Algorithm Theoretical Baseline Document for Sentinel-5 Precursor: Carbon Monoxide Total Column Retrieval

Jochen Landgraf, Joost aan de Brugh, Remco A. Scheepmaker, Tobias Borsdorff, Sander Houweling, Otto P. Hasekamp
### Document approval record

<table>
<thead>
<tr>
<th>digital signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>prepared:</td>
</tr>
<tr>
<td>checked:</td>
</tr>
<tr>
<td>Michael Buchwitz and Thomas Krings</td>
</tr>
<tr>
<td>Institute of Environmental Physics (IUP)</td>
</tr>
<tr>
<td>University Bremen, Germany</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>approved PM:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>approved PI:</th>
</tr>
</thead>
<tbody>
<tr>
<td>issue</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>0.0.1</td>
</tr>
<tr>
<td>0.0.2</td>
</tr>
<tr>
<td>0.0.2</td>
</tr>
<tr>
<td>0.5.0</td>
</tr>
<tr>
<td>0.5.0</td>
</tr>
<tr>
<td>0.9.0</td>
</tr>
<tr>
<td>0.9.1</td>
</tr>
<tr>
<td>0.10.0</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>0.11.0</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>0.11.1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>0.13.0</td>
</tr>
<tr>
<td>0.13.0</td>
</tr>
<tr>
<td>1.00</td>
</tr>
<tr>
<td>1.10</td>
</tr>
<tr>
<td>2.20</td>
</tr>
<tr>
<td>2.40</td>
</tr>
</tbody>
</table>
# Contents

Document approval record ................................................................. 2
Document change record .................................................................... 3
List of Tables .................................................................................. 5
List of Figures ................................................................................ 6

1 Introduction .................................................................................. 9
1.1 Identification .............................................................................. 9
1.2 Purpose and objectives ............................................................. 9
1.3 Document overview ................................................................. 9

2 Applicable and reference documents ............................................ 10
2.1 Applicable documents ............................................................. 10
2.2 Standard documents ............................................................... 10
2.3 Reference documents ............................................................. 10
2.4 Electronic references ............................................................. 14

3 Terms, definitions and abbreviated terms .................................... 15
3.1 Acronyms and abbreviations .................................................... 16

4 Remote Sensing of Carbon Monoxide ......................................... 18
4.1 Algorithm heritage ................................................................. 18
4.2 Carbon Monoxide level-2 requirements ..................................... 19

5 Algorithm Description ................................................................. 21
5.1 Forward model ......................................................................... 21
5.2 Inversion .................................................................................. 28
5.2.1 Methane cloud filter ............................................................. 28
5.2.2 Tikhonov regularisation for CO column retrieval .................. 29
5.2.3 Tikhonov regularisation for stability .................................... 32
5.2.4 Step control ......................................................................... 32
5.2.5 Unphysical values ............................................................... 33
5.3 Numerical Implementation and Data Product ......................... 33
5.4 Molecular optical properties .................................................. 34
5.5 Micro-physical properties of the scattering layer ...................... 35
5.6 State vector, ancillary parameters and a priori knowledge ......... 35
5.7 Data product ............................................................................ 36

6 Common aspects with other algorithms ..................................... 38

7 Error analysis ................................................................................ 39
7.1 Performance analysis for generic scenarios ............................... 39
7.2 Performance Analysis for generic scenarios ............................... 41
7.3 Geophysical test scenario for measurements over China .......... 47
7.4 Robustness of the CO retrieval with respect to uncertainties in the atmospheric input ........................................... 54
7.5 Robustness of the CO retrieval with respect to instrument artifacts ........................................................................ 58
7.6 Quality of the model derived XCH\textsubscript{4} .................................. 62
7.6.1 Simulation of XCH\textsubscript{4} ..................................................... 62
7.6.2 Uncertainty of model-derived XCH\textsubscript{4} .......................... 63
7.6.3 Discussion and conclusions .................................................. 65

8 Algorithm input and output ......................................................... 67
8.1 High level processing scheme .................................................. 67
8.2 Static input ............................................................................... 67
8.3 Dynamic input ......................................................................... 67
8.4 Algorithm output ...................................................................... 68

9 Spatial data selection approach .................................................. 70

10 A posteriori destriping of the level-2 data .................................. 71

11 Conclusion ................................................................................ 72

A Appendix: Flux method PIFM ..................................................... 75
B Appendix: SWIR Pre-Processing ................................................ 80
List of Tables

1  Setup of state vector $x$. Here, CO, CH$_4$, H$_2$O and HDO indicates the total column retrieval of the trace gases, $A_r$ and $\Delta A_r$ are the Lambertian surface albedo and its slope, $z_{\text{scat}}$ represents the center height of the scattering layer, $\tau_{\text{scat}}$ the total optical thickness of the layer, and $\Delta \lambda$ is a spectral measurement offset of the measurement. ........................................... 36
2  Microphysical properties of water and ice clouds: $n(r)$ represents the size distribution type, $r_{\text{eff}}$ and $v_{\text{eff}}$ are the effective radius and variance of the size distribution, $n = n_r - in_i$ is the refractive index. The ice cloud size distribution follows a power-law distribution as proposed by [RD1]. ............................................................................................. 39
3  Summary of the different generic test cases A-F. ................................................. 40
4  Estimated uncertainty in XCH$_4$ comparing the proposed modeling approaches. ............ 66
5  SICOR Static input. Calibration key data and irradiance L1b-product are semi-static, because they are provided once per processor run. ................................................................. 68
6  SICOR Dynamic input............................................................................. 68
7  SICOR output fields. $N_z$ is the number of layers in the model atmosphere, and is set to 50 by default. ........................................................................................................ 69
List of Figures

1. SWIR spectral transmittance along the light path of the solar beam reflected by the Earth surface into the instrument viewing direction. Simulations are performed for viewing zenith angle (VZA) = 0°, and a solar zenith angle (SZA) = 30°, and by assuming a US standard atmospheric profile. From top to bottom, the figure shows the total transmittance, the individual transmittances due to H2O, CH4, and CO, respectively. Note the different y-axis scale for CO transmittance. ................................................................. 20

2. Overall structure of the CO retrieval algorithm. For the structure of the SWIR pre-processing algorithm see Appendix B. ........................................................................................................ 22

3. Relative CH4 column above a cloud top height z_cld with respect to the total column amount, using the US standard model atmosphere. ................................................................. 28

4. Probability density function (left panel) and cumulative probability density function (right panel) of the non-scattering methane error for one year of GOSAT observations (2010) with respect to corresponding TM5 model simulations. The figure differentiates the contribution of ocean and land pixels (blue and green line) with respect to the total dataset (orange). The dataset comprises 2.4 10^6 GOSAT measurements in total under which 1.6 10^6 ocean pixels and 8 10^5 land pixels. All retrievals are performed using RemoTeC V2.1. ......................................................................... 29

5. Optimisation of the vertical layering of the two-stream radiative transfer simulation based on the initial grid. The internal grid is needed to account for pressure and temperature dependence of atmospheric absorption. For the radiative transfer, layers above and below the scattering layer can be combined to one layer each, indicated by the blue areas. .............. 34

6. Example of the CO data product. The SWIR measurements are simulated for a scene partially covered by a water cloud between 2 and 3 km with optical depth τ_cld = 30 and a surface albedo A_s = 0.05. Left panel: Difference ΔCO between the true CO column and the retrieved CO column as function of cloud fraction f_cld. Middle panel: 1σ retrieval noise estimate as function of cloud fraction f_cld. Right panel: column averaging kernel as function of altitude for different cloud fractions ................................................. 37

7. Left: atmospheric concentration profiles (bottom axis) and temperature profile (top axis) used as input for the model atmosphere. The concentrations are normalised to the concentration at ground level. Right: assumed profiles for the amount of HDO depletion (solid line, lower axis) and H2^18O depletion (dashed line, top axis). ........................................................................................................ 40

8. Retrieval bias ΔCO (upper panel) and retrieval noise (lower panel) for the clear sky test case A as a function of SZA and surface albedo A_s. .............................................................................. 42

9. Retrieval bias for case B, i.e. for water clouds above a dark surface (A_s = 0.05) as a function of cloud top height z_cld and cloud fraction f_cld for different cloud total optical depths τ_cld = 5 (upper left panel), τ_cld = 10 (upper right panel), τ_cld = 30 (lower left panel), τ_cld = 50 (lower right panel). ........................................................................................................ 43

10. Retrieval bias in case of photon trapping between a water cloud and the bright surface (case C). The CO bias is shown as a function of surface albedo A_s and cloud fraction f_cld for a cloud with optical depth τ_cld = 2 and cloud top height z_cld = 2 km (upper left), for τ_cld = 2 and z_cld = 5 km (upper right), for τ_cld = 5 and z_cld = 2 km (lower left) and for τ_cld = 5 and z_cld = 5 km (lower right). ........................................................................................................ 44

11. Retrieval bias for an aerosol loaded atmosphere of test case D. In the left panel, the CO retrieval bias is shown as a function of surface albedo and aerosol optical depth for a sulfate aerosol between the surface and 2 km altitude. The right panel shows the corresponding error analysis for an urban aerosol layer between 4 and 5 km altitude. At each panel, the lower x-axis describes the aerosol optical depth at 550 nm and the upper x-axis indicates the corresponding aerosol optical depth at 2300 nm. ........................................................................ 45

12. CO retrieval bias for measurements in presence of optically thin cirrus clouds as a function of surface albedo and cirrus optical depth (case E). The cirrus optical depth is given at 2300 nm. ........................................................................................................ 45

13. CO retrieval bias for measurements in presence of multiple cloud layers as a function of cirrus optical depth and cloud fraction of a water cloud for different surface albedo (case F). The cirrus optical depth is given at 2300 nm. ........................................................................................................ 46

14. Test ensemble to generate TROPOMI measurements over China for a 10 x 10 km^2 pixel size for 10 May, 2006. The pixel distortion towards the edge of the swath is adopted from MODIS: (upper panel) MODIS cirrus optical depth, (middle) MODIS ground albedo at 2.1 μm, (lower) MODIS water cloud optical depth. .............................................................................................. 48
(Continued) (upper) MODIS cloud fraction, (middle) MODIS cloud top height, (lower) CH4-MER and TM4 CO total column.  

Bias \( \Delta \)CO of the retrieved CO column for the test ensemble shown in Fig. 14.  

Bias of the retrieved CO column as a function of methane filter \( \Delta \text{CH}_4 \) for the China test ensemble shown in Fig. 14. The median of the methane pre-fit bias is \(-4.0\%\), the mean bias is \(-6.6\%\) with a standard deviation of 6.7%. Accordingly, the median of the CO retrieval bias is \(+1.4\%\), the mean bias is \(+1.6\%\) with a standard deviation of 2.3%. The correlation coefficient \((r)\) between \( \Delta \text{CO} \) and \( \Delta \text{CH}_4 \) is \(+0.01\).  

Same as Fig. 16 but for a non-scattering CO retrieval. The median of the CO retrieval bias is \(-4.90\%\), the mean bias is \(-8.80\%\). The correlation coefficient \((r)\) between \( \Delta \text{CO} \) and \( \Delta \text{CH}_4 \) is \(-0.05\).  

CO column retrieval noise as a function of methane filter \( \Delta \text{CH}_4 \) for the China test ensemble shown in Fig. 14. The LER distribution as a median of 0.2, a mean value of 0.2 and a standard deviation of 0.1, whereas the CO retrieval noise distribution has a median of 2.5% a mean value of 2.6%, and a standard deviation of 0.8%. The correlation coefficient \((r)\) between CO noise and LER is \(-0.5\%).  

Upper panel: CO retrieval bias as a function of a shift \( \Delta T \) of the atmospheric temperature profile. Measurement ensembles of the clear sky generic test cases A (purple lines) and the cloudy sky test case B for a cloud with an optical depth of 10 (orange lines) are considered. Here, the maximum (dashed lines), mean (solid lines) and minimum (dotted lines) bias is reported for the different temperature shifts. Middle panel: mean spectral \( \chi^2 \) of the retrieval as a function of temperature shift. Lower panel: number of converged retrievals as a function of temperature shift.  

Same as Fig. 19 but as a function of methane filter \( \Delta \text{CH}_4 \).  

Same as Fig. 19 but as a function of surface pressure error \( \Delta P \).  

Relative \( \text{H}_2\text{O} \) (left panel) and \( \text{CH}_4 \) (right panel) profiles as a function of height \( z \). Each profile is normalised to the reference profile at 500 m, which is used for the measurement simulation.  

Same as Fig. 19 but assuming erroneous relative vertical \( \text{CH}_4 \) profile from Fig. 22 in the CO retrieval.  

Same as Fig. 23 but for erroneous relative vertical \( \text{H}_2\text{O} \) profile from Fig. 22.  

CO retrieval bias as a function of a FWHM error \( (\Delta \text{FWHM}) \) of the ISRF. The measurement ensembles of the clear sky generic test cases A (purple lines) and the cloudy sky test case B for a cloud of optical depth 10 (orange lines) are considered. Here, the maximum (dashed lines), mean (solid lines) and minimum (dotted lines) bias is reported for the different FWHM errors. Middle panel: mean spectral \( \chi^2 \) of the retrieval as a function of FWHM error. Lower panel: number of converged retrievals as a function of FWHM error.  

Same as Fig. 25 but as a function of the spectral calibration error \( \delta s_2 \) as described in Eq. (73).  

Same as Fig. 25 but as a function of a spectrally constant radiometric error \( \Delta \text{offset} \). Here, \( \Delta \text{offset} \) is defined with respect to the continuum value of the spectrum.  

Same as Fig. 25 but as a function of a spectrally constant radiometric scaling error \( \Delta \text{scale} \). Here, \( \Delta \text{scale} \) is defined with respect to the continuum value of the spectrum.  

Comparison of TM5 simulated and in situ FTS observed total column \( \text{CH}_4 \) at selected sites of the TCCON network. Black: TCCON FTS, Red: TM5.  

Comparison between TM5 simulated and GOSAT retrieved \( \text{CH}_4 \). Differences (TM5 minus GOSAT) are shown for the period June 2009–June 2011. The TM5 results have been optimised using surface measurements.  

The standard deviation of total \( \text{XCH}_4 \) between years expressed in % of mean \( \text{XCH}_4 \). Standard deviations are calculated from TM5 \( \text{XCH}_4 \) fields, optimised using surface measurements, for the 15th day of the month in each year in the period 2003–2010.  

The RMS difference between NOAA optimised \( \text{CH}_4 \) and results of procedure 2 evaluated after 6 months. RMS values are calculated from the differences between the two simulations for all days of the 6th month after initialisation of procedure 2.  

As Fig. 32 for procedure 3, evaluated in different months.  

High level processing scheme for operational SSP CO data reduction. Modules that are described in this ATBD, are indicated by the red boxes.
35 Average surface albedo over five years of SCIAMACHY cloud-free land observations at 2300 nm at a resolution of 0.5° (2003–2007). ............................................................... 70
36 CO retrievals of a TROPOMI orbit granule on 27 June 2018 over the UK. Panels of the first row depict the original data, the second row shows the destriped TROPOMI CO data (FMD method left, FFD method right), and the third row illustrates the destriping mask that was subtracted from the original TROPOMI data. ........................................................................ 71
37 Overall structure of the SWIR preprocessor. ........................................................................ 80
38 CH$_4$ error of a non-scattering retrieval from the SWIR 2315–2324 nm spectral window for a water cloud with optical thickness of 5 as function of cloud height and cloud fraction (left panel, for more details see generic scenario B in Sec. 7.1) and for a cirrus cloud at 10 km height as function of surface albedo and cirrus optical thickness (right panel, for more details see generic scenario E in Sec. 7.1). .......................................................................... 82
39 CH$_4$ two-band cloud filter for the cloud scenarios of Fig. 38. The methane cloud filter relies on non-scattering methane column retrieval from strong and weak absorption features at 2363-2373 nm and 2310-2315 nm, respectively. .......................................................... 83
40 H$_2$O two-band cloud filter for the cloud scenarios of Fig. 38. The filter relies on non-scattering methane column retrieval from strong and weak absorption features at 2367-2377 nm and 2329-2334 nm, respectively. .......................................................... 83
1 Introduction

1.1 Identification

This document describes the carbon monoxide column retrieval algorithm from Sentinel-5 Precursor (S5P) measurements in the shortwave infrared (SWIR) spectral range between 2310 and 2340 nm. It is one of the deliverables of the ESA project ‘Sentinel-5 P level 2 processor development’ [AD1].

1.2 Purpose and objectives

The purpose of the document is to describe the theoretical baseline of the algorithm that is used for the operational processing of the carbon monoxide column densities from S5P measurements in the SWIR spectral range. Input, output and ancillary data are described. Additionally, the performance of the algorithm is analyzed with respect to the expected calculation times and the data product uncertainty.

1.3 Document overview

The document is structured as follows: After this introduction, references are provided in Sec. 2 and Sec. 3 contains a list of abbreviations used in this document. Sec. 4 provides a short introduction to satellite remote sensing of atmospheric CO abundance and the heritage of the presented algorithm is summarized. Moreover, we recall the level-2 requirement for the CO column product which represents the underlying criterion for the performance analysis of the presented algorithm. The theoretical concept of the CO retrieval algorithm SICOR is summarized in Sec. 5, comprising a description of the radiative transfer model and the inversion scheme. The parameters to be retrieved, ancillary data and a priori knowledge are discussed including the final data product of the algorithm. Section 7 considers the performance of the retrieval algorithm based on a set of generic measurement ensembles and a geo-physical ensemble of simulated measurements over China. Here, we investigate the CO retrieval noise and CO retrieval biases due to forward model errors, erroneous atmospheric input data and instrument artifacts. Based on this, we evaluate the algorithm performance in the context of the S5P level-1 and 2 requirements. The numerical feasibility is the subject of Sec. 9, which comprises an estimate of the numerical effort, a high level data product description and the spatial data selection criteria of the measurements to be processed. The a posteriori destriping approach that is deployed on the S5P level-2 data product within the operational processing of ESA is introduced in Sec. 10. and Sec. 11 concludes our document.

Additional material is provided in the appendices, where Appendix A discusses in detail the linearized two-stream method and Appendix B describes the SWIR preprocessing module which provides required input to both the SICOR CO algorithm and the RemoTeC CH₄ algorithm (see [RD2]).
2 Applicable and reference documents

2.1 Applicable documents

source: ESA; ref: S5P-SWESA-GS-053; date: 2012.

[AD2] GMES Sentinels 4 and 5 mission requirements document.  
source: ESA; ref: EOP-SMA/1507/JL-dr; date: 2011.

source: ESA; ref: S5p-RS-ESA-SY-0002; date: 2011.

source: KNMI; ref: S5P-KNMI-L2-0030-SD; issue: 0.0.0; date: 2014-05-22.

2.2 Standard documents

[SD1] Space Engineering – Software.  
source: ESA; ref: ECSS-Q-ST-80C; date: 2009.

source: ESA; ref: ECSS-E-ST-40C; date: 2009.

2.3 Reference documents


source: SRON; ref: SRON-S5P-LEV2-RP-001; date: 2014.

[RD3] Terms, definitions and abbreviations for TROPOMI L01b data processor.  
source: KNMI; ref: S5P-KNMI-L01B-0004-LI; date: 2011.

[RD4] Terms and symbols in the TROPOMI algorithm team.  
source: KNMI; ref: SN-TROPOMI-KNMI-049; date: 2012.


[RD18] The potential of MTG-IRS and S4-TIR to detect high pollution events at urban and regional scales. *source:* EUMETSAT; *ref:* EUM/CO/07/4600000447/SAT; *date:* 2009.


[RD59] SEOM-IAS validation report. *source:* DLR; *ref:* IAS-D09-PRJ-066.


