

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Date : 17 May 2019	Page 1/74

Sentinel 5 L2 Prototype Processors

Algorithm Theoretical Baseline Document: Carbon Monoxide Retrieval

J. Landgraf, J. aan de Brugh, T. Borsdorff, H. Hu, O.P. Hasekamp

Prepared by : Jochen Landgraf Joost aan de Brugh	
Checked by : Jochen Landgraf Joost aan de Brugh	
Approved by : Jochen Landgraf	

Citation: Landgraf, J., aan de Brugh, J., Borsdorff, T., Hu, H., Hasekamp, O.P., Sentinel 5 L2 Prototype Processors, Algorithm Theoretical Baseline Document: Carbon Monoxide Retrieval, SRON-ESA-S5L2PP-ATBD-002-v3.1-20190517-CO, SRON Netherlands Institute for Space Research, Utrecht, The Netherlands, 2019

The copyright in this document is vested in SRON. This document may only be reproduced in whole or in part, or stored in a retrieval system, or transmitted in any form, or by any means electronic, mechanical, photocopying or otherwise, either with the prior permission of SRON or in accordance with the terms of ESA Contract No. 4000118463/16/NL/AI.

Change Log

Issue	Date	Item	Comments
1.0	2017-02-15	All	Document creation
2.0	2017-10-02	Sec. 6.9	Section 'Detailed algorithm setup' added
2.0	2017-10-02	Sec. 7	Preliminary error analysis added based on S5P work
2.0	2017-10-02	Sec. 8	Section on product validation added
2.0	2017-10-02	Sec. 9	Section includes detailed input/output information and a first estimate on required computational resources.
2.1	2018-03-08	All	Applied corrections from PDR RIDs. General revision of whole document.
2.2	2018-11-14	Sec. 9 and Appendix B	Added breakpoint output, revised processing quality flags. Added algorithm settings appendix.
2.2	2018-12-15	Sec. 9	Updated computational effort.
2.2	2018-12-15	Sec. 6.9 and 9	Updated several acronyms.
3.0	2019-04-25	All	CDR RIDs are taken into account. In particular, the I/O tables are updated.

     	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Date : 17 May 2019	Page 3/74
--	--------------------------------	---	--------------

Table of Contents

Change Log	2
List of Tables	4
List of Figures	4
1 Introduction	7
1.1 Purpose and Objective	7
1.2 Document overview	7
2 Applicable and Reference Documents	8
2.1 Applicable Documents	8
2.2 Reference Documents	8
3 Terms, Definitions and Abbreviated Terms	14
3.1 Terms and Definitions	14
3.2 Acronyms and Abbreviations	15
4 Introduction to the carbon monoxide algorithm SICOR	17
4.1 Heritage	17
4.2 Requirements	18
5 Instrument Overview	20
6 Detailed Algorithm Description	21
6.1 Input	22
6.1.1 Absorption cross sections	22
6.1.2 METimage	23
6.2 Standard retrieval setup	23
6.3 Initial data filtering A	25
6.4 Non-scattering methane retrieval	25
6.4.1 Non-scattering forward model	25
6.4.2 Non-scattering inversion	27
6.5 Cloud filter B	27
6.6 Physics-based retrieval	28
6.6.1 Forward model setup	29
6.6.2 The two-stream radiative transfer model	30
6.6.3 The forward-adjoint perturbation theory applied to the two-stream method	33
6.6.4 Physics-based CO column inversion	34
6.6.5 Inversion stability measures	37
6.7 Deconvolution solar spectrum	38
6.8 State vector, ancillary parameters and a priori knowledge	39
6.9 Detailed Algorithm Setup	40
6.9.1 Serial Initialization	40
6.9.2 Parallel initialization	40
6.9.3 Retrievals	42
7 Error Analysis	46
7.1 Forward model errors	48
7.2 Atmospheric parameters	50
7.3 Instrument effects	52
7.4 Sentinel 5 orbit ensemble	52

     	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002	
		Version : 3.1	Page
		Date : 17 May 2019	4/74

