. Stiller, T. Warneke, C. Zehner, A. Bril, D. Crisp, D. W. T. Griffith, A. Kuze, C. O'Dell, S. Oshchepkov, V. Sherlock, H. Suto, P. Wennberg, D. Wunch, T. Yokota, Y. Yoshida, The Greenhouse Gas Climate Change Initiative (GHG-CCI): comparison and quality assessment of near-surface-sensitive satellite-derived CO_2 and CH_4 global data sets, *Remote Sensing of Environment*, doi:10.1016/j.rse.2013.04.024, http://authors.elsevier.com/sd/article/S0034425713003520, 2013.

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

Buchwitz et al., 2016: Buchwitz, M., Reuter, M., Schneising, O., Hewson, W., Detmers, R. G., Boesch, H., Hasekamp, O. P., Aben, I., Bovensmann, H., Burrows, J. P., Butz, A., Chevallier, F., Dils, B., Frankenberg, C., Heymann, J., Lichtenberg, G., De Mazière, M., Notholt, J., Parker, R., Warneke, T., Zehner, C., Griffith, D. W. T., Deutscher, N. M., Kuze, A., Suto, H., and Wunch, D.:, Global satellite observations of column-averaged carbon dioxide and methane: The GHG-CCI XCO_2 and XCH_4 CRDP3 data, Remote Sensing of Environment (in press), Special Issue on Essential Climate Variables, DOI: 10.1016/j.rse.2016.12.027, (link: http://dx.doi.org/10.1016/j.rse.2016.12.027), 2016.

Buchwitz et al., 2017: ESA Climate Change Initiative (CCI) Product Validation and Intercomparison Report (PVIR) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG) for data set Climate Research Data Package No. 4 (CRDP#4), Version 5.0, 9. Feb. 2017, link: http://www.esa-ghg-cci.org/?q=webfm_send/352, 2017.

Butz et al., 2009: Butz, A., Hasekamp, O. P., Frankenberg, C., Aben, I. (2009). Retrievals of atmospheric CO2 from simulated space-borne measurements of backscattered near-infrared sunlight: Accounting for aerosol effects. Appl. Opt., 48, 3322.

Butz et al., 2011: Butz, A., Guerlet, S., Hasekamp, O., et al., Toward accurate CO_2 and CH_4 observations from GOSAT, *Geophys. Res. Lett.*, doi:10.1029/2011GL047888, 2011.

Butz et al., 2012: Butz, A., Galli, A., Hasekamp, O., Landgraf, J., Tol, P., and Aben, I.: Remote Sensing of Environment, TROPOMI aboard Sentinel-5 Precursor : Prospective performance of CH_4 retrievals for aerosol and cirrus loaded atmospheres, 120, 267-276, doi:10.1016/j.rse.2011.05.030, 2012.

Chevallier et al., 2010: Chevallier, F., P. Ciais, T. J. Conway, T. Aalto, B. E. Anderson, P. Bousquet, E. G. Brunke, L. Ciattaglia, Y. Esaki, M. Fröhlich, A.J. Gomez, A.J. Gomez-Pelaez, L. Haszpra, P. Krummel, R. Langenfelds, M. Leuenberger, T. Machida, F. Maignan, H. Matsueda, J. A. Morguí, H. Mukai, T. Nakazawa, P. Peylin, M. Ramonet, L. Rivier, Y. Sawa, M. Schmidt, P. Steele, S. A. Vay, A. T. Vermeulen, S. Wofsy, D. Worthy, 2010: CO_2 surface fluxes at grid point scale estimated from a global 21-year reanalysis of atmospheric measurements. J. Geophys. Res., 115, D21307, doi:10:1029/2010JD013887

Detmers et al., 2015: Detmers, R. G., O. Hasekamp, I. Aben, S. Houweling, T. T. van Leeuwen, A. Butz, J. Landgraf, P. Koehler, L. Guanter, and B. Poulter, Anomalous carbon uptake in Australia as seen by GOSAT, Geophys. Res. Lett., 42, doi:10.1002/2015GL065161, 2015.

Van Diedenhoven et al., 2005: Van Diedenhoven, B., O.P. Hasekamp, and I. Aben, 2005: Surface pressure retrieval from SCIAMACHY measurements in the O_2 A Band: validation of the measurements and sensitivity on aerosols. *Atmos. Chem. Phys.*, **5**, 2109-2120, doi:10.5194/acp-5-2109-2005.