2.4 Electronic references

There are no electronic references.
3 Terms, definitions and abbreviated terms

Terms, definitions and abbreviated terms that are used in the development program for the TROPOMI L0 1b data processor are described in [RD3]. Terms, definitions and abbreviated terms that are used in development program for the TROPOMI L2 data processors are described in [RD4]. Terms, definitions and abbreviated terms that are specific for this document can be found below.
### 3.1 Acronyms and abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADEOS</td>
<td>Advanced Earth Observing System</td>
</tr>
<tr>
<td>AIRS</td>
<td>Atmospheric Infrared Sounder</td>
</tr>
<tr>
<td>AOT</td>
<td>Aerosol Optical Thickness</td>
</tr>
<tr>
<td>AQUA</td>
<td>A NASA Earth Science satellite mission focussing on the Earth’s water cycle</td>
</tr>
<tr>
<td>CTM</td>
<td>Chemical Transport Model</td>
</tr>
<tr>
<td>DFS</td>
<td>Degree of Freedom for Signal</td>
</tr>
<tr>
<td>S-LINTRAN</td>
<td>Scalar linearised Radiative Transfer Program for a Multi-Layered Plane-Parallel Medium</td>
</tr>
<tr>
<td>ECMWF</td>
<td>European Centre for Medium-Range Weather Forecasts</td>
</tr>
<tr>
<td>ERI</td>
<td>European Research Institute</td>
</tr>
<tr>
<td>ESRL</td>
<td>Earth System Research Laboratory</td>
</tr>
<tr>
<td>FFD</td>
<td>Fourier filter destriping</td>
</tr>
<tr>
<td>FFM</td>
<td>Fixed mask destriping</td>
</tr>
<tr>
<td>FTS</td>
<td>Fourier Transform Spectrometer</td>
</tr>
<tr>
<td>FTIR</td>
<td>Fourier Transform Infrared</td>
</tr>
<tr>
<td>FRESCO</td>
<td>Fast Retrieval Scheme for Clouds from the Oxygen A band</td>
</tr>
<tr>
<td>FWHM</td>
<td>Full Width Half Maximum</td>
</tr>
<tr>
<td>GCM</td>
<td>General Circulation Model</td>
</tr>
<tr>
<td>GMES</td>
<td>Global Monitoring for Environment and Security</td>
</tr>
<tr>
<td>GNIP</td>
<td>Global Network for Isotopes in Precipitation</td>
</tr>
<tr>
<td>GOSAT</td>
<td>Greenhouse gases Observing Satellite</td>
</tr>
<tr>
<td>IAGOS</td>
<td>In-service Aircraft for a Global Observing System</td>
</tr>
<tr>
<td>IASI</td>
<td>Infrared Atmospheric Sounding Interferometer</td>
</tr>
<tr>
<td>IMAP</td>
<td>Iterative Maximum A Posteriori</td>
</tr>
<tr>
<td>IMG</td>
<td>Interferometric Monitor for Greenhouse gases</td>
</tr>
<tr>
<td>IMLM</td>
<td>Iterative Maximum Likelihood Method</td>
</tr>
<tr>
<td>IRWG</td>
<td>Infrared Working Group</td>
</tr>
<tr>
<td>ISRF</td>
<td>Instrument Spectral Response Function</td>
</tr>
<tr>
<td>L1</td>
<td>Level-1</td>
</tr>
<tr>
<td>L2</td>
<td>Level-2</td>
</tr>
<tr>
<td>LER</td>
<td>Lambert-equivalent Reflectivity</td>
</tr>
<tr>
<td>LOS</td>
<td>Line of Sight</td>
</tr>
<tr>
<td>MACC</td>
<td>Monitoring Atmospheric Composition and Climate</td>
</tr>
<tr>
<td>MAPS</td>
<td>Measurement of Air Pollution from Satellites</td>
</tr>
<tr>
<td>MODIS</td>
<td>Moderate Resolution Imaging Spectroradiometer</td>
</tr>
<tr>
<td>MOPITT</td>
<td>Measurements of Pollution in the Troposphere</td>
</tr>
<tr>
<td>MOZAIC</td>
<td>Measurement of Ozone and Water Vapour on Airbus in-service Aircraft</td>
</tr>
<tr>
<td>NDACC</td>
<td>Network for the Detection of Atmospheric Composition Change</td>
</tr>
<tr>
<td>NOAA</td>
<td>National Oceanic and Atmospheric Administration</td>
</tr>
<tr>
<td>NPP</td>
<td>National Polar-orbiting Partnership</td>
</tr>
<tr>
<td>NRT</td>
<td>Near Real Time</td>
</tr>
<tr>
<td>OMI</td>
<td>Ozone Monitoring Instrument</td>
</tr>
<tr>
<td>PIFM</td>
<td>Practical Improved Flux Method</td>
</tr>
<tr>
<td>RemoTeC</td>
<td>Remote Sensing of Greenhouse Gases for Carbon Cycle Modelling</td>
</tr>
<tr>
<td>RMS</td>
<td>Root Mean Square</td>
</tr>
</tbody>
</table>
**S5P** Sentinel-5 Precursor

**SCIAMACHY** Scanning Imaging Absorption Spectrometer for Atmospheric Chartography

**SEOM-IAS** Scientific Exploitation of Operational Missions - Improved Atmospheric Spectroscopy Databases

**SICOR** Shortwave Infrared CO Retrieval

**SMOW** Standard Mean Ocean Water

**SNR** Signal-to-Noise Ratio

**SPEC** Standard Performance Evaluation Corporation

**SSD** Spectral Sampling Distance

**SWIR** Shortwave Infrared

**SZA** Solar Zenith Angle

**TCCON** Total Carbon Column Observing Network

**TES** Tropospheric Emission Spectrometer

**TM4** Transport Model 4

**TM5** Transport Model 5

**TOA** Top Of model Atmosphere

**TROPOMI** Tropospheric Monitoring Instrument

**VIIRS** Visible Infrared Imager Radiometer Suite

**VZA** Viewing Zenith Angle

**WFM-DOAS** Weighting Function Modified-Differential Optical Absorption Spectroscopy
4 Remote Sensing of Carbon Monoxide

Carbon monoxide (CO) is an important atmospheric trace gas and in certain urban areas, it is a major atmospheric pollutant. Measurements of its global abundance improve our understanding of tropospheric chemistry and atmospheric long range transport [RD5, RD6, RD7, RD8]. Main sources of CO are combustion of fossil fuels, biomass burning, and atmospheric oxidation of methane and other hydrocarbons. Whereas fossil fuel combustion is the main source of CO at Northern mid-latitudes, the oxidation of isoprenes and biomass burning play an important role in the tropics. Due to the long lifetime of methane (CH₄), its oxidation provides a close-to uniform background on the global CO distribution. The most important sink of CO is its reaction with the hydroxyl radical OH. The CO lifetime is several weeks to several months, which makes CO a good tracer to study long range transport processes (e.g. [RD9]). Moreover, CO is one of the highest priority chemical species measured by the Copernicus Atmosphere Monitoring Service (CAMS) project [RD10, RD11].

The first spaceborne measurements of CO were performed with the MAPS (Measurement of Air Pollution from Satellites) instrument during four flights of the space shuttle between 1981 and 1999 [RD12], and with the IMG (Interferometric Monitor for Greenhouse gases) instrument onboard ADEOS (Advanced Earth Observing System) in 1996 and 1997 [RD13]. Since 2000, long-term global data sets of CO are provided by MOPITT (Measurements of Pollution in the Troposphere) instrument (e.g. [RD14]), which monitors in two spectral ranges at 2.3 µm and 4.7 µm, respectively, using correlation radiometry. At present, the operational CO MOPITT data product relies only on measurements around the fundamental 1-0 CO absorption band at 4.7 µm. This spectral range is also employed by three other spaceborne spectrometers which measure the infrared brightness of the Earth's surface and the atmosphere: (1) AIRS (Atmospheric Infrared Sounder [RD15]) launched in 2002 onboard the Aqua satellite, (2) TES (Tropospheric Emission Spectrometer [RD16]) launched in 2004 onboard the Aura satellite and (3) IASI (Infrared Atmospheric Sounding Interferometer[RD17]) onboard a series of three METOP (Meteorological Operational) satellites. Generally, these thermal infrared measurements exhibit peak sensitivity to CO in the middle troposphere and are thus well suited to study long range atmospheric transport. However, depending on spectral resolution and the thermal contrast in the lower troposphere, the measurements show also sensitivity to CO in the lower troposphere [RD18].

CO total columns with sensitivity to the tropospheric boundary layer can be inferred from sunlight reflected by the Earth atmosphere in the 2.3 µm spectral range of the shortwave infrared (SWIR) part of the solar spectrum. The first overtone 2-0 absorption band of CO is situated between 2305 nm and 2385 nm. For clear sky measurements, this spectral range is subject to little atmospheric scattering and most of the measured light is thus reflected by the Earth’s surface. SWIR measurements are therefore sensitive to the integrated amount of CO along the light path, including the contribution of the planetary boundary layer. This makes the SWIR spectral range particularly suitable for detecting surface sources of CO. Since the launch of SCIAMACHY (Scanning Imaging Absorption Spectrometer for Atmospheric Chartography [RD19]) in the year 2002 on the Envisat satellite, a continuous time series of global CO SWIR measurements is available. Moreover, first results were reported recently using MOPITT measurements in the SWIR [RD20] and a combination of SWIR and the 4.7 µm CO absorption band [RD21]. However, new space-borne instrumentation is required to ensure continuity of SWIR measurements of CO in the future. In this respect, the Sentinel 5 Precursor mission (SSP [RD22]) provides the opportunity to extend this unique long-term global data set of CO using the same type of measurement. On top of that, the major scientific objective of these mission is to participate in the development and improvement of air quality model processes and data assimilation in support of operational services, including air quality forecasting and protocol monitoring. This scientific objective requires the combination of new instrumentation with better radiometric performances and higher spatial sampling, and the development of fast and accurate retrieval algorithms.

4.1 Algorithm heritage

Several fast algorithms are used to retrieve CO column from SCIAMACHY SWIR measurements, including the Weighting Function Modified-Differential Optical Absorption Spectroscopy (WFM-DOAS) approach ([RD23] and references therein) the Iterative Maximum A Posteriori (IMAP) approach [RD24], and the Iterative Maximum Likelihood Method (IMLM) approach ([RD25], and references therein). These algorithms retrieve vertically integrated CO column density over land and above clouds over oceans. Over ocean, the surface albedo is too low to retrieve CO under clear sky conditions. For numerical efficiency, scattering by aerosols and clouds is not considered in the radiative transfer of the IMLM retrieval approach. Both Buchwitz et al. [RD26] and Gloudemans et al. [RD25] use a priori methane information to characterise the light path through the atmosphere. The different approaches allow the detection of strongly polluted areas, the seasonal variability of global atmospheric CO and the long range transport of CO. The retrievals have been compared with
chemical transport models (CTM, e.g. [RD27, RD25]), with MOPITT retrievals (e.g. [RD23, RD28]) and ground-based FTIR (Fourier Transform Infrared) measurements [RD29, RD30]. Since CO is a weak absorber and SCIAMACHY exhibits a low signal-to-noise ratio in the 2.3 µm region, the retrieval of single-sounding CO column density is subject to large retrieval noise, typically on the order of 10-100 % [RD31]. Hence, to compare SCIAMACHY retrievals with ground-based measurements, the averaging of multiple SCIAMACHY measurements is needed. De Laat et al. [RD30] demonstrated that, at the Northern latitudes and for a typical sampling area of 8° × 8°, the absolute difference between the mean IMLM SCIAMACHY CO retrievals and ground-based measurements are close to or fall within the 2σ precision of 2 × 10^{17} molec./cm^{2}. Thus on the Northern Hemisphere, SCIAMACHY can observe CO adequately for the given sampling size. At mid to high latitudes of the Southern Hemisphere, the IMLM algorithm systematically underestimates the ground based measurements by 1−5 × 10^{17} molec./cm^{2}. Reasons for this bias are currently under investigation. Next to the large sampling areas and the Southern Hemisphere bias, one major limitation of the current SWIR CO retrieval product is that for the individual soundings, no information is provided on the vertical sensitivity of the CO column retrieval, the so-called column averaging kernel. Column averaging kernels are essential information to validate SWIR CO retrievals with other independent retrievals or CTM simulations. Last but not least, it is also an important piece of information for data assimilation in air quality forecasting employing variational schemes.

Based on SCIAMACHY heritage, we present a modified SWIR CO retrieval (SICOR) approach for S5P, which can be applied to clear sky observations over land and cloudy observations over both land and oceans. The algorithm is based on the heritage of the CO retrieval algorithm for SCIAMACHY [RD25] but improves current CO retrieval approaches for cloudy and aerosol loaded atmospheres. A typical SWIR spectrum is illustrated on the top panel of Fig. 1. It shows the total transmittance of solar light along its path from the Sun to the surface to the satellite. The transmittance is simulated using the Beer’s extinction law. In the band, the relevant absorbing species are H\textsubscript{2}O, CO and CH\textsubscript{4}, with the optical depth of CO generally much smaller than those of H\textsubscript{2}O and CH\textsubscript{4}.

Clouds and aerosols affect the sensitivity of the measurement to CO in several manners: due to multiple scattering, the path length of the observed light is enhanced in the upper part of clouds. For larger optical depth the scattering layer transmits only a small fraction of incoming light and thus the atmosphere below the layer is effectively shielded. Furthermore, light can be trapped between a scattering layer and a bright surface, which can enhance significantly the light path. The developed SICOR algorithm accounts for these effects, but also takes into consideration the computational aspects of an operational data processing. An accurate treatment of clouds and aerosols in the retrieval requires the simulation of multiple light scattering which is numerically very demanding. Therefore, we employ a two-stream radiative transfer solver, which accounts for atmospheric scattering in a simplified way. Since three decades, different two-stream methods are used to describe radiative transfer in global chemistry and climate models and they are known to be stable and numerically efficient. We adapt this method for satellite remote sensing including the linearisation of the model with respect to scattering and absorption properties in the model atmosphere. In summary, the presented SICOR algorithm builds on strong heritage of SCIAMACHY CO column retrieval and radiative transfer experience of the last 30 years and additionally adds new features to the retrieval to fully exploit the potential of TROPOMI SWIR measurement for an operational data processing.

### 4.2 Carbon Monoxide level-2 requirements

To improve our present knowledge on CO on a global scale, satellite measurements of the total CO column are needed within an accuracy of < 15% and a precision with \( \leq 10 \) % even for background CO abundance and low surface reflection in the shortwave infrared spectral range [RD22, RD32]. For the CO error budget, we assume that instrument and forward model errors contribute equally to the error budget and that all error terms add up quadratically. Herewith, both instrument and forward model errors must not exceed 8 %. These CO level-2 requirements should be considered as thresholds.
Figure 1: SWIR spectral transmittance along the light path of the solar beam reflected by the Earth surface into the instrument viewing direction. Simulations are performed for viewing zenith angle (VZA) = 0°, and a solar zenith angle (SZA) = 30°, and by assuming a US standard atmospheric profile. From top to bottom, the figure shows the total transmittance, the individual transmittances due to H₂O, CH₄, and CO, respectively. Note the different y-axis scale for CO transmittance.
5 Algorithm Description

To retrieve atmospheric CO abundances from TROPOMI SWIR measurements as a near real-time (NRT) data product, we follow an algorithm concept which is summarised in Fig. 2. It is a further development of the retrieval method of Vidot et al. [RD33] and it includes a significant improvement with respect to the description of atmospheric scattering. The retrieval algorithm requires several input fields:

- The measured Earth radiance and solar irradiance spectra including noise estimate, solar and viewing geometry, and information of geo-location.
- ECMWF temperature, water vapor and pressure profiles and geo-potential height.
- An estimate of the CH\(_4\) field using a chemistry transport model, e.g. Transport Model 5 (TM5, [RD34]).
- An estimate of the CO column from a chemistry transport model (e.g. TM5).

The retrieval is performed in two steps: first as part of the SWIR preprocessing module (see Appendix B), the vertically integrated amount of methane is retrieved from a dedicated fit window of the SWIR band between 2315 and 2324 nm using a non-scattering radiative transfer model. The extent of lightpath shortening and enhancement due to atmospheric scattering by clouds and aerosols can be indicated by comparing the retrieved CH\(_4\) column with a priori knowledge. If the difference \(\Delta \text{CH}_4\) exceeds a certain threshold, observations are strongly contaminated by clouds and are rejected. In a second step, the SICOR full physics retrieval approach is used to infer CO columns from the adjacent spectral window, 2324-2338 nm. Here, the methane absorption features are used to infer information on atmospheric scattering by clouds and aerosols, which passed the cloud filter, together with the atmospheric CO and H\(_2\)O abundances, surface albedo and a spectral calibration of the reflectance spectrum. The scattering layer has a triangular height distribution of fixed geometrical thickness, and its optical depth and height are parameters to be retrieved. This step of the retrieval relies on accurate a priori knowledge of CH\(_4\) which will be provided within an accuracy of \(\pm 3\) % by a dedicated methane forecast using the TM5 atmospheric transport model. The atmospheric scattering is described by a two-stream radiative transfer model. Finally, the retrieval product consists of a CO column estimate including a column averaging kernel and a random error estimate. A detailed description of input and output fields are given in Sec. 9.

For both retrieval approaches, the non-scattering and the scattering retrieval, a forward model \(F\) is needed which describes the measurement as a function of the state of the atmosphere, namely

\[
y = F(x, b) + e_y. \tag{1}
\]

Here, vector \(y\) has the spectral measurements as its components, state vector \(x\) represents the parameters to be retrieved, \(b\) describes parameters other than the state vector that influences the measurement, and \(e_y\) is the error of the measurement. For the retrieval of CO from SWIR measurements, the forward model is non-linear in the state vector \(x\). Therefore, the inversion problem is solved iteratively employing the Gauss-Newton method, where for each iteration step the forward model is linearized by a Taylor expansion around the solution of the previous iteration \(x_o\),

\[
F(x, b) = F(x_o, b) + \frac{\partial F}{\partial x}(x_o, b)(x - x_o) + O(x^2), \tag{2}
\]

where \(O(x^2)\) indicates second and higher order contributions of the expansion. This iteration approach shows satisfying convergence properties for moderately non-linear problems [RD35]. In the following sub-sections, both the SICOR inversion approach and the forward model are described in more depth.

5.1 Forward model

We assume that the measurement vector \(y\) contains the spectral reflectance \(r_i\) sampled at spectral points \(\lambda_i\) as its components. The forward model \(F\) simulates the measurement by a spectral convolution of both the top-of-model-atmosphere (TOA) radiance and the solar irradiance with the instrument spectral response function (ISRF),

\[
I_i^{\text{TOA}} = s \ast I_{\text{TOA}} = \int s(\lambda_i, \lambda) I_{\text{TOA}}(\lambda) d\lambda, \tag{3}
\]
and a corresponding expression for the solar spectrum. Here, $s$ is the ISRF, $\lambda_i$ is the wavelength assigned to the spectral pixel $i$ of the measurement and $I_{\text{TOA}}^i$ is corresponding radiance simulated by a radiative transfer model. The reflectance is the ratio of the radiance and the solar spectrum. Thus

$$F_i = r_i = \frac{s \cdot I_{\text{TOA}}^i}{s \cdot F_o},$$

(4)

where $F_i$ is the $i$-th component of the forward model vector $\mathbf{F}$ and $F_o$ is the solar spectrum. Furthermore, we can calculated the derivative of the forward simulation with respect to a spectral shift from Eq. 4, i.e.

$$\frac{dr_i}{d(\Delta \lambda)} = \frac{(s' \cdot I_{\text{TOA}}^i)(s \cdot F_o) - (s' \cdot F_o)(s \cdot I_{\text{TOA}}^i)}{(s \cdot F_o)^2},$$

(5)

where $s' = \frac{ds}{d(\Delta \lambda)}$ is the spectral derivative of the ISRF. The derivatives of $\mathbf{F}$ with respect to all other parameters to be retrieved, depends on the particular radiative transfer model which is used to simulate $I_{\text{TOA}}^i$.

State-of-the-art radiative transfer models account for multiple scattering in multiple propagation directions (streams) including the polarisation of light. For our application, the computational effort of such simulations is high and thus, it is not feasible to use these models for operational data processing of the CO column density. So, approximation methods have to be used. In this respect, we propose to ignore atmospheric Rayleigh scattering which contributes less than $<0.15\%$ to the total signal [RD36] and to use a numerically efficient two-stream radiative transfer model where the effect of clouds and aerosols are represented in the model atmosphere by a triangular scattering layer with a center height $z_{\text{scat}}$ and a fixed width. Singly scattered light is treated exactly by the model, whereas multiple scattered light is described in an approximated manner.

**Figure 2**: Overall structure of the CO retrieval algorithm. For the structure of the SWIR pre-processing algorithm see Appendix B.
Two-stream solvers capture the essence of atmospheric scattering and represent the simplest approximation of multiple scattering in atmospheric radiative transfer. They represent a group of approximation methods solving the radiative transfer equation for scattering atmospheres where the internal radiation field is described by two, one upward and one downward, propagation direction of the radiance field. Alternatively, one can formulate these approximations in terms of the direct irradiance and the diffuse upward and downward irradiance. Due to their little numerical cost, they are commonly used to describe radiative transfer in global circulation models and weather forecast models. For the Sentinel-5 Precursor CO column retrieval, it is used to account for atmospheric scattering by clouds and aerosols, which passed the cloud filtering. Moreover, we apply the forward-adjoint perturbation theory to the two-stream solver and derive analytical expressions for the derivatives of the measurement simulation with respect to total optical depth and single scattering albedo.

The generalised two-stream method

We assume a vertically inhomogeneous atmosphere described by \( N \) homogeneous layers. Each layer is characterised by its optical properties: the optical depth \( \tau_n \), the single-scattering albedo \( \omega_n \) and the phase function \( P_n \), with layer index \( n = 1, \ldots, N \).

Generally for an arbitrary layer \( n \), the fluxes at the layer interfaces \( n \) and \( n+1 \) are constrained by internal boundary conditions as

\[
\begin{pmatrix}
S_{n+1} \\
F_{n+1}^\uparrow \\
F_{n}^\uparrow
\end{pmatrix}
= \begin{pmatrix}
a_{1,n} & 0 & 0 \\
a_{2,n} & a_{4,n} & a_{5,n} \\
a_{3,n} & a_{5,n} & a_{4,n}
\end{pmatrix}
\begin{pmatrix}
S_n \\
F_n^\downarrow \\
F_n^\uparrow
\end{pmatrix}
\]

with \( n = 1, \ldots, N \) and coefficients \( a_{1,n}, a_{2,n}, a_{3,n}, a_{4,n}, a_{5,n} \) as given in the appendix A. Here, index \( n = 1 \) describes the top of the model atmosphere and index \( N + 1 \) indicates the surface level. Furthermore, \( S_n \) is the direct solar irradiance, \( F_n^\downarrow \) and \( F_n^\uparrow \) are the downward and upward fluxes, all defined at layer interface \( n \).

Additionally, the external boundary conditions are given as

\[
\begin{align*}
S_1 &= \mu_o F_o \\
F_1^\uparrow &= 0 \\
F_{N+1}^\uparrow &= A_s (F_{N+1}^\downarrow + S_{N+1}).
\end{align*}
\]

where \( A_s \) is the surface albedo and \( \mu_o = \cos(\Theta_o) \) with the solar zenith angle \( \Theta_o \). In the following, we discuss the practical improved flux method (PIFM) of Zdunkowski et al. [RD37] for a three layer model atmosphere to simplify matters. The approach can be generalised in a straightforward manner to a \( N \)-layer model atmosphere, which is the basis for the software implementation. Combining the internal and external boundary constraints for the three layer system, we obtain the matrix equation

\[
\mathbf{M} \mathbf{F} = \mathbf{C}
\]

with the sparse matrix

\[
\mathbf{M} = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-a_{3,1} & -a_{5,1} & 1 & 0 & 0 & -a_{4,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-a_{1,1} & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-a_{2,1} & -a_{4,1} & 0 & 0 & 1 & -a_{5,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -a_{3,2} & -a_{5,2} & 1 & 0 & 0 & -a_{4,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -a_{1,2} & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -a_{2,2} & -a_{4,2} & 0 & 0 & 1 & -a_{5,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -a_{3,3} & -a_{5,3} & 1 & 0 & 0 & -a_{4,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -a_{1,3} & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -a_{2,3} & -a_{4,3} & 0 & 0 & 1 & -a_{5,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -A_r & -A_r & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.
\]
the flux vector

\[
\mathbf{F} = \left( S_1, F_{1\uparrow}^i, F_{1\uparrow}^v, F_{2\downarrow}, F_{2\downarrow}^i, F_{2\downarrow}^v, S_3, F_{3\uparrow}^i, F_{3\uparrow}^v, S_4, F_{4\uparrow}^i, F_{4\uparrow}^v \right)^T
\]

and the right hand side

\[
\mathbf{C} = \left( \mu_o F_o, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 \right)^T.
\]

For a model atmosphere consisting of \( N \) model layers, \( \mathbf{M} \) is a \( 3(N+1) \times 3(N+1) \) matrix and \( \mathbf{F} \) and \( \mathbf{C} \) are both vectors of dimension \( 3(N+1) \). Due to the block diagonal structure of matrix \( \mathbf{M} \), Eq.(8) can be solved by a sequential substitution of the linear equations.