8	Validation	53
9	Feasibility	55
9.1	Estimated Computational Effort	55
9.2	Inputs	55
9.2.1	Static Inputs	57
9.2.2	Dynamic Inputs	58
9.3	Outputs	61
9.3.1	Static output	61
9.3.2	Dynamic output	61
9.3.3	Breakpoint Output	62
	Appendices	62
A	Appendix: Flux method PIFM	62
B	Algorithm settings	69
B.1	Filter settings	69
B.2	Atmospheric gridding	69
B.3	Retrieval settings	69
C	Usage of L1b-flagging at framework level	72

List of Tables

1	Spectral fit windows of SICOR based on the S5 band definition [AD2].	21
2	The S5 L2 CLA data product used by SICOR. For more detail see Tab. 6	23
3	Composition of of state vector \mathbf{x} . LER means Lambert-equivalent reflectivity.	39
4	Static input of the serial algorithm initialization.	41
5	Output data of the static initialization.	41
6	Dynamic input to the CO retrieval algorithm. Metimage input as specified in Tab. 2	44
7	Dynamic output data	45
8	Microphysical properties of water and ice clouds: $n(r)$ represents the size distribution type, r_{eff} and v_{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].	46
9	S5 instrument noise model	46
10	CO column retrieval sensitivity in % with respect to knowledge uncertainties of a set of atmospheric and instrument parameters for the generic clear sky, cloud and cirrus ensemble: (1) CH ₄ a priori uncertainty of TM5-NOAA runs, (2) ECMWF surface pressure uncertainty, (3) ECMWF temperature profile offset, (5) FWHM uncertainty of the ISRF, (6) spectral calibration error δ_s and (7) the radiometric offset I_{offset} and a multiplicative radiometric error I_{scal} of the level 1 data product.	52
11	Calculation settings and computation time for the S5 orbit ensemble.	55
12	Static input data of SICOR-S5.	57
13	Dynamic input of SICOR.	59
14	Output data of the static initialization.	61
15	Dynamic output data of the data processor per ground pixel.	61
16	Processing Quality Flags (PQFs) of SICOR. Four types of flags are provided: status (S), error (E), filter (F) and warning (W)	62
17	Breakpoint output of SICOR.	62
18	Filter settings used in SICOR-S5.	69
19	Atmospheric settings used in SICOR.	69
20	Retrieval settings used in SICOR.	69
21	State parameter settings used in SICOR.	70
22	Scattering settings used in SICOR.	71
23	Usage of L1b flags by algorithm.	74