Van Diedenhoven et al., 2007: van Diedenhoven, B., Hasekamp, O., Landgraf, J. (2007), Retrieval of cloud parameters from satellite-based reflectance measurements in the ultraviolet and the oxygen A-band, J. Geophys. Res., 112, D15208.

Dubovik et al:, 2006: Dubovik O., Sinyuk A., Lapyonok T., Holben B. N., Mishchenko M., Yang P., Eck T. F., Volten H., Muñoz O., Veihelmann B., van der Zande W. J., Leon J.-F., Sorokin M., Slutsker I. 2006. Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust. J. Geophys. Res. D 111, D11208.

Guerlet et al., 2013: Guerlet, S., A. Butz, D. Schepers, S. Basu, O. P. Hasekamp, A. Kuze, T. Yokota, J.-F. Blavier, N. M. Deutscher, D. W. T. Griffith, F. Hase, E. Kyro, I. Morino, V. Sherlock, R. Sussmann, A. Galli and I. Aben (2013) Impact of aerosol and thin cirrus on retrieving and validating XCO2 from GOSAT shortwave infrared measurements, J. Geophys. Res., doi: 10.1002/jgrd.50332

De Haan et al., 1987: de Haan, J. F., Bosma, P. B., Hovenier, J. W. (1987), The adding method for multiple scattering calculations of polarized light, Astron. and Astrophys., 183, 371 – 391.

Hansen et al., 1998: Hansen, P. C. (1998), Rank-deficient and discrete Ill-posed problems: Numerical aspects of linear inversion, SIAM - Monographs on Mathematical Modeling and Computation 4, Philadelphia, USA.

Hasekamp et al., 2002: Hasekamp, O. P. and Landgraf, J. (2002), A linearized vector radiative transfer model for atmospheric trace gas retrieval, J. Quant. Spectrosc. Radiat. Transfer, 75, 221–238.

Hasekamp et al., 2005a: Hasekamp, O. P. and Landgraf, J. (2005a), Linearization of vector radiative transfer with respect to aerosol properties and its use in satellite remote sensing, J. Geophys. Res, 110, D04203.

Hasekamp et al., 2005b: Hasekamp, O. P. and Landgraf, J. (2005b), Retrieval of aerosol properties over the ocean from multispectral single-viewing-angle measurements of intensity and polarization: Retrieval approach, information content, and sensitivity study, J. Geophys. Res. 110, D20207.

Hasekamp et al., 2008: Hasekamp, O. P. and Butz, A. (2008), Efficient calculation of intensity and polarization spectra in vertically inhomogeneous scattering and absorbing atmospheres, J. Geophys. Res., 113, D20309.

Jenouvrier et al., 2007: Jenouvrier, A., Daumont, L., Régalia-Jarlot, L., Tyuterev, V. G., Carleer, M., Vandaele, A. C., Mikhailenko, S., Fally, S. (2007), Fourier transform measurements of water vapor line parameters in the 4200 – 6600 cm-1 region, J. Quant. Spectrosc. Ra., 105, 326 – 355.

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.

Kuze et al., 2014: Kuze, A., Taylor, T., Kataoka, F., Bruegge, C., Crisp, D., Harada, M., Helmlinger, M., Inoue, M., Kawakami, S., Kikuchi, N., Mitomi, Y., Murooka, J., Naitoh, M., O'Brien, D., O'Dell, C., Ohyama, H., Pollock, H., Schwandner, F., Shiomi, K., Suto, H., Takeda, T., Tanaka, T., Urabe, T., Yokota, T., and Yoshida, Y. (2014), Long-term vicarious calibration of GOSAT short-wave sensors: techniques for error reduction and new estimates of radiometric degradation factors, IEEE T. Geosci. Remote, 52, 3991–4004, doi:10.1109/TGRS.2013.2278696, 2014.

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

Landgraf et al., 2001: Landgraf, J., O. Hasekamp, T. Trautmann, and M. Box (2001), A linearized radiative transfer model for ozone profile retrieval using the analytical forward-adjoint perturbation theory, J. Geophys. Res., 106, 27,291 – 27,306.

Landgraf et al., 2002: Landgraf, J., O. Hasekamp, and T. Trautmann (2002), Linearization of radiative transfer with respect to surface properties, J. Quant. Spectrosc. Radiat. Transfer, 72, 327– 339.

Landgraf et al., 2009: Landgraf, J., Tol, P., Butz, A., Hasekamp, O., Aben, I. (2009), Level 1-B requirements for the SWIR 1.6 mm and 2.3 mm band, Camelot report, SRON/EOS/RP/09-005, SRON-Utrecht, The Netherlands.