With the flux vector \( \mathbf{F} \), we can approximate the TOA radiances in the viewing direction of the instrument. For this purpose we start with the expression

\[
\mathbf{F}_{\text{TOA}}(\mu_o) = \frac{F_{N+1}}{\pi} \exp(-\tau_{\text{tot}}/\mu_o) + \frac{1}{\mu_o} \int_0^{\tau_{\text{tot}}} d\tau J(\tau, \mu_o) \exp(-\tau/\mu_o),
\]

where \( \mu_o = \cos(\Theta_o) \) with the viewing zenith angle \( \Theta_o \), \( \tau \) indicates optical depth, and \( \tau_{\text{tot}} \) is the total optical thickness of the atmosphere. Upward and downward directions are denoted by the cosine of the zenith angle \( \mu > 0 \) and \( \mu < 0 \), respectively, with the exception of the solar beam with the cosine of the solar zenith angle \(-\mu_0\) with \( \mu_0 > 0 \). The source function \( J \) describes multiply and singly scattered light, namely

\[
J(\tau, \mu) = \frac{\omega(\tau)}{2} \int_{-1}^{1} d\mu' P(\tau; \mu, \mu') I(\tau, \mu') + \frac{\omega(\tau)}{4\pi} P(\tau; \cos \Theta_o) F_o \exp(-\tau/\mu_o).
\]

Here, the scattering phase function \( P(\tau; \cos \Theta_o) \) is a function of scattering angle \( \Theta_o \), between solar beam and instrument LOS and optical depth \( \tau \) and \( P(\tau; \mu, \mu') \) is its azimuthal average with the cosine \( \mu \) and \( \mu' \) of the zenith angle of the incoming and outgoing direction. Next, we introduce the relation between the irradiances and mean radiances,

\[
F_{\downarrow\uparrow} = 2\pi \int_{-1}^{1} d\mu \mu I(\tau, \mu) = 2\pi \mu_{\downarrow\uparrow} I_{\downarrow\uparrow}
\]

where \( I_{\downarrow\uparrow} \) is the mean radiance in the upper and lower hemisphere and with the mean inclination \( \mu_{\downarrow\uparrow} \in [0, \pm1] \). Instead of \( \mu_{\downarrow\uparrow} \) one often uses the diffusivity factor \( U_{\downarrow\uparrow} = 1/\mu_{\downarrow\uparrow} \). Furthermore, we introduce the back-scattered fraction \( \beta \),

\[
\beta(\tau, \mu) = \frac{1}{2} \int_{-1}^{0} d\mu' P(\tau; \mu, \mu') = 1 - \frac{1}{2} \int_{0}^{1} d\mu' P(\tau; \mu, \mu') \quad \text{for} \ \mu > 0,
\]

which describes the mean scattering probability for any downward direction scattered into the instrument LOS, and the mean backscattering ratio

\[
\bar{\beta}(\tau) = \frac{1}{2} \int_{0}^{1} d\mu \int_{-1}^{0} d\mu' P(\tau; \mu, \mu')
\]

(see e.g. Wiscombe and Grams, 1976, [RD38]). The expansion of the phase function in Legendre polynomials

\[
P_l(\mu) = \sum_{l=0}^{\infty} \chi_l(\tau) P_l(\mu) P_l(\mu')
\]

with coefficients \( \chi_l \), provides a corresponding expansion of the back scattering ratios,
\[ \beta(\tau, \mu) = \sum_{l=0}^{\infty} \frac{\chi(l)}{2} P_l(\mu) \int_{-1}^{1} d\mu' P_l(\mu') \]  

(18)

and

\[ \tilde{\beta}(\tau) = \sum_{l=0}^{\infty} \frac{\chi(l)}{2} \int_{0}^{1} d\mu P_l(\mu) \int_{-1}^{0} d\mu' P_l(\mu') . \]  

(19)

Using Eq. (14) and (19), we can rewrite the source function as

\[ J(\tau, \mu_v) = \frac{\omega}{2\pi} [U_1(1-\tilde{\beta}(\tau))F^1(\tau) + U_2\tilde{\beta}(\tau)F^\perp(\tau)] + \frac{\omega}{4\pi\mu_o} P(\cos(\Theta_v))S_o \exp(-\tau/\mu_v). \]  

(20)

Finally, to evaluate the integration in Eq. (12) of the scattering source function over optical depth, the upward and downward fluxes \( F^\perp \) and \( F^1 \) within the scattering layer \( i \) are approximated by their mean values,

\[ F^1(\tau) = \frac{F^1_i + F^1_{i+1}}{2} \text{ for } \tau_i < \tau < \tau_{i+1}. \]  

(21)

Thus, we obtain

\[ \int_{0}^{\tau_n} d\tau J(\tau, \mu_v) \exp(-\tau/\mu_v) = \sum_{n=1}^{N} \frac{\omega_n}{4\pi\mu_o} U_1 \sum_{n=1}^{N} \frac{\omega_n}{4\pi}(1-\beta_n) t_n(\mu_v) [F^1_n + F^1_{n+1}] + U_2 \sum_{n=1}^{N} \frac{\omega_n}{4\pi} \beta_n t_n(\mu_v) [F^\perp_n + F^\perp_{n+1}] \]  

(22)

with \( P_n = P(\tau, \cos(\Theta_v)) \), \( \omega_n = \omega(\tau) \), \( \beta_n = \tilde{\beta}(\tau) \) for \( \tau_n < \tau < \tau_{n+1} \) and the auxiliary function

\[ t_n(\mu_v) = \exp\left(-\frac{\tau_{n-1}}{\mu_v}\right)[1 - \exp\left(-\frac{\Delta \tau_n}{\mu_v}\right)] \quad \text{for } n = 1, \ldots, N. \]  

(23)

Hence, the intensity at the top of the model atmosphere can be written as

\[ I_{\text{TOA}} = RF \]  

(24)

with the so-called response vector

\[ R = \frac{1}{4\pi}(u_1, v_1, w_1, \ldots, u_{N+1}, v_{N+1}, w_{N+1}) \]  

(25)

and its coefficients

\[ u_n = \frac{\omega_n}{4\pi\mu_o} \tilde{\mu} P_n(\tilde{\mu}) \]  

\[ v_n = \frac{U_1}{4\pi} [\omega_0 \beta_n t_n(\mu_v) + \omega_{n-1} \beta_{n-1} t_{n-1}(\mu_v)] \]  

\[ w_n = \frac{U_2}{4\pi} [\omega_n (1-\beta_n) t_n(\mu_v) + \omega_{n-1} (1-\beta_{n-1}) t_{n-1}(\mu_v)] \]  

(26)

with \( t_0 = t_{N+1} = 0 \). Equation (24) is an essential step towards the linearisation of the model with respect to absorption and scattering properties of the model atmosphere, which is discussed in the following.
The forward-adjoint perturbation theory applied to the generalised two-stream method

The linearisation of the two-stream method is less evident than the linearisation of the transmission model. Due to multiple scattering, the functional dependence of $I_{\text{TOA}}$ on optical depth, single-scattering albedo and on the scattering phase function characteristics is not obvious (see Eqs. (8) and (24)). In this context, the forward-adjoint perturbation theory provides a method, which is frequently used to linearise multiple scattering radiative transfer simulations with respect to absorption and scattering properties of the model atmosphere (e.g. [RD39, RD40, RD41, RD42, RD43, RD44, RD45, RD46, RD47]). Here, we summarise the main features of the perturbation theory, which is subsequently used to motivate our approach to linearize the two-stream method.

The forward-adjoint perturbation theory relies on the formulation of the radiative transfer equation in terms of a linear transport operator, namely

$$L I = Q,$$

where

$$L = \int_{4\pi} d\Omega \left\{ \mu \frac{\partial}{\partial z} + \beta_s(r) \right\} \delta(\Omega - \tilde{\Omega}) - \frac{\beta_s(r)}{4\pi} P(r, \Omega, \tilde{\Omega}) + R_s(\Omega, \tilde{\Omega}) \delta(z) \Theta(\mu) \bigl[\Theta(-\mu) \bigl| \delta\bigr],$$

represents the transport operator of scalar radiative transfer in a plane-parallel atmosphere. Here, $z$ describes altitude, the direction $\Omega = (\theta, \varphi)$ is specified by the zenith angle $\theta$ and the azimuth angle $\varphi$. Moreover, $d\Omega = \sin\theta d\theta d\varphi$ and $\delta$ is the Dirac-delta function with

$$\delta(\Omega - \tilde{\Omega}) = \frac{1}{\sin\theta} \delta(\theta - \tilde{\theta}) \delta(\varphi - \tilde{\varphi}).$$

and $\Theta$ represents the Heaviside step function. Finally, $\beta_s$ is the extinction coefficient, $\beta_u$ is the scattering coefficient, $P$ denotes the scattering phase function and $R_s$ represents the bidirectional reflection distribution function (BRDF) that characterises the surface reflection. The intensity field $I$ is a scalar function of $z$ and $\Omega$ and the solar radiation source $Q$, which illuminates the upper hemisphere of the Earth’s atmosphere, is given as

$$Q(z, \Omega) = \mu_0 F_0 \delta(z - z_{\text{TOA}}) \delta(\Omega - \Omega_0).$$

Here, $z_{\text{TOA}}$ is the height of the upper boundary of the model atmosphere and $\Omega_0 = (\theta_0, \varphi_0)$ describes the direction of the solar beam. $\theta_0$ is the local solar zenith angle and $\varphi_0$ the local solar azimuth angle. Because surface reflection is included in the radiative transfer operator $L$ and the incoming solar source is described by the source $Q$, the radiation field $I$ is subject to the free surface boundary conditions of no incoming radiation at the top and bottom of the model atmosphere.

In the context of satellite remote sensing, one is interested in the simulation of the intensity at the top of the model atmosphere in the viewing direction of the instrument, thus $I_{\text{TOA}} = I(z_{\text{TOA}}, \Omega_v)$. Here, $\Omega_v = (\theta_v, \varphi_v)$ describes the viewing direction of the satellite instrument. For the forward-adjoint perturbation theory, it is crucial to express the observation $I_{\text{TOA}}$ by the inner product of the form

$$I_{\text{TOA}} = \langle R | I \rangle = \int_0^{z_{\text{TOA}}} dz \int_{4\pi} d\Omega R(z, \Omega) I(z, \Omega).$$

The inner product $\langle \cdot | \cdot \rangle$ is defined by the integral over height and solid angle as indicated in the equation above. The appropriate response function $R$ in (31) is defined by the measurement viewing geometry.

$$R(z, \Omega) = \delta(z - z_{\text{TOA}}) \delta(\Omega - \Omega_v).$$

Additionally, the linearisation of radiative transfer with the forward-adjoint perturbation theory requires the solution of the adjoint transfer equation.
\[ L^\dagger I^\dagger = R, \]  

(33)

where \( I^\dagger \) denotes the adjoint intensity field, \( R \) is the response function in Eq. (32) and \( L^\dagger \) is the transport operator adjoint to the forward transport operator \( L \). In principle, the adjoint transport operator is identical to the forward transport operator, apart from a reversion of all directions and it is better known as the backward formulation of the radiative transfer problem.

The adjoint intensity \( I^\dagger \) has to satisfy the boundary conditions of no outgoing adjoint intensity at the top and the bottom of the atmosphere. It allows one to simulate the measurement in Eq. (31) in an alternative way:

\[ I_{\text{TOA}} = \langle I^\dagger | Q \rangle. \]  

(34)

This equation shows the physical meaning of the adjoint field: \( I^\dagger \) can be interpreted as the importance of light within the atmosphere for a given measurement.

With the solution of the forward and the adjoint transfer problem \( I_0 \) and \( I_0^\dagger \) for an atmospheric state \( x_o \), it is possible to calculate the derivatives of \( I_{\text{TOA}} \) with respect to an optical parameter \( x \) at \( x_o \) by

\[ \frac{\partial I_{\text{TOA}}}{\partial x}(x_o) = \langle I_0^\dagger | L' I_0 \rangle + \langle I_0^\dagger | Q' \rangle + \langle R' | I_0 \rangle \]  

(35)

with the partial derivative

\[ L' = \lim_{\Delta x \to 0} \frac{\Delta L}{\Delta x}, \]  

(36)

\[ R' = \frac{\partial R}{\partial x} \text{ and } Q' = \frac{\partial Q}{\partial x} \] (e.g. [RD50, RD45]).

For an exact solution of the forward and adjoint radiative transfer equation, the expression (35) represents an exact expression for the derivative and provides an efficient calculus for the forward model Jacobian. However, the two-stream method uses several approximations as described in the previous section. In turn, the forward-adjoint perturbation theory, as described above, is not appropriate for an accurate linearisation of the TOA radiance simulation as described in Eq. (24). Due to that, we apply the perturbation theory not to the radiative transfer equation but to the matrix equation (8), which is the underlying linear equation of the two-stream approach. Here, the inner product is defined by

\[ \langle x | y \rangle = x^T y. \]  

(37)

Moreover, the adjoint matrix \( M^\dagger \) is given by its transposed, \( M^T \), and so the response vector \( R \) defines the adjoint problem

\[ M^T F^\dagger = R, \]  

(38)

where \( F^\dagger \) is the adjoint flux vector. Analogous to Eq. (35), the derivative of the TOA radiance with respect to an optical parameter \( x \) can be calculated by

\[ \frac{\partial I_{\text{TOA}}}{\partial x} = F^\dagger \langle M' | F \rangle + F^\dagger \langle C' | F \rangle + \langle R' | F \rangle \]  

(39)

with the derivatives \( M' = \frac{\partial}{\partial x} M \), \( C' = \frac{\partial}{\partial x} C \), and \( R' = \frac{\partial}{\partial x} R \). With \( C \) given in Eq. (11), the derivative \( C' \) vanishes and so Eq. (39) simplifies to

\[ \frac{\partial I_{\text{TOA}}}{\partial x} = F^\dagger \langle M' | F \rangle + \langle R' | F \rangle. \]  

(40)

In general, \( x \) represents any atmospheric parameter which influences the solution \( F \). In our case, this is the optical depth \( \Delta \tau_n \), the single-scattering albedo \( \omega_n \) and the surface albedo \( \Lambda_s \). Equation (40) can be numerically implemented in a straightforward manner, which is described in detail in Appendix A.
5.2 Inversion

5.2.1 Methane cloud filter

Assuming accurate a priori knowledge, one can use methane, which is retrieved for a non-scattering model atmosphere, to filter for the presence of optically thick water clouds. A corresponding non-scattering retrieval algorithm is described in Appendix B as part of the SWIR pre-processing. To illustrate the functionality of such a cloud filter, Fig. 3 shows the amount \( \text{CH}_4 \) above a cloud with a cloud top height \( z_{\text{cld}} \). Above 2, 5 and 10 km about 80, 50 and 20 % of the total amount of methane is located. For an optically thick cloud, the indicated fraction of methane corresponds to the methane column which can be retrieved from SWIR satellite measurements assuming a non-scattering atmosphere. The situation differs when we consider optically thin clouds. Here, light can penetrate the cloud and photon trapping between the cloud and the Earth surface may cause an effective pathlength enhancement with respect to a clear-sky situation. In this case, a non-scattering retrieval of \( \text{CH}_4 \) will overestimate the true \( \text{CH}_4 \) column.

Figure 4 shows the probability density function (PDF) of the non-scattering \( \text{CH}_4 \) retrieval error and its cumulative distribution (CPDF) for one year of GOSAT data (2010). As references we use collocated \( \text{CH}_4 \) columns from TM5 model simulations. The maximum of the PDF around small retrieval error indicates scenes, which are affected little by clouds. This maximum is present in both the ocean and land PDF. For about 80 % of all observations, the methane abundance is underestimated by the non-scattering retrieval due to the presence of optically thick clouds. Here, the ocean PDF shows a relatively high probability of retrieval errors between -20 \% and -5 \% due to the presence of low stratiform clouds over ocean. For land pixels, this type of cloudiness occurs less frequent. Finally, 20 \% of all cases show an overestimation of methane by the non-scattering retrieval indicating an effective pathlength enhancement. Overall, the figure clearly demonstrates the ability of the difference between the non-scattering retrieval and methane a priori knowledge for cloud filtering of the TROPOMI dataset.

To implement a \( \text{CH}_4 \) non-scattering cloud filter based on this principle, we employ TROPOMI SWIR measurements in the spectral range 2315-2324 nm with strong methane absorption (see Fig. 1). We fit the total column of \( \text{CH}_4 \), CO, H\(_2\)O and HDO together with a surface albedo \( A_s \), its linear dependence on wavelength, and a spectral offset of the measurements. For the inversion, a standard least squares approach is used in combination with the transmission radiative transfer model described in Section B. For data selection, we
accept all cases with a non-scattering CH₄ error $\Delta$CH₄ > 25 %, and so we filter out observations with optically thick and high clouds.

5.2.2 Tikhonov regularisation for CO column retrieval

The SWIR measurements are sensitive to the total amount of CO along the path of the measured light. Within the bounds of the measurement error, no information can be retrieved on the relative vertical distribution of CO. In the presence of clouds, the measurement loses sensitivity to the amount of CO below the cloud depending on cloud parameters. To properly account for this feature, a CO profile retrieval is required which accounts for the altitude sensitivity of the measurement. However, the inversion represents an ill-posed problem, which means that a standard least squares fit of a forward model to the measurement yields a profile which is dominated by noise contribution. Thus, the inversion requires regularisation.

For this purpose, we employ the Tikhonov regularisation technique [RD51, RD52] embedded in a Gauss-Newton iteration scheme. For each iteration step, the solution is given by the least square solution with an additional side constraint,

$$\hat{x} = \min_x \left\{ \| S^{-1/2} (F(x) - y) \|^2 + \gamma^2 \| L_1 x \|^2 \right\}.$$  \hspace{1cm} (41)

$S_y$ is the error covariance matrix of the measurement $y$, where we assume uncorrelated measurement errors. The state vector $x$ contains the CO profile $x_{CO}$ which is expressed relative to a reference profile $\rho_{CO}^{ref}$, thus

$$x_{CO} = \rho_{CO} / \rho_{CO}^{ref}.$$  \hspace{1cm} (42)

For the operational implementation of the algorithm, TM5 model fields are used to extract an adequate CO reference profile. Besides the CO column, the state vector consists of a water vapor column density for two isotopes, $c_{H_2O}$ and $c_{HDO}$, a ground surface albedo $A_{surf}$, and its linear dependence on wavelength $\Delta A_{s}$, an effective cloud center height $z_{scat}$ and an effective cloud optical depth $\tau_{scat}$. Furthermore, a spectral shift $\Delta \lambda$ is fitted to account for spectral calibration errors of the measurement, hence

$$x = (x_{CO}, c_{H_2O}, c_{HDO}, A_{surf}, \Delta A_{s}, z_{scat}, \tau_{scat}, \Delta \lambda).$$  \hspace{1cm} (43)

$L_1$ describes the operator for the vertical derivative of the CO column.

$$L_1 x = \frac{d x_{CO}(z)}{dz}.$$  \hspace{1cm} (44)
For an equidistant model grid with layer thickness $\Delta z$, the discrete representation of the first derivative of the CO column is

$$L_1 x = \frac{1}{\Delta z} \begin{pmatrix} 1 & -1 & \cdots & -1 \\ \vdots & \ddots & \ddots & \vdots \\ 1 & -1 \end{pmatrix} x_{\text{CO}}.$$  \hspace{1cm} (45)

The choice of the regularisation parameter $\gamma$ is of crucial importance for the inversion. If $\gamma$ is chosen too large, the noise contribution to the solution of the measurement is low, but the least squares residual norm deviates significantly from its minimum. On the other hand, if $\gamma$ is chosen too small, the measurement is fitted well but the solution norm is high, and so, the solution is overwhelmed by noise. The part of the profile that is dominated by noise defines the effective null-space of the problem. Thus, $\gamma$ should be chosen such that the two minimisations are well balanced. For atmospheric profile retrievals, it has been demonstrated in several studies that an appropriate value for $\gamma$ can be found using the L-curve method [RD53, RD54]. However, the method becomes less stable when the information on the vertical distribution of the trace gas is limited as in our case, where about one independent piece of information of the CO profile can be retrieved. Therefore, we decided to regularize the solution in Eq. (41) with respect to its degree of freedom for signal (DFS) of the retrieved CO profile. For Eq. (41), this corresponds to a regularisation parameter $\gamma \to \infty$ (e.g. [RD55]).