List of Figures

1	Spectral transmittance along the light path of the solar beam reflected by the Earth surface into the instrument viewing direction for the 2.3 μm band. Simulations are performed for viewing zenith angle (VZA) of 0° , and a solar zenith angle (SZA) of 30° , and by assuming a US standard atmospheric profile. From top to bottom, the figure shows the total transmittance, the individual transmittances due to H_2O , CH_4 , and CO , respectively. Note the different y-axis scale for CO transmittance.	19
2	Overall processing flowchart of the CO retrieval algorithm.	22
3	Relative CH_4 column above a cloud top height z_{cld} with respect to the total column amount, using the US standard model atmosphere.	26
4	CH_4 error of a non-scattering retrieval from the SWIR-3 2315–2324 nm spectral window for a cloud with optical thickness of 5 as function of cloud height and cloud fraction (left panel) and for a cirrus cloud at 10 km height as function of surface albedo and cirrus optical thickness (right panel).	28
5	Probability density function (left panel) and cumulative probability density function (right panel) of the difference ΔCH_4 of one year of GOSAT non-scattering retrievals (2010) compared with the corresponding TM5 model simulations. The figure differentiates between the contribution of ocean and land pixels (blue and green areas). The dataset comprises $2.4 \cdot 10^6$ GOSAT measurements in total, under which $1.6 \cdot 10^6$ ocean pixels and $8 \cdot 10^5$ land pixels. All retrievals are performed using RemoTeC V2.1.	29
6	SICOR algorithm architecture including the algorithm initialization and the pixel processing. Here, two parallelization options are foreseen: first, the initialization of the algorithm with respect to the detector spatial channel index of the measurements and second, the actual retrievals for the pixels to be processed within a data granule.	40
7	Serial SICOR initialization, which is executed once per processor call. The solid green boxes indicate static input, the dashed green boxes indicate initialization results accessible for the remaining SICOR software.	41
8	Parallel initialization: The ISRF is interpolated for each spectral pixel. Subsequently, the spectral shift between solar reference spectrum and observation is determined. This is needed as input for the subsequent deconvolution of the solar spectrum.	42
9	Software architecture of the S5 CO retrieval tool. The shaded boxes indicate two specific calls of the retrieval model (non-scattering retrieval and physics-based retrieval), which is explained in more detail in Fig. 10.	43
10	Software architecture of the non-scattering and physics-based retrieval.	44
11	Example of the S5 CO data product and its performance as a function of cloud fraction f_{cld} . The SWIR-3 measurements are simulated for a scene partially covered by a cloud between 2 and 3 km with optical depth $\tau_{\text{cld}} = 5$, a surface albedo $A_s = 0.05$, a solar zenith angle of 50° and a viewing zenith angle of 40° . Left panel: CO retrieval bias b_{CO} . Middle panel: retrieval noise σ_{CO} . Right panel: column averaging kernel for different cloud fractions as indicated in the legend. The grey area indicates the position of the cloud.	47
12	Retrieval bias b_{CO} (left panel) and retrieval noise σ_{CO} (right panel) for the clear sky conditions (without aerosol, clouds and cirrus) and for a viewing zenith angle (VZA) of 0 degree as a function of solar zenith angle (SZA) and surface albedo A_s	48
13	Left panel: Retrieval bias in case of a cloud atmosphere. The CO bias is shown as a function of surface albedo A_s and cloud fraction f for a cloud between 4 and 5 km altitude with optical depth $\tau_{\text{scat}} = 2$ and a VZA of 0 degree. Right panel: CO retrieval bias for measurements in presence of optically thin cirrus, which overcasts the entire scene, as a function of surface albedo and cirrus optical depth that defined at 2300 nm. The grey area indicates measurement simulations, which were rejected by the cloud filter.	48
14	Retrieval bias b_{CO} in cloudy atmospheres in case of strongly enhanced CO concentrations. Measurement simulations are performed for a surface albedo $A_s = 0.05$, SZA and VZA of 50 and 0 degree and for overcast sky with a cloud at 1–2 km altitude with an optical depth of $\tau_{\text{cld}} = 2$ (pink) and $\tau_{\text{cld}} = 5$ (blue). Additionally, we consider a case of partially cloud cover with cloud fraction $f_{\text{cld}} = 0.1$ at 4–5 km altitude with $\tau_{\text{cld}} = 2$ (yellow). The CO profile represents the US standard atmosphere with a perturbation at the indicated altitude z_{per} enhancing the total amount of CO by 50%.	49

15	Difference between CH ₄ total column dry air mixing ratios from TM5-NOAA simulations and GOSAT retrievals for the period June 2009 to December 2012.	50
16	CO bias due to a priori errors in CH ₄ for the clear sky measurement ensemble of Fig. 12. For each CH ₄ error, the CO bias probability function is shown. The CO error sensitivity is estimated by a linear regression through all data points (solid line).	51
17	Fraction of GOSAT non-scattering retrievals which are accepted by the cloud filter $\Delta\text{CH}_4 < 25\%$. The analysis is based on one year (2010) of GOSAT non-scattering retrievals (RemoTeC V2.1).	56

 	<p style="text-align: center;">S5L2PP Carbon Monoxide ATBD</p>	<p>Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 7/74</p>
--	--	--

1 Introduction

1.1 Purpose and Objective

This document describes the carbon monoxide (CO) column retrieval algorithm using Sentinel-5 measurements in the shortwave infrared (SWIR) spectral range between 2310 and 2340 nm. The algorithm name is SICOR-S5 and it is one of the deliverables of the ESA project 'Sentinel-5 level 2 prototype processor' [AD3]. The purpose of the document is to describe the theoretical baseline of the algorithm, the input and ancillary data, and the output that is generated. Additionally, we estimate the algorithm runtime, the product precision and accuracy.