Landgraf et al., 2010: Landgraf Jochen, Paul Tol, Otto P. Hasekamp, "Spatial Re-Sampling of Spectral Radiance Measurements of the O2 A-Band", Technote SRON-TROPSC-TN-2010-014, Netherlands Institute for Space Research (SRON), Utrecht, The Netherlands, 2010

Mishchenko et al., 1999: Mishchenko, M. I., Geogdzhayev, I. V. , Cairns, B., Rossow, W. B., Lacis, A. A. (1999), Aerosol retrievals over the ocean using channel 1 and 2 AVHRR data: A sensitivity analysis and preliminary results, Appl. Opt., 38, 7325 – 7341.

Peters et al., 2007: Peters, Wouter and Jacobson, Andrew R. and Sweeney, Colm and Andrews, Arlyn E. and Conway, Thomas J. and Masarie, Kenneth and Miller, John B. and Bruhwiler, Lori M. P. and Pétron, Gabrielle and Hirsch, Adam I. and Worthy, Douglas E. J. and van der Werf, Guido R. and Randerson, James T. and Wennberg, Paul O. and Krol, Maarten C. and Tans, Pieter P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, PNAS, 104, 48, doi: 10.1073/pnas.0708986104

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

Rodgers et al., 2000: Rodgers, C.D. (2000), Inverse methods for atmospheric sounding: Theory and practice, World Sc., River Edge, USA.

Rodgers et al., 2003: Rodgers, C. D., and B. J. Connor (2003), Intercomparison of remote sounding instruments, J. Geophys. Res., 108, 4116, doi:10.1029/2002JD002299, D3.

Rothman et al., 2009: Rothman, L. S., Gordon, I. E., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., Boudon, V., Brown, L. R., Campargue, A., Champion, J. P., et al. (2009), The HITRAN 2008 molecular spectroscopic database, J. Quant. Spectrosc. Ra., 110, 533 – 572.

Scheepmaker et al., 2012: Scheepmaker, R. A., Frankenberg, C., Galli, A., Schrijver, H., Fally, S., Deutscher, N. M., Wunch, D., Warneke, T., Aben, I. (2012), Improved water vapour spectroscopy in the 4174 – 4300 cm-1 region and its impact on SCIAMACHY's HDO/ H2O measurements, Atmos. Meas. Tech., in preparation.

Schepers et al., 2012: Schepers, D., et al. (2012), Methane retrievals from Greenhouse Gases Observing Satellite (GOSAT) shortwave infrared measurements: Performance comparison of proxy and physics retrieval algorithms, J. Geophys. Res., 117, D10307, doi:10.1029/2012JD017549.

Stier et al., 2005: Stier, P., Feichter, J., Kinne, S., Kloster, S., Vignati, E., Wilson, J., Ganzeveld, L., Tegen, I., Werner, M., Balkanski, Y., Schulz, M., Boucher, O., Minikin, A., Petzold, A. (2005), The aerosol-climate model ECHAM5-HAM, Atmos. Chem. Phys., 5, 1125 – 1156.

Tikhonov et al., 1963: Tikhonov, A. (1963), On the solution of incorrectly stated problems and a method of regularization, Dokl., Akad., Nauk SSSR, 151, 501-504.

Tran et al., 2006: Tran, H., Boulet, C., Hartmann, J. M. (2006), Line mixing and collision-induced absorption by oxygen in the A band: Laboratory measurements, model, and tools for atmospheric spectra computations, J. Geophys. Res., 111, D15210.

Tran et al., 2008: Tran, H., and J.-M. Hartmann (2008), An improved O2A band absorption model and its consequences for retrievals of photon paths and surface pressures, J. Geophys. Res., 113, D18104, doi:10.1029/2008JD010011.

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.

Velazco et al. 2011: Velazco, V. A., Buchwitz, M., Bovensmann, H., Reuter, M., Schneising, O., Heymann, J., Krings, T., Gerilowski, K., and Burrows, J. P.: Towards space based verification of CO2 emissions from strong localized sources: fossil fuel power plant emissions as seen by a CarbonSat constellation, Atmos. Meas. Tech., 4, 2809-2822, 2011.

Wunch et al. 2010: Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jiménez, R., Keppel-Aleks, G., Kort, E. A.,

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

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

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: info@copernicus-climate.eu

climate.copernicus.eu copernicus.eu ecmwf.int