For a forward model which is linearized around the state vector $x_o$,

$$F(x, b) = F(x_o, b) + K(x - x_o),$$  \hspace{1cm} (46)

the solution of Eq. (41) is given by:

$$\hat{x} = G\tilde{y}$$  \hspace{1cm} (47)

with the gain matrix

$$G = \left( K^T S_y^{-1} K + \gamma^2 L_1^T L_1 \right)^{-1} K^T S_y^{-1}$$  \hspace{1cm} (48)

and $\tilde{y} = y - F(x_o) + K x_o$. Here, $K = \partial F(x_o)/\partial x$ is the Jacobian of the forward model $F$ with respect to state vector $x$ at $x_o$. Due to the regularisation, the retrieved profile $\hat{x}$ is a smoothed version of the true profile $x_{\text{true}}$. The smoothing can be characterised by the averaging kernel,

$$A = \frac{\partial \hat{x}}{\partial x_{\text{true}}} = GK,$$  \hspace{1cm} (49)

and so, the retrieved state vector $\hat{x}$ in Eq. (47) can be written as:

$$\hat{x} = Ax_{\text{true}} + e_x.$$  \hspace{1cm} (50)

where $e_x = Ge_y$ represents the error on the retrieved CO profile caused by the error $e_y$. For a proper error characterisation of the inversion, the retrieval error covariance matrix $S_x$ is needed. This can be calculated from the measurement covariance matrix $S_y$ by

$$S_x = GS_y G^T.$$  \hspace{1cm} (51)

Because of the non-linearity of the forward problem, the inversion is solved iteratively and so a convergence criterion is needed to terminate the iteration. For the CO retrieval, we consider the difference in $\chi^2$ between two consecutive iteration steps with

$$\chi^2_n = \frac{1}{N} \sum_{i} \frac{(y_i - F_i(x_n))^2}{e_i^2}.$$  \hspace{1cm} (52)
Here \( \varepsilon_i \) is the standard deviation (error) of the \( i \)th measurement, \( \varepsilon_i^2 = S_{x_i} \). Convergence is achieved when \(|\chi^2_n - \chi^2_{n-1}| < \epsilon \) for a given choice of \( \epsilon \). The value of \( \epsilon \) can only be determined in a reliable manner using real measurements during the mission consolidation phase. Furthermore, a minimum number of iterations is always performed. If convergence is not achieved within a certain number of iterations, the inversion is terminated and the measurement is rejected.

Delivering the retrieved profile as a final product is misleading because it suggests height information which is not provided. Therefore, we vertically integrate the retrieved profile as:

\[
\hat{c} = C\hat{x} = \int \hat{\rho}(z)dz, \tag{53}
\]

where \( \hat{c} \) is the vertically integrated amount of CO and \( C \) represents the corresponding integration. Using this formulation, we can characterise the effect of regularisation on the CO column by:

\[
\hat{c} = A_c x_{\text{true}} + e_c, \tag{54}
\]

where \( A_c = CA \) is the column averaging kernel, and \( e_c \) is the error on the retrieved column. The retrieval noise on the retrieved column is given by the standard deviation

\[
\sigma_c = \sqrt{CS_x^T}. \tag{55}
\]

In this manner, all diagnostic tools of the retrieval of the state vector \( \hat{x} \) can be transformed to the corresponding diagnostics for the retrieved column \( \hat{c} \).

Borsdorff et al. [RD55] showed that for the definition of the CO state vector \( x_{\text{CO}} \) in Eq. (42) and for a regularisation parameter \( \gamma \to \infty \), the solution of the Tikhonov minimisation problem (41) is identical to an unregularised least squares approach,

\[
\hat{x} = \min_x ||S_y^{-1/2}(F(x) - y)||^2, \tag{56}
\]

where the state vector contains the total CO column instead of the vertically resolved CO profile. The numerical implementation of the least squares approach is beneficial with respect to numerical stability and computational cost and so it is used for the presented algorithm.

For this purpose, the Jacobian of the measurement simulation with respect to the CO column \( c \) is calculated by means of a scaling of the reference profile using the derivatives with respect to an altitude resolved CO profile \( \rho \), namely

\[
K_{col} = \frac{\partial F_i}{\partial c} = \sum_j K_{prof}^j \rho_j^{ref} c_{ref}^j, \tag{57}
\]

Here, \( K_{prof}^j = \frac{\partial F_i}{\partial \rho_j} \) describes the profile Jacobian and \( \rho_j \) represents the CO subcolumn of a model layer \( j \). A direct analytical calculation of the derivative \( \frac{\partial F_i}{\partial c} \) is not possible due to scattering and the temperature dependence of the trace gas absorption. \( K_{col} \) together with the derivative of the measurements with respect to the other elements of the state vector defines the least squares Jacobian \( K_{lsq} \). So for a linearized forward model, the least squares solution in Eq. (56) is

\[
\hat{x} = G_{lsq}y \tag{58}
\]

with the gain matrix

\[
G_{lsq} = (K_{lsq}^T S_y^{-1} K_{lsq})^{-1} K_{lsq}^T S_y^{-1}. \tag{59}
\]

To infer the column averaging kernel defined in Eq. (50) from the least squares solution, we follow the approach by [RD55]: We extract the gain vector \( \mathbf{g}^{col} \) from the gain matrix of the least-squares fit \( G_{lsq} \) which belongs to the fitted CO column. Interpreting the profile scaling approach as a particular case of a regularised profile retrieval using Tikhonov regularisation of the first order with an infinite strong regularisation, [RD55] showed that the total column averaging kernel can be calculated as

\[
A^{col} = \mathbf{g}^{col} K^{prof}. \tag{60}
\]
Here, the spatial sampling of the column averaging kernel depends on the corresponding sampling of $K_{\text{prof}}$. For the full mathematical proof, the reader is referred to [RD55].

The column averaging kernel allows us to quantify that part of the CO column which cannot be inferred from the measurement using a given reference profile, namely

$$e_n = (C - A)c \rho$$  \hspace{1cm} (61)

This is the so-called null space contribution of a profile $\rho$ and is also known as smoothing error of the retrieval [RD35]. [RD55] have shown that the null space contribution of the reference profile $\rho_{\text{ref}}$ is zero, $A_{\text{col}}^{\text{ref}} = C \rho_{\text{ref}}$.

So for the particular case that the relative profile used for the scaling is correct, the retrieved column can be interpreted as an estimate of the true column, as expected. In all other cases, equation (60) and (61) have to be accounted for a proper interpretation of the retrieval product.

### 5.2.3 Tikhonov regularisation for stability

The solution of the minimisation problem (Eq. (56)) is vulnerable to state variables to which the spectrum is insensitive. For example, in case of a measurement over a scene which is fully overacted by a optically thick cloud, the measurement has no sensitivity to the surface albedo. In this case, one or more eigenvalues of $K^T K$ approach zero, leading to extremely large values in $G$. Therefore, a side constraint similar to that of Eq. (41) is desirable. However, a first order derivative matrix $L_1$ with this state vector makes no sense, because the state vector does not contain vertically resolved columns.

Instead, we take a regularisation matrix based on the identity matrix. In order to have the least effect on our main product, the CO column, we restrict regularisation to those state variables to which the modeled reflectance can be insensitive. Those are the state variables related to the scattering layer and the surface albedo. So,

$$\hat{x} = \min_x \left\{ ||S_{\gamma}^{-1/2}(F(x) - y)||^2 + \gamma^2 ||Wx||^2 \right\}$$  \hspace{1cm} (62)

with state vector $x$ defined as

$$x = (c_{\text{CO}}, c_{\text{H}_2\text{O}}, c_{\text{HDO}}, A_{\text{surf}}, \Delta A_s, \xi_{\text{scat}}, \tau_{\text{scat}}, \Delta \lambda) .$$  \hspace{1cm} (63)

and $W$ is defined as

$$W_{ij} = 1 \quad \text{for} \quad i = j \in (A_{\text{surf}}, \Delta A_s, \xi_{\text{scat}}, \tau_{\text{scat}}) \quad = 0 \quad \text{otherwise}.$$  \hspace{1cm} (64)

The gain matrix is defined similarly to Eq. (48):

$$G = (K^T S_{\gamma}^{-1}K + \gamma^2 W^TW)^{-1} K^T S_{\gamma}^{-1} .$$  \hspace{1cm} (65)

and the averaging kernel $A = GK$ differs from the identity matrix for $\gamma \neq 0$. Because of the zero elements in $W$ for the carbon monoxide column $c_{\text{CO}}$, the null-space contribution for the total carbon monoxide column can be neglected.

### 5.2.4 Step control

Tikhonov regularisation protects the inversion for variables to which the measurement is insensitive. However, instability can also originate from strong nonlinearities in the forward model. The inversion algorithm requires a linearized forward model, assuming that the nonlinear behaviour can be treated by using an iterative approach. However, a system with too strong nonlinearities will not converge. The risk of an unstable inversion is largest during the first few iteration because of large step sizes.

This issue is treated by introducing a step size reduction factor $\lambda$ (also explained by Butz et al. [RD56]). The state retrieved during iteration $i$ is the weighted mean of the state retrieved during iteration $i-1$ and the calculated state during iteration $i$, where the reduction factor $\lambda$ determines the weights:
\[
x_i = \frac{1}{1 + \lambda} x_i + \frac{\lambda}{1 + \lambda} x_{i-1}
\]  \hspace{1cm} (66)

Each inversion step is evaluated by a forward simulation. A step size is accepted if either the \(\chi^2\) decreases with respect to the previous iteration, or it increases slightly. When a step is rejected, \(\lambda\) is increased and the new state is calculated with Eq. (66). When a step is accepted, \(\lambda\) is reduced and the next iteration is performed.

The idea of this reduction of \(\lambda\) is that the step is most strongly reduced at the first iterations, where \(x_i\) is further away from the minimum of the cost function and the proposed step size of a linear inversion is generally large.

### 5.2.5 Unphysical values

A radiative transfer model cannot handle unphysical values for the state variables (for instance, a negative cloud optical depth). However, the inversion module based in a linearized forward model may provide negative and so unphysical values in the state vector. This should be prevented to warrant a proper and stable performance of the iteration.

To avoid the risk of unphysical values of the state vector, two interceptor mechanisms are added to the algorithm:

1. After each iteration, we add the null-space contribution \((I - A)x_{a\text{ prior}}\) with an a priori estimate of the state vector \(x_{a\text{ prior}}\) during the iteration. This contribution is not added to the final result.

2. If a state vector still contains unphysical values for a given iteration, the last step size is reduced such that this state becomes physical (for instance, zero cloud optical depth). In such a case the state variable that tended towards unphysical values is left out of the inversion for a few (typically 3) iterations. After those iterations, that variable is returned to the inversion, allowing it to tend to physical (positive) values again.

This last interception means an ad hoc regularisation of the inversion for a few iterations. It affects only the search path of the mineralization and the interpretation of the retrieval product does not change if in the last iteration the initial definition of the state vector is used. However, it is possible that convergence is achieved at an iteration in which one or more variables are excluded from the inversion. Two reasons may cause this result: First, the true value of the state vector is close to zero and negative values are caused by noise propagation in the inversion. In this case, setting the parameter to zero does not significantly affect the CO retrieval results. Second, negative values are caused by forward model errors and measurement errors. Here, the interpretation is not straightforward and depends on the particular error source. In Sec. 7.3, we discuss the effect of forward model errors in more detail in the context of the a geo-physical China measurement ensemble.

### 5.3 Numerical Implementation and Data Product

Based on SCIAMACHY heritage and numerical experience, we calculate the gain matrix \(G\) (see Eq. 65) using Cholevsky decomposition, which is roughly twice as efficient as the LU decomposition [RD57] with good numerical stability.

Furthermore, the numerical performance of the retrieval can be improved by optimising the vertical layering of the model grid used for the radiative transfer simulation. The layering determines the numerical speed of the simulation because it is directly linked to the dimension of matrix equation (8). To optimize the numerical efficiency of the two-stream solver, we first calculate absorption cross sections on a 1 km vertical grid to account for pressure and temperature dependence of atmospheric absorption in the SWIR spectral range. Because Rayleigh scattering is ignored in the radiative transfer simulation, all scattering takes place within the atmospheric scattering layer and between the scattering layer and the model Earth surface. Thus for the simulation, we can combine the atmosphere layers above the scattering layer and between this layer and the surface to one layer each by integrating the optical depth (see Fig. 5). This significantly reduces the number of vertical layers in the radiative transfer simulation (typically to less than 10), depending on the number of internal layers covered by the scattering layer. Although the radiative transfer simulations are performed on an aggregated vertical grid, the forward model Jacobian can be calculated on the original 1 km vertical grid taking into account the layer integration of optical depth.
Figure 5: Optimisation of the vertical layering of the two-stream radiative transfer simulation based on the initial grid. The internal grid is needed to account for pressure and temperature dependence of atmospheric absorption. For the radiative transfer, layers above and below the scattering layer can be combined to one layer each, indicated by the blue areas.

5.4 Molecular optical properties

The molecular absorption cross sections of CH$_4$, H$_2$O, HDO and CO are pre-calculated as a function of pressure, temperature and for a spectral sampling of $5 \cdot 10^{-3}$ cm$^{-1}$ deploying the Scientific Exploitation of Operational Missions - Improved Atmospheric Spectroscopy Databases (SEOM-IAS) which is an ESA Project that revised the line list parameters/absorption cross sections of O$_3$, CO, CH$_4$, H$_2$O, HDO, and SO$_2$ with the objective to improve the quality of the Sentinel-5P data products ([https://www.wdc.dlr.de/seom-ias/](https://www.wdc.dlr.de/seom-ias/)) [RD58]. The CH$_4$ and H$_2$O line lists of SEOM-IAS were tested by fitting atmospheric spectra recorded by FTIR spectrometry, resulting in significantly improved residuals in spectral sections dominated by CH$_4$ and H$_2$O compared to HITRAN 2012 [RD59]. Furthermore, SEOM-IAS achieves the best spectral fit quality (root-mean-squared (rms) differences between simulated and measured TROPOMI SWIR spectra) in comparison with the HITRAN 2008, HITRAN 2012, and HITRAN 2016 spectroscopic databases [RD60]. SEOM-IAS is significantly reducing the bias between TROPOMI CO and the validation measurements of the TCCON Network [RD60] in comparison with the previously used HITRAN 2008 spectroscopic database [RD61].

The cross sections are stored in lookup table files (NetCDF format). Here, the format of the lookup table files are identical for the CO and CH$_4$ retrieval algorithms to guarantee flexibility with respect to updates of the SWIR spectroscopy. For each individual retrieval with a corresponding pressure and temperature profile, cross sections are calculated by a bilinear interpolation on the pressure and temperature grid of the lookup table.

Because calculation time must be minimised, the wavenumber grid on which the forward model calculations are performed must be rather coarse, possibly too coarse to properly resolve all absorption lines, especially at low pressures. When sampling on a selection of wavenumbers, it is possible that important absorption lines are skipped by the sampling, which is obviously not desirable for the retrieval.

The challenge is to represent the entire spectroscopy of the molecules with a limited number of wavenumbers. Therefore, effective cross sections should be found for each model wavenumber point. Together, all model wavenumber points should represent the entire spectroscopy of the molecules.

As a first definition of effective cross section we considered an average with a triangular weight function around the sample wavenumber $k$:
where the triangular weight function \( T_i(k) \) is defined as

\[
T_i(k) = \begin{cases} 
    k - k_{i-1} & \text{for } k_{i-1} < k < k_i \\
    k_{i+1} - k & \text{for } k_i < k < k_{i+1} \\
    0 & \text{otherwise}
\end{cases}
\]

assuming an equidistant wavenumber grid \( (k_i) \) in the model. This approach, however, tends to overestimate the representative cross section systematically. The reason is that not the cross sections themselves should be averaged, but their effects on the measured quantity, the reflectance. The reflectance is a convex \( (e^{-x}) \) shape function of the cross sections, and therefore, the reflectance for an average cross section is lower than the average reflectance for the cross sections. Therefore, conventional triangle convolution tends to overestimate the absorbing power of the molecules, leading to underestimations in the retrievals.

Weighted averaging of the cross section using the reflectance as a weighting kernel is not feasible because it requires forward model simulations on a fine wavenumber grid. A pragmatic solution is to tune down the average cross sections by using the generalised mean rather than that of the arithmetic mean, i.e.

\[
\sigma_i = \sqrt{\frac{\int T_i(k)\sigma(k)dk}{\int T_i(k)dk}},
\]

This introduces a tuning parameter \( m \) that theoretically can be any value between \(-\infty\) to \(\infty\). If \( m = 1 \), the arithmetic mean of Eq. (67) is reproduced. The limit of \( m \rightarrow 0 \) describes the geometric mean. This geometric mean results in much too low values for the cross sections. Therefore, the best value of \( m \) should be between 0 and 1. The optimal value of \( m \) depends on several factors, such as the concentrations of the absorbers in the atmosphere. As baseline, \( m \) is taken to be 0.85, which is obtained from a test with a standard cloudless atmosphere.

5.5 Micro-physical properties of the scattering layer

The micro-physical properties of the scattering layer are of minor relevance for this study. This has been demonstrated by using the single-scattering albedo and the phase function for a highly absorption dust aerosol (type HA DUST 14 of Torres et al. [RD62]), although in all test scenarios measurements were calculated for water and ice clouds. Because of the minor relevance of the micro-physical properties of the scattering layers, these properties have been simplified to a single value for the single-scattering albedo and the asymmetry parameter and a simplified wavelength-dependence of the extinction optical thickness based on the Ångström parameter:

\[
\tau(\lambda) = \tau(\lambda_0) \left( \frac{\lambda}{\lambda_0} \right)^{-\alpha}.
\]

Here, \( \lambda_0 \) is a fixed reference wavelength and \( \alpha \) is the Ångström parameter. The phase function for single scattering is approximated with the Henyey-Greenstein phase function, depending only on the asymmetry parameter. The effect of different scattering characteristics will be investigated in a later phase.

5.6 State vector, ancillary parameters and a priori knowledge

Table 1 summarizes the state vector definition for the different retrieval steps, which are described at the beginning of this section. A priori knowledge about the following atmospheric parameter is needed:

1. CO profile.
2. CH$_4$ profile.
Table 1: Setup of state vector $x$. Here, CO, CH$_4$, H$_2$O and HDO indicates the total column retrieval of the trace gases, $A_s$ and $\Delta A_s$ are the Lambertian surface albedo and its slope, $z_{\text{scat}}$ represents the center height of the scattering layer, $\tau_{\text{scat}}$ the total optical thickness of the layer, and $\Delta \lambda$ is a spectral measurement offset of the measurement.

<table>
<thead>
<tr>
<th>Methane pre-fit</th>
<th>CO final retrieval</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO</td>
<td>x</td>
</tr>
<tr>
<td>CH$_4$</td>
<td>x</td>
</tr>
<tr>
<td>H$_2$O</td>
<td>x</td>
</tr>
<tr>
<td>HDO</td>
<td>x</td>
</tr>
<tr>
<td>$A_s$</td>
<td>x</td>
</tr>
<tr>
<td>$\Delta A_s$</td>
<td>x</td>
</tr>
<tr>
<td>$z_{\text{scat}}$</td>
<td>x</td>
</tr>
<tr>
<td>$\tau_{\text{scat}}$</td>
<td>x</td>
</tr>
<tr>
<td>$\Delta \lambda$</td>
<td>x</td>
</tr>
</tbody>
</table>

3. H$_2$O profile.
4. HDO profile.
5. Geo-location of the SWIR measurement.
6. Temperature profile.
7. Surface elevation on SWIR pixel level.

When temperature is given on a coarse spatial resolution (e.g. the $1^\circ \times 1^\circ$ ECMWF spatial resolution), one needs the surface pressure of the SWIR pixel additionally to allocate the temperature profile in height. Alternatively, geo-potential height of the SWIR ground pixel can be used for this purpose. We assume a triangular scattering layer with a fixed FWHM of 2.5 km. Because the retrieval method does not rely on any statistical regularisation method, no estimate of the statistical variability (e.g. covariance) of the a priori information is needed.

5.7 Data product

Although several atmospheric parameters are retrieved, parameters like the height and optical depth of the scattering layer and the Lambertian surface albedo are effective parameters and so, they cannot be interpreted as an estimate of the truth. Thus for the level-2 data product of this algorithm, we focus only on the CO relevant parameters which are:

1. The retrieved total column of CO.
2. The corresponding column averaging kernel.
3. The statistical noise estimate.

Figure 6 shows an example of the CO retrieval product over land for a cloudy atmosphere and a surface albedo $A_s = 0.05$. It shows the difference between the true CO column and retrieved CO column corrected for retrieval noise (left panel), the 1-$\sigma$ estimate of the corresponding retrieval noise (middle panel) and the column averaging kernel (right panel). The retrieval scales the true relative CO profile. The left panel shows only small differences ($\leq \pm 1\%$) between the true and the retrieved profile for cloud fractions $f_{\text{cloud}} = 0$–1. The differences are due to the description of atmospheric scattering in the retrieval. The shielding of the atmosphere below the cloud does not reduce the retrieved CO column as a function of increased cloud fraction because the sensitivity of the measurement with respect to the CO abundance above the cloud is used to retrieve the total CO column by an appropriate scaling of the relative profile. At the same time, the retrieval noise on the CO column decreases due to the gain in SNR for increasing cloud coverage. The change of retrieval sensitivity...
Figure 6: Example of the CO data product. The SWIR measurements are simulated for a scene partially covered by a water cloud between 2 and 3 km with optical depth $\tau_{\text{cld}} = 30$ and a surface albedo $A_s = 0.05$. Left panel: Difference $\Delta CO$ between the true CO column and the retrieved CO column as function of cloud fraction $f_{\text{cld}}$. Middle panel: $1\sigma$ retrieval noise estimate as function of cloud fraction $f_{\text{cld}}$. Right panel: column averaging kernel as function of altitude for different cloud fractions. With cloud coverage is clearly illustrated by the column averaging kernels shown in the right panel of Fig. 6. Here, the color of lines indicate the cloud fraction. When the cloud fraction is greater than zero, the column averaging kernel starts to increase above the top of the retrieved cloud height and at the same time the retrieval sensitivity decreases below the cloud such that the net effect on the retrieval CO column nearly cancels out.
6 Common aspects with other algorithms

Both, the SICOR algorithm and the operational TROPOMI CH₄ RemoTeC algorithm [RD2] are implemented in one common processing structure as discussed in Sec. 8.1. This makes it possible to optimize synergistic use of data structures as e.g. a common lookup table for atmospheric absorption cross sections and input meteo fields.
Table 2: Microphysical properties of water and ice clouds: $n(r)$ represents the size distribution type, $r_{\text{eff}}$ and $\nu_{\text{eff}}$ are the effective radius and variance of the size distribution, $n = n_r - in_i$ is the refractive index. The ice cloud size distribution follows a power-law distribution as proposed by [RD1].