1.2 Document overview

Chapter 2 provides the document's references and Chapter 3 contains a list of abbreviations used in this document. Chapter 4 gives a short introduction to satellite remote sensing of atmospheric CO abundance and the heritage of the presented algorithm is summarised. Moreover, we recall the level-2 requirement for the CO column product, which represents the underlying criterion for the performance analysis of the SICOR-S5 algorithm. The theoretical concept of the CO retrieval algorithm SICOR is given in Chapter 6, 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. The numerical feasibility is the subject of Chapter 9, which includes an estimate of the numerical effort, a high level data product description and the spatial data selection criteria of the measurements to be processed. Chapter 7 gives a detailed error analysis, considers the performance of the retrieval algorithm based on a set of generic measurement ensembles and a geo-physical ensemble of simulated measurements of one S5 orbit. Here, we investigate the CO retrieval noise and CO retrieval biases due to forward model errors, erroneous atmospheric input data and possible instrument artifacts. Based on this, we evaluate the algorithm performance in the context of the S5 level-1 and 2 requirements. Finally, Chapter 8 discusses a strategy for the product validation. Additional material is provided in the appendices, where Appendix A discusses in detail the linearized two-stream method 2S-LINTRAN.

	S5L2PP	Reference : SRON-ESA-S5L2PP-ATBD-002
	Carbon Monoxide ATBD	Version : 3.1 Page
		Date : 17 May 2019 8/74

2 Applicable and Reference Documents

2.1 Applicable Documents

- [AD1] S5L2PP Project Management Plan.
source: S&T; **ref:** ST-ESA-S5L2PP-PMP-001; **issue:** 2.4; **date:** 2019-05-17.
- [AD2] GMES Sentinels 4 and 5 mission requirements document.
source: ESA; **ref:** EOP-SMA/1507/JL-dr; **date:** 2011.
- [AD3] Sentinel-5 Level-2 Prototype Processor Development - Statement of Work.
source: ESA; **ref:** S5-SOW-ESA-GR-0130; **issue:** 1.7; **date:** 2018-06-28.
- [AD4] Sentinel-5 Level-2 Prototype Processor Development Requirements Specification.
source: ESA; **ref:** S5-RS-ESA-GR-0131; **issue:** 1.7; **date:** 2018-06-29.

2.2 Reference Documents

- [RD1] A. J. Heymsfield and C. M. R. Platt; A Parameterization of the Particle Size Spectrum of Ice Clouds in Terms of the Ambient Temperature and the Ice Water Content. *J. Atmos. Sci.*; **41** (1984), 846; 10.1175/1520-0469(1984)041<0846:APOTPS>2.0.CO;2.
- [RD2] S5L2PP Terms, Definitions and Abbreviations.
source: S&T; **ref:** ST-ESA-S5L2PP-LST-001; **issue:** 2.1; **date:** 2018-08-28.
- [RD3] A. Hollingsworth, R. J. Engelen, A. Benedetti *et al.*; Toward a Monitoring and Forecasting System For Atmospheric Composition: The GEMS Project. *Bull. Amer. Meteor. Soc.*; **89** (2008), 1147.
- [RD4] H.H. Levy; Normal atmosphere: Large radical and formaldehyde concentrations predicted. *Science*; **173** (1971), 141.
- [RD5] J.A. Logan, M.J. Prather, S.C. Wofsy *et al.*; Tropospheric chemistry: A global perspective. *J. Geophys. Res.*; **86** (1981), 7210.
- [RD6] D.T. Shindell, G. Faluvegi, D.S. Stevenson *et al.*; Multi-model simulations of carbon monoxide: Comparison with observations and projected near-future changes. *J. Geophys. Res.*; **111** (2006), D19306, doi: 10.1029/2006JD007100.
- [RD7] D.P. Edwards, L.K. Emmons, D.A. Hauglustaine *et al.*; Observations of carbon monoxide and aerosols from the Terra satellite: Northern Hemisphere variability. *J. Geophys. Res.*; **109** (2004), D24202, doi: 10.1029/2004JD004727.
- [RD8] H.G. Reichle Jr and V.S. Connors; The mass of CO in the atmosphere during October 1984, April 1994, and October 1994. *J. Atmos. Sci.*; **56** (1999), 307.
- [RD9] H. Kobayashi, A. Shimota, K. Kondo *et al.*; Development and evaluation of the Interferometric Monitor for Greenhouse Gases: a high throughput Fourier transform infrared radiometer for nadir Earth observations. *Appl. Opt.*; **38** (1999), 6801.
- [RD10] M.N. Deeter, L. K. Emmons, G. L. Francis *et al.*; Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument. *J. Geophys. Res.*; **108** (2003), 4399, doi:10.1029/2002JD003186.
- [RD11] W.W. McMillan, C. Barnet, L. Strow *et al.*; Daily global maps of carbon monoxide from NASA's Atmospheric Infrared Sounder. *Geophys. Res. Lett.*; **32** (2005), L11801, doi:10.1029/2004GL021821.
- [RD12] C. P. Rinsland, M. Luo, J. A. Logan *et al.*; Nadir measurements of carbon monoxide distributions by the Tropospheric Emission Spectrometer instrument onboard the Aura Spacecraft: Overview of analysis approach and examples of initial results. *Geophys. Res. Lett.*; **33** (2006), L22806, doi:10.1029/2006GL027000.
- [RD13] S. Turquety, J. Hadji-Lazaro, C. Clerbaux *et al.*; Operational trace gas retrieval algorithm for the Infrared Atmospheric Sounding Interferometer. *J. Geophys. Res.*; **109** (2004), D21301, doi:10.1029/2004JD004821.

	S5L2PP	Reference : SRON-ESA-S5L2PP-ATBD-002
	Carbon Monoxide ATBD	Version : 3.1 Page
		Date : 17 May 2019 9/74

- [RD14] 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.
- [RD15] H. Bovensmann, J. P. Burrows, M. Buchwitz *et al.*; SCIAMACHY: Mission Objectives and Measurement Modes. *Journal of Atmospheric Sciences*; **56** (1999), 127.
- [RD16] M.N. Deeter, D. P. Edwards, J. C. Gille *et al.*; CO retrievals based on MOPITT near-infrared observations. *J. Geophys. Res.*; **114** (2009), D04303, doi:10.1029/2008JD010872.
- [RD17] H. M. Worden, M. N. Deeter, D. P. Edwards *et al.*; Observations of near-surface carbon monoxide from space using MOPITT multispectral retrievals. *J. Geophys. Res.*; **115** (2010), D18314, doi:10.1029/2010JD014242.
- [RD18] P. Veefkind; TROPOMI on the ESA Sentinel-5 Precursor: a GMES mission for Global Observations of the Atmospheric Composition for Climate and Air Quality Applications. *Remote Sens. Environ.*; **120** (2012), 70.
- [RD19] Paul Ingmann, Ben Veihelmann, Jörg Langen *et al.*; Requirements for the GMES Atmosphere Service and ESA's implementation concept: Sentinels-4/-5 and -5p. *Remote Sensing of Environment*; **120** (2012), 58 ; <https://doi.org/10.1016/j.rse.2012.01.023>. The Sentinel Missions - New Opportunities for Science; URL <http://www.sciencedirect.com/science/article/pii/S0034425712000673>.
- [RD20] Sentinel-5/UVNS L1BPP Algorithm Theoretical Basis Document (ATBD). **source:** AIRBUS; **ref:** GS5.RP.ASG.UVNS.00044; **issue:** 7.0; **date:** 2017-05-08.
- [RD21] M. Buchwitz, I. Khlystova, H. Bovensmann *et al.*; Three years of global carbon monoxide from SCIAMACHY: comparison with MOPITT and first results related to the detection of enhanced CO over cities. *Atmos. Chem. Phys.*; **7** (2007), 2399.
- [RD22] C Frankenberg, U. Platt and T. Wagner; Retrieval of CO from SCIAMACHY onboard ENVISAT: detection of strongly polluted areas and seasonal patterns in global CO abundances. *Atmos. Chem. Phys.*; **4** (2005), 8425.
- [RD23] A.M.S. Gloudemans, A.T.J. de Laat, H. Schrijver *et al.*; SCIAMACHY CO over land and oceans: 2003-2007 interannual variability. *Atmos. Chem. Phys.*; **9** (2009), 3799.
- [RD24] M. Buchwitz, R. de Beek, S. Noel *et al.*; Atmospheric carbon gases retrieved from SCIAMACHY by WFM-DOAS: version 0.5 CO and CH₄ and impact of calibration improvements on CO₂ retrieval. *Atmos. Chem. Phys.*; **6** (2006), 2727.
- [RD25] S. Turquety, C. Clerbaux, K. Law *et al.*; CO emission and export from Asia: an analysis combining complementary satellite measurements (MOPITT, SCIAMACHY and ACE-FTS) with global modeling. *Atmos. Chem. Phys.*; **8** (2008) (17), 5187.
- [RD26] A. T. J. de Laat, A. M. S. Gloudemans, I. Aben *et al.*; Global evaluation of SCIAMACHY and MOPITT carbon monoxide column differences for 2004-2005. *J. Geophys. Res.*; **115** (2010), D06307, doi:10.1029/2009JD012698.
- [RD27] B. Dils, M. De Maziere, J. F. Muller *et al.*; Comparisons between SCIAMACHY and ground-based FTIR data for total columns of CO, CH₄, CO₂ and N₂O. *Atmos. Chem. Phys.*; **6** (2006), 1953.
- [RD28] A. T. J. de Laat, A. M. S. Gloudemans, H. Schrijver *et al.*; Validation of five years (2003-2007) of SCIAMACHY CO total column measurements using ground-based spectrometer observations. *Atmos. Meas. Tech.*; **3** (2010), 1457.
- [RD29] A. T. J. de Laat, A. M. S. Gloudemans, I. Aben *et al.*; Scanning Imaging Absorption Spectroscopy for Atmospheric Chartography carbon monoxide total columns: Statistical evaluation and comparison with chemistry transport model results. *J. Geophys. Res.*; **112** (2007), D12310, doi:10.1029/2006JD008256.
- [RD30] T. Borsdorff, P. Tol, J. E. Williams *et al.*; Carbon monoxide total columns from SCIAMACHY 2.3 μ m atmospheric reflectance measurements: towards a full-mission data product (2003-2012). *amt*; **9** (2016), 227; 10.5194/amt-9-227-2016.
- [RD31] T. Borsdorff, J. aan de Brugh, H. Hu *et al.*; Carbon monoxide column retrieval for clear-sky and cloudy atmospheres: a full-mission data set from SCIAMACHY 2.3 μ m reflectance measurements. *amtd*; (2017); 10.5194/amt-2016-355.