<table>
<thead>
<tr>
<th></th>
<th>water clouds</th>
<th>ice clouds</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n(r)$</td>
<td>gamma</td>
<td>$(r/r_1)^{-3.85}$</td>
</tr>
<tr>
<td>$r_{\text{eff}}$ [µm]</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>$\nu_{\text{eff}}$</td>
<td>0.10</td>
<td>-</td>
</tr>
<tr>
<td>$n_r$</td>
<td>1.28</td>
<td>1.26</td>
</tr>
<tr>
<td>$n_i$</td>
<td>$4.7 \cdot 10^{-4}$</td>
<td>$2.87 \cdot 10^{-4}$</td>
</tr>
</tbody>
</table>

7 Error analysis

To assess the performance of the CO retrieval algorithm, we have generated simulated measurements for various test cases using the S-LINTRAN version 2.0 radiative transfer model [RD63]. S-Lintran is a scalar plane-parallel radiative transfer model based on the discrete ordinate method [RD64]. It fully accounts for multiple elastic scattering by particles (clouds and cirrus) and molecules, and the interaction with the reflecting Earth surface. This version of LINTRAN does not account for the polarisation properties of light, which are of minor relevance for the purpose of this study. The optical properties of water clouds are calculated using Mie theory with microphysical cloud properties given in Table 2. For ice clouds the ray tracing model of Hess et al. [RD65, RD66] is employed assuming hexagonal, columnar ice crystals randomly oriented in space. Finally, cirrus and water clouds are described by a cloud top and base height, and a cloud optical thickness. Furthermore, we assume that cirrus fully cover the observed ground scene, whereas for water clouds, partially cloud coverage is possible utilising the independent pixel approximation [RD67] for the simulation.

The radiance spectra are superimposed by measurement noise using the TROPOMI noise model of [RD68]. For this purpose, we assumed an observed ground scene of $7 \times 7$ km² and a telescope aperture of $6 \times 10^{-6}$ m². The optical transmittance of the instrument is adjusted such that, for a spectral sampling of 0.1 nm, a signal to noise ratio of 120 is achieved in the continuum of the spectrum for a dark reference scene (surface albedo $A_s = 0.05$, viewing zenith angle VZA = 0° and SZA = 70°). The instrument noise is estimated to be 346 electrons per signal. It includes noise due to thermal background, dark current of the detector, readout noise and analog-to-digital converter noise.

To characterise the retrieval performance, we consider separately the retrieval noise $\sigma_{\text{CO}}$ of the CO column (see Eq. (55)) and the CO bias defined by

$$
\Delta \text{CO} = C_{\text{CO}} - G_{\text{CO},e} - A_{\text{CO}} \rho_{\text{CO, true}},
$$

where $A_{\text{CO}}$ is the CO column averaging kernel from Eq. (54), $\rho_{\text{CO, true}}$ is the true CO profile from the test ensemble, $C_{\text{CO}}$ represents the retrieved CO column, and $G_{\text{CO},e}$ estimates the noise of the CO column using the CO relevant contribution of the gain matrix $G_{\text{CO}}$.

First, we analyse generic ensembles of simulated measurements by systematically varying the atmospheric state and measurement geometry. Second, we investigate the retrieval performance for a geo-physical measurement ensemble comprising measurement simulations over China. Here, we combine the atmospheric trace gas abundances from chemical transport simulations with spatially and temporally allocated observations of the MODIS AQUA satellite of cirrus and water cloud occurrences and surface albedo.

7.1 Performance analysis for generic scenarios

For all generic cases, which we describe in more detail below, we assume the same atmospheric model. We assume the US standard atmosphere [RD69] for the profiles of dry air density, pressure, water and CO. The CH₄ profile is taken from the European background profile of the TM4 + CHIMERE chemical transport model [RD70], interpolated to the same pressure grid and converted from mixing ratios to densities using the air densities from the US standard atmosphere. We separated the water profile into individual profiles for the three isotopic components with absorption features in the TROPOMI SWIR range: H₂¹⁶O, H₂¹⁸O and HDO.

In Fig. 7, we show the (normalised) atmospheric H₂O, CO, CH₄ and air density profiles, the temperature profile and the profiles for the HDO and H₂¹⁸O depletion. For all cases A-F, the measurement spectra are
simulated for a water vapor column of $4.79 \cdot 10^{22}$ molec./cm$^2$, a CH$_4$ column of $3.68 \cdot 10^{19}$ molec./cm$^2$, and a CO column of $2.38 \cdot 10^{18}$ molec./cm$^2$. In addition to this model atmosphere, we defined different cases to test the influence of clouds, aerosol, surface reflection and solar geometry. All simulations assumed a fixed sensor viewing zenith angle of 40° and an azimuthal difference between the sensor and the sun of 60°. These test cases are described below and an overview of the test cases is given in Table 3.

Case A is defined to study the performance of the algorithm under clear sky conditions. The ensemble contains simulated measurements without clouds and aerosol, but with a variable surface albedo (in the SWIR range) between 0.03–0.6 (in 10 non-equidistant steps i.e. $A_s = 0.03, 0.05, 0.075, 0.10, 0.15, 0.20, 0.25, 0.3, 0.4, 0.5, 0.6$) and a variable SZA between 0.0–70.0° in 14 steps of 5°.

Case B is defined to study the effect of clouds on the retrieval algorithm. The ensemble contains simulated measurements for a fixed cloud optical thickness ($\tau_{\text{cloud}}$) of 5, 10, 30 or 50 and a variable cloud top height between 1–8 km (in 7 steps of 1 km) and a variable cloud fraction between 0.0–1.0 (in 10 steps of 0.1). The clouds have a constant geometrical thickness of 1 km. The surface albedo is fixed at 0.05 and the SZA is fixed at 50°.

Case C allows us to study the effect of "photon trapping". We have varied the surface albedo between 0.03–0.6
(in 10 steps) together with the cloud fraction (between 0.0–1.0 in 20 steps) for four different water clouds: a cloud optical thickness $\tau_{\text{cld}}$ of 2 and 5, combined with a cloud top height of 2 and 5 km. The geometrical cloud thickness was 1 km and the SZA was fixed at 50°.

Case D is to study the impact of different aerosol loads. We have varied the surface albedo between 0.03–0.6 (in 10 non equidistant steps, same as for Case A) together with the aerosol optical thickness (AOT) at 550 nm between 0.0–1.0 in 20 steps. We assumed two different scenarios: one scenario contained a sulfate-type aerosol in the boundary layer between 0–2 km, the other scenario contained an urban-type aerosol between 4–5 km. Both scenarios also contained a background aerosol layer with an AOT of 0.1 (at 2300 nm) at the surface, exponentially decreasing with altitude $z$ with a factor of the form $0.9998^z$ with $z$ in meters. The SZA was again fixed at 50°.

Case E is defined to study the impact of ice (i.e. cirrus) clouds. We have varied the surface albedo between 0.03–0.6 (in 10 steps) and the cirrus optical depth at 2300 nm between 0.0–1.0 in steps of 0.05. The cirrus cloud fraction was 100% for a layer between 9 and 10 km. The SZA was again fixed at 50°.

Case F is defined to study the effect of multiple layer clouds and the interaction with surface reflection. We have considered a cirrus cloud between 9–10 km with optical depth between 0.0–1.0. and a water cloud between 2–3 km with a cloud fraction between 0.0–1.0 (in 10 steps of 0.1). The water cloud has an optical depth of 5. Moreover, no aerosols are present and we consider four different surface albedos of 0.05, 0.10, 0.30, and 0.60. The SZA is fixed at 50° and VZA is 40°.

### 7.2 Performance Analysis for generic scenarios

To test the iterative approach, the retrieval is initialised with an atmospheric state, which differs from the truth by 7.2 Performance Analysis for generic scenarios

Case F is defined to study the effect of multiple layer clouds and the interaction with surface reflection. We have considered a cirrus cloud between 9–10 km with optical depth between 0.0–1.0. and a water cloud between 2–3 km with a cloud fraction between 0.0–1.0 (in 10 steps of 0.1). The water cloud has an optical depth of 5. Moreover, no aerosols are present and we consider four different surface albedos of 0.05, 0.10, 0.30, and 0.60. The SZA is fixed at 50° and VZA is 40°.

Both retrieval bias and retrieval noise are within the level-2 requirement specification for CO (see Sec. 4.2).

The case B retrieval performance for water clouds is shown in Fig. 9. Here, the grey areas indicate observations which are rejected by the methane cloud filter using a threshold value of $\Delta$ CH$_4$ of 25%. The data filtering depends on cloud optical depth and thus on the transmission of the cloud layer. For example, measurements with a cloud top height at 3 km are rejected by the cloud filter for a cloud coverage of $f_{\text{cld}} \geq 0.6$ for $\tau_{\text{cld}} = 5$, of $f_{\text{cld}} \geq 0.4$ for $\tau_{\text{cld}} = 10$, and of $f_{\text{cld}} \geq 0.3$ for $\tau_{\text{cld}} = 30$ and 50. For observations accepted by the cloud filter, the retrieval bias is within ±1%. This illustrates the capability of the retrieval to account for water clouds above dark surfaces. For all successful retrievals of case B, the retrieval noise (not shown) is better than the corresponding clear sky value because of the enhanced signal due to the presence of clouds.

In the previous case B, the effect of light path enhancement due to photon trapping, i.e. multiple reflection between the Earth surface the cloud is minor because of the low surface reflection. To estimate the algorithm performance in case of photon trapping, we show in Fig. 10 the Case C CO retrieval bias for optically thin water clouds at two different altitude levels with cloud top height $\tau_{\text{cld}} = 2$ and 5 km as a function of surface albedo and cloud fraction. In this case, measurements are only rejected for the cloud at 5 km altitude for increasing cloud fraction and low surface albedo. The cloud filtering is stricter for the optically thicker cloud ($\tau_{\text{cld}} = 5$), because the shielding of the atmosphere below the cloud becomes more efficient with increasing cloud optical thickness.

Photon trapping is most efficient for optically thin clouds above bright surfaces and can compensate the light path shortening due to the reflection of light at the cloud top. Overall, in case of convergence the CO bias is between ±2% and only for a few case (6 out of 336 measurement simulations) the retrieval did not converge. Thus, due to the retrieval of an scattering layer the effect of photon trapping can be account for by the retrieval.

Multiple scattering by atmospheric aerosol also affects the light path of the observed light. To investigate this effect, we consider the retrieval performance for case D. Figure 11 shows the CO retrieval bias as a function...
Figure 8: Retrieval bias $\Delta CO$ (upper panel) and retrieval noise (lower panel) for the clear sky test case A as a function of SZA and surface albedo $A_s$.

of surface albedo and aerosol optical depth. Here, the aerosol optical depth is indicated at 550 nm and 2300 nm. The ratio of both optical depths is mainly governed by the size of the aerosol particles. For both aerosol types with different layer heights, the retrieval bias does not exceed 1 %. From these results, we conclude that also scattering by aerosols can be treated appropriately by the retrieval code.

Next, we consider Case E measurement simulations for an atmosphere with a cirrus layer between 9 and 10 km and varying optical depth. Similar to Fig. 10, the CH$_4$ pre-fit rejects measurements for larger optical depth and low surface albedo. The effect of photon trapping at high surface albedo and small cirrus optical depth results in cirrus biases of less than 1 % (see Fig. 12). So, this example indicates that the presence of optically thin cirrus clouds can be handled by the retrieval algorithm.

As a last test ensemble, Case F, Fig. 13 shows the CO bias for multiple cloud layers, i.e. a cirrus between 9 and 10 km and a water cloud between 2 and 3 km. For the low albedo scenes most of the data are rejected by the cloud filter, where for high albedo most measurements pass the cloud filter. Here, multiple scattering
Figure 9: Retrieval bias for case B, i.e. for water clouds above a dark surface ($A_s = 0.05$) as a function of cloud top height $z_{cld}$ and cloud fraction $f_{cld}$ for different cloud total optical depths $\tau_{cld} = 5$ (upper left panel), $\tau_{cld} = 10$ (upper right panel), $\tau_{cld} = 30$ (lower left panel), $\tau_{cld} = 50$ (lower right panel).

between the surface, the water cloud and the cirrus takes place. The retrieval bias is in all cases below 1 %.

This demonstrates nicely the strength of the algorithm to deal with multiple scattering of sunlight even for multiple scattering layers.

In summary, the retrieval performance for the measurement ensembles of case A to F are compliant with the CO level-2 requirements as described in Sec. 4.2. The bias analysis showed that the effect of water clouds, which partially cover the observed ground scene, of optically thin cirrus, of atmospheric aerosol loading and of multiple cloud layers can be treated with sufficient accuracy by a triangular scattering layer implemented in a simplified two-stream radiative transfer model.
Figure 10: Retrieval bias in case of photon trapping between a water cloud and the bright surface (case C). The CO bias is shown as a function of surface albedo $A_s$ and cloud fraction $f_{cld}$ for a cloud with optical depth $\tau_{cld} = 2$ and cloud top height $z_{cld} = 2$ km (upper left), for $\tau_{cld} = 2$ and $z_{cld} = 5$ km (upper right), for $\tau_{cld} = 5$ and $z_{cld} = 2$ km (lower left) and for $\tau_{cld} = 5$ and $z_{cld} = 5$ km (lower right).
Figure 11: Retrieval bias for an aerosol loaded atmosphere of test case D. In the left panel, the CO retrieval bias is shown as a function of surface albedo and aerosol optical depth for a sulfate aerosol between the surface and 2 km altitude. The right panel shows the corresponding error analysis for an urban aerosol layer between 4 and 5 km altitude. At each panel, the lower x-axis describes the aerosol optical depth at 550 nm and the upper x-axis indicates the corresponding aerosol optical depth at 2300 nm.

Figure 12: CO retrieval bias for measurements in presence of optically thin cirrus clouds as a function of surface albedo and cirrus optical depth (case E). The cirrus optical depth is given at 2300 nm.
**Figure 13**: CO retrieval bias for measurements in presence of multiple cloud layers as a function of cirrus optical depth and cloud fraction of a water cloud for different surface albedo (case F). The cirrus optical depth is given at 2300 nm.
7.3 Geophysical test scenario for measurements over China

To test the algorithm performance for more realistic circumstances, we have simulated a measurements ensemble for China. Here, the measurement geometry and the atmospheric state of the observed scene are taken from several data sources. The viewing and solar geometry as well as the ground pixel size is adapted from a MODIS AQUA granule over China. The MODIS information is resampled on a pixel size of $10 \times 10$ km$^2$ at sub-satellite point, which is similar to the TROPOMI pixel size of $7 \times 7$ km$^2$. The pixel distortion towards the outer swath is adopted from the MODIS observation geometry. The surface albedo is taken from the MODIS MCD43C4 product for the Nadir Reflectance Band 7 at 2105–2155 nm and ground elevation is inferred from the ETOPO5 database of land and sea-floor elevations on a 5-minute latitude/longitude grid. Both surface albedo and ground elevation are interpolated to the MODIS granule grid using bilinear interpolation. The MODIS MYD06 data product with a horizontal resolution of $5 \times 5$ km$^2$ is used for cloud information, where we evaluate cloud top height from the MODIS cloud top pressure. Cirrus optical thicknesses are derived from the MODIS cirrus reflectance product and viewing geometry. Here, we use the algorithm by Desser et al. and Yang et al. [RD71, RD72]. The cirrus is located at 9–10 km all over the scene. One has to be aware that the cirrus data are rather uncertain due to various assumptions made in the cirrus algorithm, but also due to the MODIS data itself. Moreover, we added the MODIS aerosol product MOD04 at a spatial resolution of $10 \times 10$ km$^2$ at sub-satellite point to an overall rural background aerosol of an total optical depth of 0.1 at 2300 nm. ECMWF data on pressure, temperature, water vapor and air density stratification with a temporal resolution of 6 hours are interpolated to the instance of the MODIS measurement. Finally, the CHIMERE [RD73] and TM4 chemistry models [RD74] provide CO and CH$_4$ profiles, which are subsequently interpolated to the MODIS granule grid and to the ECMWF local pressure scales. Here, the CHIMERE chemical transport model provides information in the lower troposphere, and the TM4 fields are used to extend the CHIMERE fields to the upper troposphere and stratosphere. Examples of the ensemble fields are given in Fig. 14.

With this input and using the S-Lintran model as described above, measurements are simulated which subsequently are used as input to the SICOR retrieval algorithm. The corresponding SICOR CO bias $\Delta$CO is shown in Figure 15. About 28 % of the simulated observations are rejected by the CH$_4$ cloud filter causing the data gaps in the figure. For about 2 % of all converged cases, scattering had to be omitted in the inversion to avoid unphysical input to the radiative transfer model (see discussion in Sec. 5.2.5). Over the South China Sea, the retrieval converges in the presence of low water clouds. Other measurements in this region are rejected because of the methane filter or because of a too low signal. Furthermore, in the Northwest of China the CO biases exceed 5 % because of the occurrence of cirrus above a bright surface. Biases of $\leq$ 5 % occur around the city Wuhan due to the presence of tropospheric aerosols.

Figure 16 shows the bias distribution as a function of the methane filter $\Delta$CH$_4$. Here, the CO bias distribution is not centered around zero but has a mean bias of about 1.6 % and a median bias of about 1.4 %. This overall offset can be attributed to the triangular convolution of the molecular absorption cross sections as described in Sec. (5.4). The CO bias distribution has a narrow standard deviation of 2.3 %. It is asymmetric and skewed towards positive values. However, the figure shows no significant dependence of the bias on the CH$_4$ pre-fit bias with a correlation coefficient $r = 0.01$. This confirms the stable performance of the retrieval for different cloud conditions.

To demonstrate the gain in retrieval accuracy by accounting for atmospheric scattering, Fig. 17 shows the corresponding bias distribution for a non-scattering CO retrieval, where we assumed in both retrieval steps a clear sky, non-scattering atmosphere. The figure shows larger CO biases with a mean bias of -8.8 % and a bias median of -4.9 % with a strong correlation between the CO bias and the methane filter $\Delta$CH$_4$ with a correlation coefficient $r = 0.95$. The more clouds are present in the observed scene, the more methane and CO is underestimated by this approach. This correlation disappears completely when we retrieve the scattering parameters $\tau_{scat}$ and $\delta_{scat}$ in the second retrieval step.

Finally, Fig. 18 shows the distribution of the CO retrieval noise as a function of the LER in the continuum of the spectrum. As expected, the distribution shows a clear correlation of CO bias with the LER and the data are clustered around two maxima at LER-values of 0.1 and 0.4. The form of the CO bias distribution has its reason in the specific water cloud and cirrus distribution of our ensemble. The mean CO retrieval noise is 2.6 % and an insignificant amount of data points exceed the 6 % level. This low CO noise level may have its origin in the use of the MODIS 2.1 $\mu$m albedo to represent the surface reflection at the 2.3 $\mu$m band. The distribution of the LER values has its maximum at 0.1 and nearly no data points below 0.05. SCIAMACHY observations have indicated that the MODIS 2.1 $\mu$m albedo slightly overestimates the corresponding albedo at the 2.3 $\mu$m band by 0.03-0.06, which may cause a too low CO retrieval noise estimate for the China ensemble. However, the China test ensemble confirms the overall good retrieval performance which was already indicated by the generic test cases.
Figure 14: Test ensemble to generate TROPOMI measurements over China for a $10 \times 10$ km$^2$ pixel size for 10 May, 2006. The pixel distortion towards the edge of the swath is adopted from MODIS: (upper panel) MODIS cirrus optical depth, (middle) MODIS ground albedo at 2.1 $\mu$m, (lower) MODIS water cloud optical depth.
Figure 14: (Continued) (upper) MODIS cloud fraction, (middle) MODIS cloud top height, (lower) CHIMERE and TM4 CO total column.
Figure 14: (Continued) MODIS aerosol optical depth at 2300 nm.

Figure 15: Bias $Δ$CO of the retrieved CO column for the test ensemble shown in Fig. 14.
**Figure 16**: Bias of the retrieved CO column as a function of methane filter $\Delta CH_4$ for the China test ensemble shown in Fig. 14. The median of the methane pre-fit bias is −4.0 %, the mean bias is −6.6 % with a standard deviation of 6.7 %. Accordingly, the median of the CO retrieval bias is +1.4 %, the mean bias is +1.6 %, with a standard deviation of 2.3 %. The correlation coefficient ($r$) between $\Delta CO$ and $\Delta CH_4$ is +0.01.
Figure 17: Same as Fig. 16 but for a non-scattering CO retrieval. The median of the CO retrieval bias is $-4.90\%$, the mean bias is $-8.80\%$. The correlation coefficient ($r$) between $\Delta CO$ and $\Delta CH_4$ is 0.95.
Figure 18: CO column retrieval noise as a function of methane filter $\Delta$CH$_4$ for the China test ensemble shown in Fig. 14. The LER distribution as a median of 0.2, a mean value of 0.2 and a standard deviation of 0.1, whereas the CO retrieval noise distribution has a median of 2.5% a mean value of 2.6%, and a standard deviation of 0.8%. The correlation coefficient ($r$) between CO noise and LER is $-0.5$. 
Figure 19: Upper panel: CO retrieval bias as a function of a shift $\Delta T$ of the atmospheric temperature profile. Measurement ensembles of the clear sky generic test cases A (purple lines) and the cloudy sky test case B for a cloud with an optical depth of 10 (orange lines) are considered. Here, the maximum (dashed lines), mean (solid lines) and minimum (dotted lines) bias is reported for the different temperature shifts. Middle panel: mean spectral $\chi^2$ of the retrieval as a function of temperature shift. Lower panel: number of converged retrievals as a function of temperature shift.

7.4 Robustness of the CO retrieval with respect to uncertainties in the atmospheric input

In this section, we discuss the sensitivity of the CO retrieval with respect to uncertainties of atmospheric input data. In particular, we investigated the sensitivity with respect to

- A shift of the assumed temperature profile (see Fig. 19.)
- Errors in the ground pressure and therefore a scaling in the pressure profile (see Fig. 20.)
- Wrong a priori knowledge on the total amount of methane (see Fig. 21.)
- Uncertainties in the relative profile of $\text{CH}_4$ and $\text{H}_2\text{O}$ (see Figs. 22-24.)

To limit the numerical effort, we restrict the error analysis to the generic case A, describing clear sky conditions, and case B for a water cloud with optical depth $\tau_{\text{cld}} = 10$. For each of these data sets, we calculated the mean, the maximum and the minimum CO bias as a function of the corresponding sensitivity parameter.