	S5L2PP	Reference : SRON-ESA-S5L2PP-ATBD-002
	Carbon Monoxide ATBD	Version : 3.1 Page
		Date : 17 May 2019 10/74

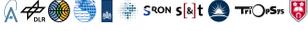
- [RD32] Sentinel-5/UVNS L1BPP Algorithm Theoretical Baseline Document.
source: AIRBUS; **ref:** GS5.RP.ASG.UVNS.00044; **issue:** 8.0; **date:** 03-10-2017.
- [RD33] J. Vidot, J. Landgraf, O.P. Hasekamp *et al.*; Carbon monoxide from shortwave infrared reflectance measurements: A new retrieval approach for clear sky and partially cloudy atmospheres. *Remote Sens. Environ.*; **120** (2012), 255.
- [RD34] Algorithm Theoretical Baseline Document for Sentinel-5 Precursor: Carbon Monoxide Total Column Retrieval.
source: SRON; **ref:** SRON-S5P-LEV2-RP-002; **date:** 2016.
- [RD35] J. Landgraf, J. aan de Brugh, R. Scheepmaker *et al.*; Carbon monoxide total column retrievals from TROPOMI shortwave infrared measurements. *Atmospheric Measurement Techniques*; **9** (2016), 4955; 10.5194/amt-9-4955-2016.
- [RD36] Sentinel 5 L2 Prototype Processors: Algorithm Theoretical Baseline Document S5-METimage Product.
source: STFC Rutherford Appleton Laboratory; **ref:** RAL-ESA-S5L2PP-ATBD-002; **issue:** 3.0; **date:** 15-12-2018.
- [RD37] J. Danielson and D. Gesch; Global multi-resolution terrain elevation data 2010 (GMTED2010). *U.S. Geological Survey Open-File Report*; (2011), 2011.
- [RD38] A. Jenouvrier, L. Daumont, L. Régalia-Jarlot *et al.*; Fourier transform measurements of water vapor line parameters in the 4200–6600 cm^{-1} region. *J. Quant. Spectrosc. Radiat. Transfer*; **105** (2007), 326; 10.1016/j.jqsrt.2006.11.007.
- [RD39] L.S. Rothman, I.E. Gordon, A. Barbe *et al.*; The HITRAN 2008 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transfer*; **110** (2009), 533.
- [RD40] R. A. Scheepmaker, C. Frankenberg, A. Galli *et al.*; Improved water vapour spectroscopy in the 4174–4300 cm^{-1} region and its impact on SCIAMACHY HDO/H₂O measurements. *Atmos. Meas. Tech.*; **6** (2013) (4), 879; 10.5194/amt-6-879-2013. URL <http://www.atmos-meas-tech.net/6/879/2013/>.
- [RD41] C. Frankenberg, D. Wunch, G. Toon *et al.*; Water vapor isotopologue retrievals from high-resolution GOSAT shortwave infrared spectra. *Atmospheric Measurement Techniques*; **6** (2013), 263; 10.5194/amt-6-263-2013.