Erroneous temperature input affects the CO retrieval due to wrong molecular absorption cross sections. Figure 19 shows the CO bias as a function of a constant shift of the atmospheric temperature profile. At $\Delta T = 0$ the maximum, minimum and mean bias characterises the bias distribution in Fig. 8 and in the upper right panel of Fig. 9. The small CO bias drift of about 0.1 % per K, shown in Figure 19, goes along with only marginal changes in the fit quality and the convergence. Only at temperature shifts $\Delta T > 3.5$ K we see a reduced convergence for test case A. Because of this low sensitivity, we expect no limitations of the overall algorithm performance due to uncertainties in the atmospheric temperature input.

Next, we consider the estimated surface pressure which must be representative for the observed ground scene. Errors in the surface pressure result in relative errors of the pressure profile. Similar to temperature uncertainty, this input error affects the retrieval due to wrong molecular absorption cross sections. Figure 20 shows the CO retrieval bias due to this error source for a scaling of the pressure profile between $\pm 3\%$. Within a scaling of $\pm 1\%$ the scaling affects only little the CO performance with a CO bias trend of 1.15 % for 1 % scaling
of the pressure profile. For a scaling < −1 %, the number of successful retrievals is significantly reduced for the clear sky scenarios indicating a critical loss of retrieval performance. However, a priori knowledge of the surface pressure within 1 % is reasonable and so we expect no critical performance issues due to uncertainties in the ground pressure.

Due to our retrieval concept, which uses the atmospheric CH₄ absorptions to retrieve effective scattering layer height and optical depth, the sensitivity of the retrieval with respect to uncertainties in our CH₄ a priori knowledge is of particular relevance. Figure 21 shows a linear one-to-one relationship between the CO bias and the CH₄ a priori uncertainty. So, to keep the CO bias due to this error source within ±3% the vertically integrated amount of CH₄ must be known within 3 % as well. This puts a stringent requirement on the methane a priori knowledge accuracy. In Sec. 7.6, we discuss the current accuracy of CH₄ model prediction.

Finally, we consider the CO retrieval sensitivity to errors in the assumed relative profile of relevant atmospheric absorbers. For CO, the difference between the relative CO profile and its truth causes a null-space contribution of the retrieval which is characterised by the column averaging kernel (see discussion in Sec. 5.2.2 and [RD55]). Hence, in the following we consider only the CO retrieval biases due to deficit knowledge of the relative profile of CH₄ and H₂O. For this purpose, an ensemble of relative profiles, shown in Fig. 22, is used. With increasing profile index I_{H₂O,CH₄} = 1···10 the underestimation of the relative abundance in the lower troposphere changes gradually to an overestimation, which comes along with a corresponding but opposite change in the middle troposphere. Both for CH₄ and H₂O, the true relative profile lies between profile I_{H₂O,CH₄} = 4 and 5. In general, the change in the relative H₂O profiles in the ensemble is larger than that of the CH₄ profiles, which reflects the stronger atmospheric variability of water. However, the assumed variability is based on ad hoc assumptions and should not be interpreted as an estimate of the a priori knowledge uncertainty, which is very hard to quantify. However, we believe that the considered range exceeds the actual uncertainty, yielding a conservative error estimate.

Figure 23 and 24 present the CO bias caused by the wrongly assumed vertical distribution of CH₄ and H₂O. For the clear sky scenarios of Case A, the dependence of the CO bias on the vertical CH₄ profile is small. Also the spectral χ² depends little on the profile. However, in case of a water cloud, where the retrieved properties of the scattering layer depend on methane absorption, the dependence on the vertical profile becomes significant with CO biases up to 5-10 %. The scattering layer height is adjusted by the retrieval such that the absorption by CH₄ above the cloud is described properly by the forward model. So, assuming a wrong relative profile directly relates to a wrong layer height and subsequently, it affects the retrieved CO abundance and its column averaging kernel. Here, critical CO biases are obtained for the profiles I_{CH₄} = 7,8,9,10 with enhanced CH₄ concentration at the lowest model layer of ≥ 20 %. The crossing point of

![Figure 20](image_url)
the individual relative CH₄ profiles $I_{CH_4} = 1 \ldots 10$ is at 2.5 km altitude, which is close to the cloud top height assumed in the measurement simulations. This makes the retrieval particular sensitive to the changes in the relative profile, because potential positive and negative deviations of the relative profile above the cloud do not cancel out. In that sense, we consider the vertical structures, shown in Fig. 22, as a worst case assumption. The retrieval sensitivity differs when we consider a perturbation of the relative H₂O profile. Because the water absorption is not used to determine scattering properties, the CO retrieval has little sensitivity to the vertical distribution of water vapor. The figure indicates that the bias on the retrieved CO column only varies between 1 to 2 % for the different profile $I_{H_2O} = 1 \ldots 10$. Bearing in mind that the ensemble of relative water profiles includes a large variation of ±70 % close to the surface, we consider this error source of minor relevance for the overall retrieval performance.
Figure 23: Same as Fig. 19 but assuming erroneous relative vertical CH$_4$ profile from Fig. 22 in the CO retrieval.

Figure 24: Same as Fig. 23 but for erroneous relative vertical H$_2$O profile from Fig. 22.
Robustness of the CO retrieval with respect to instrument artifacts

The accuracy of the CO retrieval product may also suffer from instrument artifacts. The acceptance level of instrument artifacts is described in detail by [RD32] [AD3]. This section is not meant to consolidate these studies but to exemplarily consider errors on five instrument-related parameters and their relevance for the CO retrieval quality. It is important to realize that the SWIR spectral range is also used to retrieve \( \text{CH}_4 \) vertical column amounts, and the required instrument performance in this spectral range is driven by both, the required CO and \( \text{CH}_4 \) retrieval uncertainty. Analogous to the previous section, the performance analysis is based on the generic ensemble of case A and case B with a water cloud of optical depth \( \tau_{\text{cld}} = 10 \). We consider measurement artifacts due to

- An error of the FWHM of the ISRF (see Fig. 25).
- Spectral calibration errors yielding a spectral squeeze (see Fig. 26).
- A spectrally constant radiometric offset (see Fig. 27).
- A spectrally constant scaling error (see Fig. 28).

To simulate the measurement spectra of the test ensembles, line-by-line spectra are convoluted with a Gaussian ISRF with an FWHM of 0.25 nm (see Eq. 3). Errors in the ISRF may introduce spectral residuals to the spectral fit and affect the CO retrieval performance. Figure 25 shows the CO retrieval sensitivity to errors of the FWHM of the ISRF between \( \pm 5 \% \), where the perturbed ISRF is applied to the Earthshine and solar spectrum. Here the retrieval bias changes only very little. However, when the FWHM is underestimated by more than 2 \% the spectral \( \chi^2 \) of the fit increases systematically and at the same time the number of successfully converged retrievals is reduced by a factor of 2.6. Thus, to warrant the CO retrieval quality, the FWHM of the ISRF has to be known at least with an uncertainty of 2 \% within the CO spectral fit windows. This is in agreement with the Mission and System Requirement Document (see [RD32] and [AD3]), which requests the FWHM knowledge of the ISRF with an uncertainty of 1 \%. 

\[ \text{Figure 25: CO retrieval bias as a function of a FWHM error (}\Delta \text{FWHM}\text{) of the ISRF. The measurement ensembles of the clear sky generic test cases A (purple lines) and the cloudy sky test case B for a cloud of optical depth 10 (orange lines) are considered. Here, the maximum (dashed lines), mean (solid lines) and minimum (dotted lines) bias is reported for the different FWHM errors. Middle panel: mean spectral } \chi^2 \text{ of the retrieval as a function of FWHM error. Lower panel: number of converged retrievals as a function of FWHM error.} \]
Next, we consider an erroneous spectral calibration of the measurement. Assuming a correct instrument calibration \( \lambda_i \), where spectral detector \( i \) is assigned to wavelength \( \lambda_i \), we consider an erroneous calibration

\[
\lambda'_i = \lambda_i + \delta s_1 + \frac{\lambda_i - \lambda_m}{\lambda_r - \lambda_m} \delta s_2.
\]

Here, \( \lambda_l = 2305 \text{ nm} \) and \( \lambda_r = 2385 \text{ nm} \) describe the shortwave and longwave edge of the SWIR band and \( \lambda_m = 2345 \text{ nm} \) is its spectral center. So, \( \delta s_1 \) and \( \delta s_2 \) characterise spectral calibration errors which result in an overall spectral shift and a spectral squeeze. Both are defined over the full spectral range and the corresponding squeeze for the CO fit windows (2315-2338 nm) is about one-third of \( \delta s_2 \).

Because a spectral shift is a default fit parameter of the CO inversion, the retrieval is not sensitive to the calibration error \( \delta s_1 \). This was confirmed by retrieval simulations for the range \(-0.1 \text{ nm} \leq \delta s_1 \leq +0.1 \text{ nm} \), which is the spectral range of plus-minus one spatial sampling distance (SSD) of the SWIR spectrometer (not shown). The situation differs when a calibration error \( \delta s_2 \) is introduced to the measurement simulation. Figure 26 shows a nearly linear dependence of the CO bias in range \(-0.1 \times \text{ SSD} \leq \delta s_2 \leq 0.1 \times \text{ SSD} \) with a relative CO bias dependence of 1 % per 0.1 SSD. The required knowledge of the spectral position of the spectral channels is 2 pm which corresponds to 1/50 SSD [AD3]. Thus for an instrument satisfying this requirement, a calibration error conform the spectral squeeze in Eq. (73) provides a minor contribution to the CO error budget.

Finally, we consider the effect of two radiometric errors: a spectrally constant radiometric offset which e.g. can be caused by spectral stray light and detector dark currents, and a spectrally constant radiometric scaling due to an overall radiometric calibration error of the instrument gain. The CO bias has only a weak sensitivity to a radiometric offset, as shown in Fig. 27. However for an offset > 0.6 %, the number of successful convergence is reduced. At the same time no significant effect is seen for the converged cases in spectral residuals. This means that a radiometric offset mainly effects the stability of the inversion and so to assure a stable inversion the offset must not exceed 0.6 % of the spectral continuum. For the radiometric gain, Fig. 28 indicates no sensitivity to a scaling error within the range of \( \pm 5 \% \) of the spectral continuum value. The required SWIR radiometric accuracy of the instrument is 0.1 % for a radiometric offset and 3% for the overall radiometric accuracy, which both is sufficient to keep the CO accuracy within its required error budget.

The reported error sensitivity of the CO retrieval with respect to instrument artifacts is relatively low, and the required instrument performance is mainly driven by the CH\(_4\) column retrieval [RD2]. For CH\(_4\), spectrally strong absorption features are needed to determined atmospheric scattering characteristics simultaneously with the abundance of water vapor and CH\(_4\) [RD2]. Here, the radiometric accuracy in the deep absorption lines drives
Figure 27: Same as Fig. 25 but as a function of a spectrally constant radiometric error $\Delta I_{\text{offset}}$. Here, $\Delta I_{\text{offset}}$ is defined with respect to the continuum value of the spectrum.

Figure 28: Same as Fig. 25 but as a function of a spectrally constant radiometric scaling error $\Delta I_{\text{scale}}$. Here, $\Delta I_{\text{scale}}$ is defined with respect to the continuum value of the spectrum.

mainly the radiometric requirements of the instrument. For the retrieval of CO, we have selected a spectral fit window with only moderate and weak absorption of CH$_4$ and H$_2$O to reduce interference with the weak absorption lines of CO. Due to this, the sensitivity to radiometric errors is weaker for the CO retrieval. However to retrieve atmospheric scattering properties, we employ a priori knowledge about the CH$_4$ abundance and use the corresponding spectral features to obtain a measurement sensitivity to the height and optical depth.
of a scattering layer. This puts a stringent requirement on our a priori knowledge of methane and so it shifts challenges from instrument calibration to atmospheric modeling. In the next section we discuss the accuracy of state-of-the-art chemical transport modeling to predict the CH$_4$ concentration.
7.6 Quality of the model derived XCH$_4$

An important element of the CO retrieval approach is the use of an atmospheric transport model to estimate the methane abundance in the observed scene. Figure 21 shows the error on the retrieved CO column assuming an error of $\pm 3\%$ on the methane a priori knowledge. In both cases, the CO retrieval bias is clearly affected by this error. Overall, a $\pm 3\%$ CH$_4$ uncertainty causes an additional error on the retrieved CO column of about 3%. Additionally, it affects the number of successful retrievals. So, it is essential to make a sophisticated estimate of the methane uncertainties. For this purpose, we consider the dry air mole fraction XCH$_4$$_{mod}$ predicted by the global transport model TM5. Moreover, we propose different options for operational delivery of CH$_4$$_{mod}$ in the context of the TROPOMI mission including a comparison of the expected performance.

7.6.1 Simulation of XCH$_4$

The TM5 model [RD34] has been used in several studies of CH$_4$ (e.g. [RD75, RD76, RD77]). This specific version of the model has been developed to estimate the global sources and sinks of CH$_4$ from atmospheric measurements using a variational inverse modeling technique. In short, the method starts with a first guess, or a priori estimate, on the basis of published emission inventories and output from land surface models. A forward run of the TM5 model yields global maps of the spatio-temporal distribution of CH$_4$ in the atmosphere, which are sampled according to the times and locations of available measurements. The inverse modeling technique optimizes the sources and sinks by minimising a least squares cost function, consisting of the residual differences between model and measurements and deviations from the first guess. This procedure has been used to optimize TM5 on the basis of measurements from the NOAA-ESRL global monitoring network and of CH$_4$ retrievals from SCIAMACHY and GOSAT.
7.6.2 Uncertainty of model-derived XCH$_4$

The overall uncertainty of model derived XCH$_4$ can be separated into two components: 1. Limitations in the state-of-the-art knowledge of XCH$_4$. 2. Limitations of the applied modeling procedure to reproduce the state-of-the-art. The second contribution refers to limiting conditions, mostly due to requirements on the timeliness of data delivery, which prompt for a non optimal approach.

Contribution 1 is assessed by comparing inverse modeling optimised XCH$_4$ to high precision measurements from the Total Carbon Column Observing Network (TCCON) of ground based Fourier Transform Spectrometers (FTS). Figure 29 shows examples of such comparisons for a selection of sites spanning a wide range in latitude. In these comparisons, the model has been optimised using measurements from 46 background sites of the NOAA-ESRL network.

The root mean square difference in diurnal mean XCH$_4$ between the TCCON measurements and the NOAA optimised model ranges from 8 to 22 ppb between sites, with a mean of 13 ppb for the 10 sites included in the analysis. As can be seen, the RMS differences are explained in part by a small bias of the model, which increases towards southern latitudes. Without the contribution of this systematic offset, which could in principle be corrected for, the mean RMS difference reduces to 8 ppb. However, it should be realised that the TCCON network is not designed to be well representative for typical conditions encountered over land worldwide. Therefore the optimised model may on average be further from the truth than indicated by the TCCON residuals. Limited means are available for quantifying this difference. Figure 30 shows results from an attempt in this direction using retrievals from the Greenhouse Gas Observing Satellite (GOSAT). The results are largely consistent with the TCCON analysis, showing a systematic underestimation of TM5 simulated XCH$_4$ increasing towards southern latitudes. The RMS difference between TM5 and GOSAT amounts to 20 ppb. Part of this difference is explained by the uncertainty of GOSAT XCH$_4$ retrievals, which has been estimated at 15 ppb [RD78].

From Fig. 30, it can be concluded that the use of GOSAT in the optimisation of TM5 reduces the uncertainty of model simulated XCH$_4$. Besides measurement continuity, also timeliness needs further consideration. Commonly, inverse modeling derived estimates lag behind real-time by approximately a year. This has to do with the availability of various types of inputs that are required, including meteorological driving fields, a priori emission estimates, and measurements. The most important limitation at present is the release of new surface measurements, which cannot easily be accelerated because of standard calibration and verification procedures adopted by the laboratories involved. Alternatively, other modeling procedures can be followed to achieve near real time model estimates of XCH$_4$, such as:

![Figure 30: Comparison between TM5 simulated and GOSAT retrieved XCH$_4$. Differences (TM5 minus GOSAT) are shown for the period June 2009–June 2011. The TM5 results have been optimised using surface measurements.](image-url)
Figure 31: The standard deviation of total XCH$_4$ between years expressed in % of mean XCH$_4$. Standard deviations are calculated from TM5 XCH$_4$ fields, optimised using surface measurements, for the 15$^{th}$ day of the month in each year in the period 2003–2010.

1. Use of inversion-optimised TM5 estimates of a different year.

2. Option 1 extrapolated to the current year using a priori emissions.

3. Option 1 extrapolated to the current year using a posteriori emissions of the previous year.

It is clear that each of these options is less accurate than the NOAA optimised model discussed so far, which brings us to the estimation of uncertainty component 2 mentioned at the start of this section. For procedure 1, the uncertainty component 2 is the variability in total column XCH$_4$ caused by year-to-year variations in meteorology and the inter-annual variation of methane sources and sinks. The size of this error is estimated from results of a multi-year inversion, by calculating how XCH$_4$ on a given day of the year varied between the years. Figure 31 shows results of this procedure applied to selected days over the year.

The results highlight regions with a high variability in atmospheric transport, such as the storm tracks in the northern and southern hemisphere. The largest variations are found over South East Asia, due to large regional sources of methane. Some enhanced variability can be seen near the equator, where the north-south gradient in XCH$_4$ is largest. Standard deviations are on average well within 1% (18 ppb), regionally increasing up to 1.5% (27 ppb). Sporadically, standard deviations up to 3% are found, associated with biomass burning events.

Procedure 2 largely eliminates the meteorological error component, which is quite prominent in procedure 1, by using actual meteorological driving fields (see Fig. 32). As can be seen in this case, however, systematic differences show up that are caused by inaccuracies on the latitudinal distribution of the a priori fluxes, which are no longer corrected by the inversion. These differences increase towards high northern latitudes and over time, and amount up to 25 ppb after a year since initialisation (using optimised methane fields).

The impact of systematic differences between a priori and a posteriori fluxes affecting procedure 1 can be reduced by extrapolating inverse modeling results using a posteriori instead of a priori fluxes. Obviously, the inconsistency that remains in this case (procedure 3) is that the fluxes are optimised for a different year. As shown in Fig. 33, this procedure avoids growing offsets in the latitudinal gradient over time. Instead, the largest variations show up in the tropics, in response, for example, to inter-annual variability of tropical sources such as wetlands and biomass burning. It should be realised that tropical sources of methane are relatively poorly constrained, both a priori and by the limited availability of surface measurements in the tropics. Compared
Figure 32: The RMS difference between NOAA optimised CH$_4$ and results of procedure 2 evaluated after 6 months. RMS values are calculated from the differences between the two simulations for all days of the 6th month after initialisation of procedure 2.

Figure 33: As Fig. 32 for procedure 3, evaluated in different months.

with Fig. 31, the deviations may seem larger than expected since one source of error (the transport variability error) has been eliminated. This is explained by the fact that emissions are optimised monthly and therefore the anomalies persist during the month. Therefore the residuals contributing to the RMS have an important systematic component, whereas the standard deviations in Fig. 31 represent random variations. The implication is that $> 1\sigma$ perturbations show up in Fig. 33, whereas Fig. 31 only shows $1\sigma$ intervals. Figure 33 indicates that for procedure 3, the error compared to the optimised model is larger by up to 1.5%.

7.6.3 Discussion and conclusions

As demonstrated in the previous section, the accuracy of model-derived XCH$_4$ estimates depend on the optimisation approach. Options for model optimisation, in turn, depend on timeliness requirements on the delivery of data products. Table 4 lists uncertainties for each of the proposed options, calculated from the available statistics for three regions: 1. Global land, 2. South East Asia, 3. Tropical America. The most accurate approach to estimating XCH$_4$,mod is to optimize the model using both surface and satellite measurements of
Table 4: Estimated uncertainty in XCH₄ comparing the proposed modeling approaches.

<table>
<thead>
<tr>
<th>Method</th>
<th>Land 2σ (ppb (%))</th>
<th>S.E. Asia 2σ (ppb (%))</th>
<th>Trop. America 2σ (ppb (%))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimised TM5</td>
<td>26 (1.5)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Procedure 1</td>
<td>32 (1.9)</td>
<td>38 (2.2)</td>
<td>35 (2.0)</td>
</tr>
<tr>
<td>Procedure 2</td>
<td>30 (1.7)</td>
<td>32 (1.9)</td>
<td>30 (1.7)</td>
</tr>
<tr>
<td>Procedure 3</td>
<td>28 (1.6)</td>
<td>30 (1.7)</td>
<td>28 (1.6)</td>
</tr>
</tbody>
</table>

CH₄, which limits the uncertainty of XCH₄,mod to within 1% (18 ppb). When using only surface measurements, the uncertainty is estimated at 26 ppb (at 95% confidence), based on comparisons with TCCON data. This should be considered a lower limit, because of the limited representativity of the TCCON network. On the other hand, part of the error is systematic and can to first order be corrected on the basis of available TCCON and GOSAT measurements. The overall uncertainty is difficult to quantify but should remain within 2% (35 ppb).

Without the possibility to resort to optimised model estimates, additional errors are introduced, depending on the adopted procedure. The least demanding approach, with respect to timeliness requirements, is to use results from the most recent optimised year. In this case, the additional error is on average within 20 ppb (at 95% confidence), increasing to 28 ppb (1.6%) over intense source regions of South East Asia. Combining these uncertainties with the uncertainty of the optimisation approach (taken as 26 ppb), yields an overall uncertainty of 32–38 ppb (1.9–2.2%) assuming no interdependence between the errors. The transport model uncertainty can be further reduced by using actual meteorology, which limits the timeliness to about half a year behind real-time. This approach works best using optimised instead of a priori fluxes for the years that were analysed, which yields estimated uncertainties in the range of 28–30 ppb (1.6–1.7%). Our assessment of the 2nd uncertainty component for procedure 1–3 may be somewhat optimistic, because of the limited resolution of the transport model and potential underestimation of inter-annual variability of CH₄ emissions using surface data. Acknowledging these limitations in our approach, an uncertainty of 2% seems a reasonable save margin that should be achievable for most conditions encountered throughout the global domain.
8 Algorithm input and output

8.1 High level processing scheme

The operational CO and CH₄ algorithms use a similar processing scheme, which for CO is summarised in Fig. 34. Red boxes indicate modules that are described in this ATBD, light green boxes are modules that are developed in the context of the overall processing framework (see e.g. [AD4]). The algorithm input is divided in static input and dynamic input. Static input is read once, during algorithm initialisation, whereas dynamic input is different for each pixel.

Figure 34: High level processing scheme for operational S5P CO data reduction. Modules that are described in this ATBD, are indicated by the red boxes.

8.2 Static input

Table 5 shows all input information that remains the same during one level-2 processor run. Those consist of data that remains constant, such as cross sections, and data that is updated not more frequently than a restart of the processor. The latter kind of data is referred to as semi-static data.

All data should be accompanied with the domain on which they are defined. For example, the cross section lookup tables have a grid of temperatures, pressures and wavenumbers, the ISRFs have a domain of wavelength relative to the centre of the ISRFs and irradiance spectra are defined on wavelength grids.

8.3 Dynamic input

The dynamic input is listed in table 6. These data are specific for each measurement pixel. Dynamic data coming from different sources should be co-located to the horizontal position and observation time of the Level 1b product. Nearest neighbor approximation is also possible. Vertical domains and sampling may differ for the different input fields. The original ECMWF and TM5 data should be given on their original vertical hybrid grid.
Table 5: SICOR Static input. Calibration key data and irradiance L1b-product are semi-static, because they are provided once per processor run.

<table>
<thead>
<tr>
<th>Name/Data</th>
<th>Symbol</th>
<th>Unit</th>
<th>Source</th>
<th>Pre-process needs</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculation settings</td>
<td></td>
<td></td>
<td>SRON</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Absorption cross sections(^1)</td>
<td></td>
<td>(\sigma) (\text{cm}^2)</td>
<td>[RD58, RD59]</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Reference irradiance spectrum</td>
<td>(I_{\text{ref}})</td>
<td>(\text{mol s}^{-1} \text{m}^{-2} \text{nm}^{-1} \text{sr}^{-1})</td>
<td>KNMI</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ISRF</td>
<td>(S_R)</td>
<td>(\text{nm}^{-1})</td>
<td>Calibration key data</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Solar irradiance SWIR</td>
<td>(I)</td>
<td>(\text{mol s}^{-1} \text{m}^{-2} \text{nm}^{-1})</td>
<td>SSP L1b product</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Noise solar irradiance SWIR</td>
<td>(\epsilon_I)</td>
<td>(\text{mol s}^{-1} \text{m}^{-2} \text{nm}^{-1})</td>
<td>SSP L1b product</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Irradiance pixel mask</td>
<td>(l_I)</td>
<td>–</td>
<td>SSP L1b product</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

\(\text{1}\) A lookup table for CO, CH\(_4\), H\(_2\)O and HDO on a grid of temperatures, pressures and wavenumbers, covering the SWIR band.

Like for static input data, all dynamical input fields must contain information on the corresponding domains. For the radiance measurement, the domain is a wavelength grid, for vertical profiles, the domain is defined by hybrid a- and b-coefficients. The CO algorithm needs the a- and b-coefficients on the level interfaces. This applies for both the ECMWF data and the TM5 data.

Table 6: SICOR Dynamic input.

<table>
<thead>
<tr>
<th>Name/Data</th>
<th>Symbol</th>
<th>Unit</th>
<th>Source</th>
<th>Pre-process needs</th>
<th>Backup if not available</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earth radiance SWIR</td>
<td>(R)</td>
<td>(\text{mol s}^{-1} \text{m}^{-2} \text{nm}^{-1} \text{sr}^{-1})</td>
<td>SSP L1b product</td>
<td>–</td>
<td>No retrieval</td>
<td></td>
</tr>
<tr>
<td>Noise Earth radiance SWIR</td>
<td>(\epsilon_R)</td>
<td>(\text{mol s}^{-1} \text{m}^{-2} \text{nm}^{-1} \text{sr}^{-1})</td>
<td>SSP L1b product</td>
<td>–</td>
<td>No retrieval</td>
<td></td>
</tr>
<tr>
<td>Radiance pixel mask</td>
<td>(l_R)</td>
<td>–</td>
<td>SSP L1b product</td>
<td>–</td>
<td>No retrieval</td>
<td></td>
</tr>
<tr>
<td>Solar zenith angle</td>
<td>(\theta_0)</td>
<td>degree</td>
<td>SSP L1b product</td>
<td>–</td>
<td>No retrieval</td>
<td></td>
</tr>
<tr>
<td>Viewing zenith angle</td>
<td>(\theta_v)</td>
<td>degree</td>
<td>SSP L1b product</td>
<td>–</td>
<td>No retrieval</td>
<td></td>
</tr>
<tr>
<td>Relative azimuth angle(^1)</td>
<td>(\varphi)</td>
<td>degree</td>
<td>SSP L1b product</td>
<td>–</td>
<td>No retrieval</td>
<td></td>
</tr>
<tr>
<td>Longitude</td>
<td>(\text{lat})</td>
<td>degree</td>
<td>SSP L1b product</td>
<td>–</td>
<td>No retrieval</td>
<td></td>
</tr>
<tr>
<td>Latitude</td>
<td>(\text{lon})</td>
<td>degree</td>
<td>SSP L1b product</td>
<td>–</td>
<td>No retrieval</td>
<td></td>
</tr>
<tr>
<td>SWIR mean pixel elevation</td>
<td>(z_s)</td>
<td>m</td>
<td>GMTED 2010</td>
<td>Interpolate</td>
<td>Same as (z_{s,0})</td>
<td></td>
</tr>
<tr>
<td>Temperature profile</td>
<td>(T)</td>
<td>K</td>
<td>ECMWF</td>
<td>Interpolate</td>
<td>Previous day</td>
<td></td>
</tr>
<tr>
<td>Specific humidity profile</td>
<td>(q)</td>
<td>(\text{kg kg}^{-1})</td>
<td>ECMWF</td>
<td>Interpolate</td>
<td>Previous day</td>
<td></td>
</tr>
<tr>
<td>Surface pressure</td>
<td>(p_0)</td>
<td>Pa</td>
<td>ECMWF</td>
<td>Interpolate</td>
<td>Previous day</td>
<td></td>
</tr>
<tr>
<td>ECMWF surface elevation</td>
<td>(z_{s,0})</td>
<td>m</td>
<td>ECMWF</td>
<td>Interpolate</td>
<td>Previous day</td>
<td></td>
</tr>
<tr>
<td>CO mixing ratio profile</td>
<td>(m_{\text{CO}})</td>
<td>1</td>
<td>TM5</td>
<td>Interpolate</td>
<td>Previous day</td>
<td></td>
</tr>
<tr>
<td>CH(_4) mixing ratio profile</td>
<td>(m_{\text{CH}_4})</td>
<td>1</td>
<td>TM5</td>
<td>Interpolate</td>
<td>Previous day</td>
<td></td>
</tr>
</tbody>
</table>

\(\text{1}\) Zero degrees means that the instrument is looking towards the sun.

8.4 Algorithm output

For each spatial sounding, the level-2 output is generated and the corresponding SICOR output data structure is listed in table 7. For the operational data processing, the SICOR CO product is processed as part of both the near-real-time processor and the offline processor. Here, a subset of output fields will not be calculated in the near-real-time processor and will be filled with NetCDF fill values. The total output product is 476 bytes per pixel. More details about the SICOR CO data product is given in the CO product user manual [RD79].
Table 7: SICOR output fields. \( N_z \) is the number of layers in the model atmosphere, and is set to 50 by default.

<table>
<thead>
<tr>
<th>Name/Data</th>
<th>Symbol</th>
<th>Unit</th>
<th>Description</th>
<th>Data type</th>
<th>Number of values per pixel</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO column</td>
<td>vco</td>
<td>mol m(^{-2})</td>
<td>Float</td>
<td>1</td>
<td>Main product</td>
<td></td>
</tr>
<tr>
<td>Noise CO column</td>
<td>eco</td>
<td>mol m(^{-2})</td>
<td>Float</td>
<td>1</td>
<td>Main product</td>
<td></td>
</tr>
<tr>
<td>CO column averaging kernel</td>
<td>avg</td>
<td>1</td>
<td>Float</td>
<td>( N_z )</td>
<td>Main product</td>
<td></td>
</tr>
<tr>
<td>Layer interface pressure</td>
<td>press</td>
<td>Pa</td>
<td>Float</td>
<td>1</td>
<td>Auxiliary product</td>
<td></td>
</tr>
<tr>
<td>( \text{H}_2\text{O} ) column</td>
<td>vh2o</td>
<td>mol m(^{-2})</td>
<td>Float</td>
<td>1</td>
<td>Auxiliary product</td>
<td></td>
</tr>
<tr>
<td>Noise ( \text{H}_2\text{O} ) column</td>
<td>eh2o</td>
<td>mol m(^{-2})</td>
<td>Float</td>
<td>1</td>
<td>Auxiliary product</td>
<td></td>
</tr>
<tr>
<td>HDO column</td>
<td>vhdo</td>
<td>mol m(^{-2})</td>
<td>Float</td>
<td>1</td>
<td>Auxiliary product</td>
<td></td>
</tr>
<tr>
<td>Noise HDO column</td>
<td>ehdo</td>
<td>mol m(^{-2})</td>
<td>Float</td>
<td>1</td>
<td>Auxiliary product</td>
<td></td>
</tr>
<tr>
<td>Prefit ( \text{CH}_4 ) column</td>
<td>vch4(_p)</td>
<td>mol m(^{-2})</td>
<td>Float</td>
<td>1</td>
<td>Cloud filter product</td>
<td></td>
</tr>
<tr>
<td>Weak-band ( \text{H}_2\text{O} ) column</td>
<td>vh2o(_w)</td>
<td>mol m(^{-2})</td>
<td>Float</td>
<td>1</td>
<td>Cloud filter product</td>
<td></td>
</tr>
<tr>
<td>Strong-band ( \text{H}_2\text{O} ) column</td>
<td>vh2o(_s)</td>
<td>mol m(^{-2})</td>
<td>Float</td>
<td>1</td>
<td>Cloud filter product</td>
<td></td>
</tr>
<tr>
<td>Weak-band ( \text{CH}_4 ) column</td>
<td>vch4(_w)</td>
<td>mol m(^{-2})</td>
<td>Float</td>
<td>1</td>
<td>Cloud filter product</td>
<td></td>
</tr>
<tr>
<td>Strong-band ( \text{CH}_4 ) column</td>
<td>vch4(_s)</td>
<td>mol m(^{-2})</td>
<td>Float</td>
<td>1</td>
<td>Cloud filter product</td>
<td></td>
</tr>
<tr>
<td>Scatterer optical depth 2330 nm</td>
<td>tausat</td>
<td>1</td>
<td>Float</td>
<td>1</td>
<td>Auxiliary product</td>
<td></td>
</tr>
<tr>
<td>Scatterer centre height</td>
<td>zscat</td>
<td>m</td>
<td>Float</td>
<td>1</td>
<td>Auxiliary product</td>
<td></td>
</tr>
<tr>
<td>Albedo 2325 nm</td>
<td>alb(_1)</td>
<td>1</td>
<td>Float</td>
<td>1</td>
<td>Auxiliary product</td>
<td></td>
</tr>
<tr>
<td>Albedo 2335 nm</td>
<td>alb(_2)</td>
<td>1</td>
<td>Float</td>
<td>1</td>
<td>Auxiliary product</td>
<td></td>
</tr>
<tr>
<td>Spectral shift</td>
<td>spec(_{off})</td>
<td>nm</td>
<td>Float</td>
<td>1</td>
<td>Auxiliary product</td>
<td></td>
</tr>
<tr>
<td>Residual ( \chi^2 )</td>
<td>chi2</td>
<td>1</td>
<td>Float</td>
<td>1</td>
<td>Quality product</td>
<td></td>
</tr>
<tr>
<td>Number of spectral bins ( N_y )</td>
<td>ny</td>
<td>1</td>
<td>Integer</td>
<td>1</td>
<td>Auxiliary product</td>
<td></td>
</tr>
<tr>
<td>Degree of freedom for signal DFS</td>
<td>dfs</td>
<td>1</td>
<td>Float</td>
<td>1</td>
<td>Quality product</td>
<td></td>
</tr>
<tr>
<td>Iterations</td>
<td>niter</td>
<td>1</td>
<td>Integer</td>
<td>1</td>
<td>Quality product</td>
<td></td>
</tr>
<tr>
<td>Processing quality flag</td>
<td>PQF</td>
<td>–</td>
<td>Integer</td>
<td>1</td>
<td>Quality product</td>
<td></td>
</tr>
</tbody>
</table>
9 Spatial data selection approach

For CO, we aim to process data over land and ocean, where over ocean only cloudy scenes are considered. Here, measurements are selected with a Lambert-equivalent reflectivity (LER) > 0.03 at the continuum of the spectrum (2315 nm). The LER is provided as a standard output of the SWIR pre-processor (see Appendix B). For land pixels without inland waters, the LER value rarely goes below this threshold as indicated by the global distribution of the surface albedo, shown in Fig. 35. However, the LER value may fall below the threshold value for land pixels which are covered partly by inland waters or clear sky ocean pixels. In the SWIR spectral range, water surfaces reflect too little light to obtain useful measurements for these scenes. This LER filter reduces the number of calculations which depends mainly on the cloud coverage of ocean pixels.

Figure 35: Average surface albedo over five years of SCIAMACHY cloud-free land observations at 2300 nm at a resolution of 0.5° (2003–2007).
Figure 36: CO retrievals of a TROPOMI orbit granule on 27 June 2018 over the UK. Panels of the first row depict the original data, the second row shows the destriped TROPOMI CO data (FMD method left, FFD method right), and the third row illustrates the destriping mask that was subtracted from the original TROPOMI data.

10 A posteriori destriping of the level-2 data

The TROPOMI CO retrievals from single orbits show a striping pattern along the flight path, which is a well-known feature for observations of push-broom spectrometers (e.g., OMI [RD80],[RD61] reported that the CO stripes can exceed 5 ppb and can hamper, e.g., the detection of small point sources and the estimate of emissions from fire plumes. The origin of the stripy pattern is not yet understood and is changing with time from orbit to orbit. [RD61] suggested an empirical destriping approach that is applied on the CO data fields (see left column of Fig 36). This method removes first the background of the CO field by a median smoothing in cross-track direction and then determines per orbit a fixed stripe pattern for correction by a median along the flight path. This method already reduces a major part of the stripes in the CO data and is denoted as fixed mask destriping (FMD). Analyzing TROPOMI CO orbit observation, we found that the stripe patterns changes to some extent also along the flight path, which cannot be captured by this approach. Therefore, we developed an alternative approach that is based on a Fourier filter destriping (FFD) [RD60] (see right column of Fig 36) and already deployed it for scientific studies e.g. [RD81, RD82]. Due to its robustness the FMD destriping approach is implemented for ESA’s operational processing of TROPOMI data in the first instance, the FFD approach is planed for a future update of the processor.
11 Conclusion

We presented the theoretical baseline of the CO column retrieval algorithm SICOR, which is developed for the operational near real-time (NRT) and offline (OFL) data processing of the Sentinel 5 Precursor mission. To improve our current knowledge on CO using satellite observations, the total column has to be retrieved with maximum accuracy 15 % and a maximum precision of 10 %, even for background CO atmospheric abundance and low surface reflection. The retrieval approach is based on the so-called physics-based retrieval concept, which aims to retrieve jointly atmospheric trace gas abundances and scattering properties of the observed atmosphere. The retrieval approach relies on a loose cloud filtering. In conjunction with the large data rate of the Sentinel 5 Precursor mission, this means a serious burden on the numerical efficiency of the CO algorithm.

The SICOR algorithm relies on a two-step retrieval from TROPOMI SWIR measurements. In the first step as part of the SWIR pre-processing, a non-scattering retrieval of the total amount of CH$_4$ in the spectral range 2315-2324 nm is performed to filter on the presence of high and optically thick clouds. In a second step, the CO column is inferred from SWIR measurements of an adjacent spectral window 2324-2338 nm. Here, a priori knowledge on the atmospheric methane abundance is used to retrieve effective cloud parameters from methane absorption bands simultaneously with atmospheric CO and H$_2$O abundances. The algorithm uses two highly efficient numerical modules. The inversion infers the trace gas column information by a least squares fitting of a scaling of a CO reference profile. From a theoretical point of view, the approach is identical to a regularized CO profile retrieval using Tikhonov regularization of first order and infinite regularization strength. This analogy provides us an analytical approach to calculate the column averaging kernel, which describes the sensitivity of the retrieved column to changes in the corresponding trace gas profile as a function of altitude. Moreover, the retrieval employs a two-stream radiative transfer model, which is linearized with respect to the parameters to be retrieved. The two-stream approximation is a simple approximation to account for multiple light scattering in atmospheric radiative transfer and its numerical implementation comes along with low computational cost.

For a HP dc7900 SFF hardware with Intel® Core™2 Duo 1390 CPU E8400 at 3.00 GHz and 4 GB RAM, the computational burden of this retrieval approach is estimated to be 0.17 s using the the Intel FORTRAN compiler and 0.24 s using the GNU gfortran compiler. Assuming the intel compiler and a data stream of 260 spectra/s and a total usable observation time of 50 min. per orbit, 22 cores of the described hardware are needed to keep up the TROPOMI data stream.

An extensive performance analysis demonstrated a robust algorithm with stable numerical implementation, where we considered both erroneous atmospheric input fields and instrument malfunctions. The retrieval performance was analyzed for 6 generic measurement ensembles, where solar geometry, surface albedo and aerosol, cirrus and cloud parameters were varied systematically over a realistic parameter domain. Furthermore, we investigated the performance for an ensemble of about 26 000 simulated measurements over China, combining different data sources that describe the observed scenes as realistically as possible. For this purpose, we used MODIS cloud, aerosol and surface information, ECMWF pressure, water and temperature and CO and CH$_4$ profiles, simulated with the chemical transport models TM4 and CHIMERE. Measurements are simulated with the S-LINTRAN version 2.0 radiative transfer model, which accounts for multiple scattering by water clouds and cirrus. Additionally, we considered the robustness of the algorithm with respect to serval potential malfunction of the instrument, i.e. erroneous instrument spectral response, spectral calibration, a constant radiometric offset and a multiplicative radiometric error. Overall, the analysis showed a very stable numerical performance of the SICOR CO retrieval and the accuracy and precision of the CO data product is fully compliant with the level 2 user requirements. In this context, we identified the a priori knowledge of the atmospheric methane abundance as the most critical input parameter and to keep the corresponding CO retrieval bias within 3 %, the vertically integrated column density of CH$_4$ must be known with an accuracy of also 3 %. We showed that the forecast skills of present state-of-the-art chemical transport models are sufficient to predict methane with the required accuracy.

We characterized the input and output interfaces of the algorithm with the overall processing frame. Besides standard measurement information, the algorithm requires the following information on the atmospheric state:

1. Relative profile of CO and H$_2$O.
2. Temperature-pressure profile.
3. Surface pressure on SWIR pixel level.
5. Geo-potential height of the different input profiles.
6. First guess of the total amount of CO and H$_2$O.

As a compact retrieval product, which reflects both the need of the data user and the information content of the SWIR measurements, the vertically integrated CO column density is provided together with its retrieval noise and the column averaging kernel for each individual measurement.

Based on the analysis of the CO data produced since the S5p satellite launch, we conclude that the mission ground segment provides a reliable high-quality CO data product, detecting the distribution of CO on global, regional and city scales with daily global coverage. Routine validations with ground based TCCON observations indicates a CO data quality compliant with the mission requirements (https://mpc-vdaf.tropomi.eu/index.php/carbon-monoxide). Due to the high radiometric performance combined with the daily global coverage, the TROPOMI SWIR spectrometer opens doors for new research of natural and man-made emission of CO and air quality monitoring, one of the missions main objectives.
Acknowledgements

We would like to thank Michael Buchwitz and Thomas Krings (IUP, University Bremen, Germany) for their constructive comments that aided us to improve this document.
### A Appendix: Flux method PIFM

For non-conservative scattering in model layer $n$, most two-stream methods rely on a system of flux differential equations of the form

\[
\begin{align*}
\frac{dF^\uparrow}{d\tau} &= \alpha_{1,n} F^\uparrow - \alpha_{2,n} F^\downarrow - \alpha_{3,n} \frac{S}{\mu_o} \\
\frac{dF^\downarrow}{d\tau} &= \alpha_{2,n} F^\uparrow - \alpha_{1,n} F^\downarrow - \alpha_{4,n} \frac{S}{\mu_o} \\
\frac{dS}{d\tau} &= -(1 - \omega_n f_n) \frac{S}{\mu_o}
\end{align*}
\]  

(74) (75) (76)

where the factor $(1 - \omega_n f_n)$ results from the delta-scaling approximation and the coefficients $\alpha_{1,n}$ to $\alpha_{4,n}$ are defined by

\[
\begin{align*}
\alpha_{1,n} &= U_1 (1 - \omega_n (1 - \beta_n)) \\
\alpha_{2,n} &= U_1 \beta_n \omega_n \\
\alpha_{3,n} &= (1 - f_n) \omega_n \beta_n (\mu_o) \\
\alpha_{4,n} &= (1 - f_n) \omega_n (1 - \beta_n (\mu_o))
\end{align*}
\]  

(77) (78) (79) (80)

with the fraction of radiation contained in the forward peak $f_n$, the single scattering albedo $\omega_n$, the fractional mean backward scattering coefficient of diffuse light $\beta_n$, the backward scattering coefficient of primary scattered solar radiation $\beta_n (\mu_o)$, and the diffusivity factors of upward and downward radiation $U_\uparrow$, respectively.

For a $N$-layer atmosphere, the general solution of the two-stream model of Zdunkowski et al. [RD37] can be expressed by a linear combination of the internal boundary conditions

\[
\begin{pmatrix}
S_n \\
F_n^\uparrow \\
F_n^\downarrow
\end{pmatrix}
= 
\begin{pmatrix}
\alpha_{1,n} & 0 & 0 & \alpha_{3,n} & \alpha_{5,n} \\
\alpha_{2,n} & \alpha_{4,n} & a_{5,n} \\
\alpha_{3,n} & a_{5,n} & a_{4,n}
\end{pmatrix}
\begin{pmatrix}
S_{n+1} \\
F_{n+1}^\uparrow \\
F_{n+1}^\downarrow
\end{pmatrix}
\]

(see also Eq. (6)) with coefficients

\[
\begin{align*}
\alpha_{1,n} &= \exp \left[ \frac{(1 - \omega_n f_n) \Delta \tau_n}{\mu_o} \right] \\
\alpha_{2,n} &= -a_{4,n} \gamma_{2,n} - a_{5,n} \gamma_{1,n} a_{1,n} + \gamma_{2,n} a_{1,n} \\
\alpha_{3,n} &= -a_{5,n} \gamma_{2,n} - a_{4,n} \gamma_{1,n} a_{1,n} + \gamma_{1,n} \\
a_{4,n} &= E_n \frac{1 - M_n^2}{1 - E_n^2 M_n^2} \\
a_{5,n} &= M_n \frac{1 - E_n^2}{1 - E_n^2 M_n^2}
\end{align*}
\]  

(81) (82) (83) (84) (85)

and with

\[
\begin{align*}
E_n &= \exp(-\epsilon_n \Delta \tau_n); \quad M_n = \frac{a_{2,n}}{a_{1,n} + \epsilon_n}; \quad \epsilon_n = \sqrt{\alpha_{1,n}^2 + a_{2,n}^2} \\
\gamma_{1,n} &= \frac{(1 - \omega_n f_n) a_{3,n} - \mu_o (a_{1,n} a_{3,n} + a_{2,n} a_{4,n})}{(1 - \omega_n f_n) - \epsilon_n^2 \mu_o^2} \\
\gamma_{2,n} &= \frac{(1 - \omega_n f_n) a_{4,n} - \mu_o (a_{1,n} a_{4,n} + a_{2,n} a_{3,n})}{(1 - \omega_n f_n) - \epsilon_n^2 \mu_o^2}
\end{align*}
\]  

(86) (87) (88)

The resonance at $\omega_n \mu_o = (1 - \omega_n f_n)$ can be avoided by changing $\mu_o$ with a small increment.

The numerical effort to calculate the coefficients $\beta$ and $\beta (\mu_o)$ in Eq. (16) and (15) and the fraction $f$ scattered in forward direction can be significant, and so different approaches are suggested to estimate this qualities in a simplified manner. Zdunkowski et al. [RD37] proposed the following approximations:
\[ \ddot{\beta}_n = \frac{3 - p_{1,n}}{8} \]  
(89)

\[ \beta_n(\mu_n) = \frac{1}{2} - \frac{\mu_n p_{1,n} - 3f_n}{4(1 - f_n)} \]  
(90)

\[ f_n = \frac{p_{1,n}^2}{2\tau} \]  
(91)

with the diffusivity factors

\[ U_1 = U_1 = U = 2 \]  
(92)

which is referred as the practical improved flux method (PIFM). Here, the first expansion coefficient of the scattering function \( p_{1,n} \) in model layer \( n \) is also called the asymmetry factor. Due to this definition, the optical properties of the model layer are characterised by the single scattering albedo \( \omega_n \), the optical depth \( \Delta \tau_n \), the asymmetry factor, and the phase function \( P_\delta(\cos \Theta) \) in single scattering geometry.

With these definitions, we can consider the derivatives of matrix \( M \) and the response vector \( R \) in more detail. Starting with the derivative \( \frac{\partial}{\partial \omega_n} \), we obtain

\[ \frac{\partial \alpha_{1,n}}{\partial \omega_n} = -U(1 - \dot{\beta}_n) \]  
(93)

\[ \frac{\partial \alpha_{2,n}}{\partial \omega_n} = -U \dot{\beta}_n \]  
(94)

\[ \frac{\partial \alpha_{3,n}}{\partial \omega_n} = (1 - f_n)\beta_n(\mu_n) \]  
(95)

\[ \frac{\partial \alpha_{4,n}}{\partial \omega_n} = (1 - f_n)(1 - \beta_n(\mu_n)) \]  
(96)

Furthermore,

\[ \frac{\partial \epsilon_n}{\partial \omega_n} = \frac{1}{\epsilon_n} (\alpha_{1,n}^2 - \alpha_{2,n}^2 + \alpha_{3,n}^2 - \alpha_{4,n}^2) \]  
(97)

\[ \frac{\partial E_n}{\partial \omega_n} = \epsilon_n' \tau_n E_n \]  
(98)

\[ \frac{\partial M_n}{\partial \omega_n} = \frac{\alpha_{2,n}'(\alpha_{1,n} + \epsilon_n) - \alpha_{3,n}(\alpha_{1,n}' + \epsilon_n')}{(\alpha_{1,n} + \epsilon_n)^2} \]  
(99)

and

\[ \frac{\partial \gamma_{1,n}}{\partial \omega_n} = \frac{-f_n \alpha_{3,n} + (1 - \omega_n \epsilon_n \rho_n) \alpha_{4,n}' - \mu_n(\alpha_{3,n}' \alpha_{3,n} + \alpha_{1,n} \alpha_{1,n} + \alpha_{2,n} \alpha_{2,n} + \alpha_{4,n} \alpha_{4,n} + \alpha_{3,n} \alpha_{4,n} + \alpha_{2,n} \alpha_{4,n})((1 - \omega_n \epsilon_n \rho_n)^2 - \epsilon_n^2 \mu_n^2)}{(1 - \omega_n \epsilon_n \rho_n)^2 - \epsilon_n^2 \mu_n^2)^2} \]  
(100)

\[ \frac{\partial \gamma_{2,n}}{\partial \omega_n} = \frac{f_n \alpha_{4,n} - (1 - \omega_n \epsilon_n \rho_n) \alpha_{4,n}' - \mu_n(\alpha_{4,n}' \alpha_{4,n} + \alpha_{1,n} \alpha_{3,n} + \alpha_{2,n} \alpha_{4,n} + \alpha_{3,n} \alpha_{4,n} + \alpha_{2,n} \alpha_{3,n} + \alpha_{4,n} \alpha_{4,n})((1 - \omega_n \epsilon_n \rho_n)^2 - \epsilon_n^2 \mu_n^2)}{(1 - \omega_n \epsilon_n \rho_n)^2 - \epsilon_n^2 \mu_n^2)^2} \]  
(101)

To simplify matters in this context, at the right hand side of the equations we use the prime mark to indicate derivatives with respect to the same variable as given at the left hand side. For example, in the equation above \( \epsilon_n' \) means the derivative \( \frac{\partial}{\partial \omega_n} \).

The derivatives with respect to the optical depth \( \Delta \tau_n \) are
For the derivative with respect to the asymmetry factor $p_{1,n}$, we obtain
\begin{align}
\frac{\partial \alpha_{1,n}}{\partial p_{1,n}} &= -\frac{U \omega_n}{8} \\
\frac{\partial \gamma_{1,n}}{\partial p_{1,n}} &= \frac{\omega_n}{p_{1,n}} \left[ \beta_n(\mu_n)(1 - f_n) - \beta_n(\mu_n) \frac{2p_{1,n}}{9} \right] \\
\frac{\partial \gamma_{2,n}}{\partial p_{1,n}} &= -\frac{\omega_n}{p_{1,n}} \left[ \beta_n(\mu_n)(1 - f_n) + (1 - \beta_n(\mu_n)) \frac{2p_{1,n}}{9} \right]
\end{align}

with
\begin{equation}
\beta_n(\mu_n) = \frac{\partial \beta_n(\mu_n)}{\partial p_{1,n}} = -\frac{\mu_0}{4} \left[ 1 - \frac{2p_{1,n}}{9} \right] \frac{p_{1,n} - 3f_n - 2p_{1,n}}{(1-f_n)^2 - 9}
\end{equation}

and
\begin{align}
\frac{\partial \epsilon_n}{\partial p_{1,n}} &= -\frac{U \omega_n}{8 \epsilon_n} (\alpha_{1,n} - \alpha_{2,n}) \\
\frac{\partial \tau_n}{\partial p_{1,n}} &= \epsilon_n' E_n \\
\frac{\partial M_n}{\partial p_{1,n}} &= \frac{\alpha_n'(1 + \epsilon_n) + \alpha_{2,n}(\alpha_n' + \epsilon_n')}{(\alpha_{1,n} + \epsilon_n)^2} \\
\frac{\partial \gamma_{1,n}}{\partial p_{1,n}} &= \frac{1 - \omega_n f_n \alpha_{3,n} - \omega_n a_{3,n} M_{1,n} - 3f_n - 2p_{1,n}}{(1 - \omega_n f_n)^2 - \epsilon_n'^2 \mu_n^2} \\
\frac{\partial \gamma_{2,n}}{\partial p_{1,n}} &= \frac{1 - \omega_n f_n \alpha_{3,n} - \omega_n a_{3,n} M_{1,n} - 3f_n - 2p_{1,n}}{(1 - \omega_n f_n)^2 - \epsilon_n'^2 \mu_n^2} \\
\frac{\partial \gamma_{3,n}}{\partial p_{1,n}} &= \frac{(1 - \omega_n f_n)^2 - \epsilon_n'^2 \mu_n^2}{(1 - \omega_n f_n)^2}
\end{align}

With Eq. 93-114, we can calculate the derivatives of the matrix elements $a_{1,n}$, $a_{2,n}$, $a_{3,n}$, $a_{4,n}$ and $a_{5,n}$ in Eq. (81)-(85):
with respect to the atmospheric parameters \( \Delta \tau_n, \omega_n, \beta_n \) and the phase function \( P_n \). Using Eq. (26), we obtain:

\[
\frac{\partial u_n}{\partial \omega_m} = \delta_{nm} P_n(\bar{\mu})
\]

\[
\frac{\partial v_n}{\partial \omega_m} = U^\top \left[ \delta_{nm}(1 - \beta_n)\eta_n(\mu_v) + \delta_{n-1,m}(1 - \beta_{n-1})\eta_{n-1}(\mu_v) \right]
\]

\[
\frac{\partial w_m}{\partial \omega_m} = U^\top \left[ \delta_{n,m}\beta_n\eta_n(\mu_v) + \delta_{n-1,m}\beta_{n-1}\eta_{n-1}(\mu_v) \right]
\]

and

\[
\frac{\partial u_n}{\partial P_m} = \delta_{n,m}\omega_n\eta_n(\bar{\mu})
\]

\[
\frac{\partial v_n}{\partial P_m} = 0
\]

\[
\frac{\partial w_m}{\partial P_m} = 0
\]

and

\[
\frac{\partial u_n}{\partial \beta_m} = 0
\]

\[
\frac{\partial v_n}{\partial \beta_m} = -U^\top \left[ \delta_{n,m}\omega_n\eta_n(\mu_v) + \delta_{n-1,m}\omega_{n-1}\eta_{n-1}(\mu_v) \right]
\]

\[
\frac{\partial w_m}{\partial \beta_m} = U^\top \left[ \delta_{n,m}\omega_n\eta_n(\mu_v) + \delta_{n-1,m}\omega_{n-1}\eta_{n-1}(\mu_v) \right]
\]

Finally to calculate the Jacobian in Eq. (39), we need the corresponding derivatives of the response vector \( R \) with respect to the atmospheric parameters \( \Delta \tau_n, \omega_n, \beta_n \) and the phase function \( P_n \). Using Eq. (26), we obtain:

\[
R = \frac{1}{4\pi} (u_1, v_1, w_1, \cdots, u_{N+1}, v_{N+1}, w_{N+1})
\]
and

\[
\frac{\partial u_n}{\partial \Delta \tau_m} = \frac{1}{4\pi \mu_o} \omega_n P_n \nu_n^{t_n,m}(\mu) \\
\frac{\partial v_n}{\partial \Delta \tau_m} = \frac{U^\dagger}{4\pi} \left[ \omega_n (1 - \beta_n) \nu_n^{t_n,m}(\mu) + \omega_{n-1} (1 - \beta_{n-1}) \nu_{n-1,m}^{t_{n-1,m}}(\mu) \right] \\
\frac{\partial w_n}{\partial \Delta \tau_m} = \frac{U^\dagger}{4\pi} \left[ \omega_n \beta_n \nu_n^{t_n,m}(\mu) + \omega_{n-1} \beta_{n-1} \nu_{n-1,m}^{t_{n-1,m}}(\mu) \right]
\]

(134)

with the derivative

\[
t_n^{t_n,m}(\mu) = \frac{\partial \nu_n(\mu)}{\partial \Delta \tau_m} = \begin{cases} -\frac{1}{\mu} \nu_n(\mu) & \text{for } m < n \\ \frac{1}{\mu} \exp\left(-\frac{\tau_{n-1}}{\mu}\right) \exp\left(-\frac{\Delta \tau}{\mu}\right) & \text{for } m = n = 1, \cdots, N \\ 0 & \text{for } m > n \end{cases}
\]

(135)

of the auxiliary function \(\nu_n\) in Eq. (23).
B Appendix: SWIR Pre-Processing

To optimise synergies between the different SWIR retrievals, i.e. the CO and HDO/H$_2$O retrieval described in this document and the CH$_4$ retrieval as described by [RD2], a SWIR pre-processing module is developed. Exploiting spectral information from the SWIR spectral ranges, the module provides the following auxiliary products:

- A cloud filter based on a non-scattering retrieval of the total amount of CH$_4$ and a corresponding column estimate from the TM5 forecast (CH$_4$ a priori cloud filter)
- A cloud filter using a non-scattering retrieval of the total amount of CH$_4$ in a strong and weak methane absorption band (CH$_4$ two-band cloud filter)
- A cloud filter using a non-scattering retrieval of the total amount of H$_2$O in a strong and weak water absorption band (H$_2$O two-band cloud filter)

Technically, the pre-processing module is implemented in the CO prototype software to reduce interfaces with the processing framework as indicated in Fig. 2. However, functionally the module does not depend on the CO retrieval. The overall algorithm structure is summarised in Fig. 37. Via the CO interface to the processing framework, the TROPOMI measurements in the SWIR spectral range are allocated in combination with the meteo fields for temperature, pressure and water vapour abundance, information on the pixel's orography and the model forecast of the CH$_4$ and CO profiles. The different auxiliary products are processed in a hierarchical manner to optimise the computational effort. In a first step, only measurements are accepted with a continuum LER value exceeding a certain threshold. Due to this, only measurements are processed with an acceptable signal level. For example, clear sky ocean measurements are rejected by this inquiry due to the very low ocean surface reflection in the SWIR spectral range. Subsequently, the CH$_4$ a priori filter is processed. A comparison

Figure 37: Overall structure of the SWIR preprocessor.
of a non-scattering methane column retrieval with a model a priori methane estimate is used to reject spectral observations contaminated by high and optically thick clouds. This cloud filter is required for the SICOR CO retrieval, which is designed to process clear sky SWIR observation and measurements of scenes with low water clouds and optically thin cirrus and aerosol loads (see Sec. 5). For the processing of CH$_4$ and HDO/H$_2$O an even stricter cloud screening is required. Therefore, the remaining auxiliary products, viz. the CH$_4$ and H$_2$O two-band cloud filters, are only processed for observations which pass the CH$_4$ a priori filter. The retrieval of the auxiliary products are described in more detail below.

Non-scattering total column retrieval

The cloud filters of the SWIR processing are based upon a non-scattering retrieval of a total column abundance from dedicated spectral windows in the SWIR spectral range. Here, we employ the inversion module as described in Sec. 5.2 to infer the total column abundance of the relevant trace gases using the profile scaling approach together with the retrieval of an effective Lambertian surface albedo and a spectral shift of the forward model. Cloud properties are not retrieved from the measurement and so a regularisation of the least squares solution to maintain the numerical stability, as described in Sec. 5.2.3, is not required. Moreover, the forward model employs a radiative transmission model that ignores atmospheric scattering. Here, sunlight is reflected at the Earth surface into the satellite line of sight (LOS) and is attenuated by atmospheric absorption along its path. Using this approximation, the simulated radiance at the TOA $I_{\text{TOA}}(\lambda)$ is given by:

$$I_{\text{TOA}}(\lambda) = A_s(\lambda) \frac{\mu_o F_0}{\pi} \exp\left(-\frac{1}{\tilde{\mu}} \tau_{\text{tot}}(\lambda)\right),$$

(136)

where $A_s(\lambda)$ is the surface albedo, $\mu_o = \cos(\Theta_o)$ with the solar zenith angle $\Theta_o$. This cosine is corrected for the sphericity of the Earth according to Kasten and Young (1989) [RD83]. $F_0$ is the solar irradiance and

$$\frac{1}{\tilde{\mu}} = \frac{\mu_o + \mu_v}{\mu_o \mu_v},$$

(137)

is the air mass factor with $\mu_v = \cos(\Theta_v)$ and viewing zenith angle $\Theta_v$. The total optical thickness $\tau_{\text{tot}}$ is given by

$$\tau_{\text{tot}}(\lambda) = \sum_k \int \sigma_k(z,\lambda) \rho_k(z) \, dz,$$

(138)

where $z$ indicates the altitude, index $k$ represents the relevant absorbers CO, CH$_4$, H$_2$O and HDO, $\rho_k(z)$ is the concentration of absorber $k$ at altitude $z$, $\sigma_k(z,\lambda)$ is the corresponding wavelength-dependent absorption cross sections.

In the following, we assume that the relative profile

$$\rho_k^{\text{rel}} = \frac{\rho_k}{c_k},$$

(139)

of absorber $k$ is constant, where

$$c_k = \int \rho_k(z) \, dz$$

(140)

is the column density of this absorber. So,

$$\frac{\partial \tau_{\text{tot}}}{\partial c_k} = \frac{1}{c_k} \int \sigma_k(z,\lambda) \rho_k(z) \, dz,$$

(141)

and thus the derivative with respect to the total column amount $c_k$ of a trace gas $k$ is given by

$$\frac{\partial I_{\text{TOA}}}{\partial c_k} = -\frac{I_{\text{TOA}}}{\tilde{\mu} c_k} \int \sigma_k(z) \rho_k(z) \, dz.$$

(142)
Finally, the derivative of $I_{TOA}$ with respect to surface albedo $A_s$ is

$$\frac{\partial I_{TOA}}{\partial A_s} = \frac{\mu_o F_o}{\pi} \exp\left(\frac{1}{\mu} \tau_{\text{tot}}\right).$$ (143)

To account for the spectral instrument response, the TOA radiance and its derivatives have to be convoluted with the ISRF as described in Eqs.(3) and (5).

**CH$_4$ a priori cloud filter**

The methane a priori cloud filter relies on a non-scattering methane column retrieval. In the presence of high and optically thick clouds, the lightpath gets effectively shortened due to reflection of light by clouds. Thus assuming a clear sky model atmosphere, the retrieved methane total column is underestimated depending on cloud occurrence in the observed scene. Figure 38 shows the induced error in the non-scattering methane column retrieved from the spectral window 2315–2324 nm for two cloudy scenes. For a water cloud with an optical depth of 5, the methane error increases with cloud coverage and cloud height. In case of a cirrus at 10 km height, the non-scattering CH$_4$ column can be used to identify cloud cirrus contamination with optical depth > 0.5 for low and moderate surface reflection. However for bright surfaces, the lightpath shortening due the reflection of light by the cirrus is compensated by an enhancement of the lightpath because of multiple reflection of light between the cirrus layer and the surface. In this case, the non-scattering CH$_4$ column is less suited for cirrus detection. Keeping in mind that the current accuracy of the methane column forecast is in the order of 2–3 % using state-of-the-art chemical transport models (see Sec. 7.6), the CH$_4$ a priori cloud filter is a powerful tool to screen measurement with respect to the presence of high and optically thick clouds.

![Figure 38: CH$_4$ error of a non-scattering retrieval from the SWIR 2315–2324 nm spectral window for a water cloud with optical thickness of 5 as function of cloud height and cloud fraction (left panel, for more details see generic scenario B in Sec. 7.1) and for a cirrus cloud at 10 km height as function of surface albedo and cirrus optical thickness (right panel, for more details see generic scenario E in Sec. 7.1).](image)

**CH$_4$ two-band cloud filter**

Complementary to the CH$_4$ a priori filter, the SWIR pre-processor includes a cloud filter based on two methane non-scattering retrievals that utilise two different spectral bands with a strong and weak absorption band, respectively. This method relies on the fact that in a strong absorption band, photons along an enhanced light path contribute less to the total signal than in a weak absorption band. Thus, we expect that the non-scattering methane column, which are inferred from a weak absorption, exceeds the corresponding column retrieved from a strong absorption band. Moreover, the difference between the two columns indicates changes of the lightpath due to atmospheric scattering and so can be used as a cloud filter. In Fig. 39, this difference is depicted for the cloud scenarios of Fig. 38 using the strong CH$_4$ absorption at 2363-2373 nm and the weak absorption at 2310-2315 nm. The figure indicates that the methane two-band retrieval is well suited to detect cirrus cloud even above bright surfaces and complements the CH$_4$ a priori filter. Therefore, this cloud filter is particular valuable for cirrus screening needed by the CH$_4$ and HDO/H$_2$O data processing.
Figure 39: CH\textsubscript{4} two-band cloud filter for the cloud scenarios of Fig. 38. The methane cloud filter relies on non-scattering methane column retrieval from strong and weak absorption features at 2363-2373 nm and 2310-2315 nm, respectively.

H\textsubscript{2}O two-band cloud filter

The concept of the two-band CH\textsubscript{4} cloud filter can be applied to H\textsubscript{2}O as well. The two-band cloud filter does not rely on a-priori knowledge, so the poor a-priori knowledge of H\textsubscript{2}O does not hamper the application of the two-band cloud filter for water. Compared to CH\textsubscript{4}, H\textsubscript{2}O absorption lines are narrower, resulting in a different sensitivity towards pressure broadening close to the surface. The H\textsubscript{2}O two-band cloud filter is set up analogous to the CH\textsubscript{4} two-band cloud filter. Here, the window 2329-2334 nm contains weak H\textsubscript{2}O absorption and the window 2367-2377 nm includes strong H\textsubscript{2}O absorption features.

Figure 40: H\textsubscript{2}O two-band cloud filter for the cloud scenarios of Fig. 38. The filter relies on non-scattering methane column retrieval from strong and weak absorption features at 2367-2377 nm and 2329-2334 nm, respectively.

Figure 40 shows the relative difference in the retrieved water columns using the weak and strong absorption bands. Compared to the CH\textsubscript{4} two-band cloud filter in Fig. 39, the H\textsubscript{2}O two-band filter shows larger sensitivity to clouds and so can be used to compliment the other cloud filters.