- [RD42] W.H. Press, S.A. Teukolsky, W.T. Vetterling *et al.*; *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge University Press, 1992); 2nd edition.
- [RD43] A. Butz, A. Galli, O.P. Hasekamp *et al.*; ROPOMI aboard Sentinel-5 Precursor: Prospective performance of CH₄ retrievals for aerosol and cirrus loaded atmospheres. *Remote Sens. Environ.*; **120** (2012), 267.
- [RD44] F. Kasten and T. Young; Revised optical air mass tables and approximation formula. *Appl. Opt.*; **28** (1989), 4735.
- [RD45] OMI aerosol retrieval algorithm.
source: KNMI; **ref:** ATBD-OMI-03; **date:** 2001.
- [RD46] A.M.S. Gloudemans, H. Schrijver, O.P. Hasekamp *et al.*; Error analysis for CO and CH₄ total column retrieval from SCIAMACHY 2.3 μm spectra. *Atmos. Chem. Phys.*; **8** (2008), 3999.
- [RD47] W.G. Zdunkowski, Welch R.M. and Korb G.; An Investigation of the Structure of Typical Two-stream-methods for the Calculation of Solar Fluxes and Heating Rates in Clouds. *Contrib. Atmos. Phys.*; **53** (1979), 147.
- [RD48] W. J. Wiscombe and G. W. Grams; The backscattered fraction in two-stream approximations. *Journal of Atmospheric Sciences*; **33** (1976), 2440.
- [RD49] G.I. Marchuk; Equation for the Value of Information from Weather Satellites and Formulation of Inverse Problems. *Cosmic Res.*; **2** (1964), 394.
- [RD50] M.A. Box, S.A.W. Gerstl and C. Simmer; Application of the adjoint formulation to the calculation of atmospheric radiative effects. *Beitr. Phys. Atmos.*; **61** (1988), 303.
- [RD51] J. Landgraf, O.P. Hasekamp, M. Box *et al.*; A Linearized Radiative Transfer Model Using the Analytical Perturbation Approach. *J. Geophys. Res.*; **106** (2001), 27291.

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002
		Version : 3.1 Page
		Date : 17 May 2019 11/74

- [RD52] O.P. Hasekamp and J. Landgraf; A linearized vector radiative transfer model for atmospheric trace gas retrieval. *J. Quant. Spectrosc. Radiat. Transfer*; **75** (2002).
- [RD53] V.V. Rozanov, D. Diebel, R.J.D. Spurr *et al.*; GOMETRAN: A radiative transfer model for the satellite project GOME, the plan-parallel version. *J. Geophys. Res.*; **102** (1997), 16,683.
- [RD54] V.V. Rozanov, T. Kurosu and J.P. Burrows; Retrieval of Atmospheric Constituents in the UV-Visible: A new Quasi-Analytical Approach for the Calculation of Weighting Functions. *J. Quant. Spectrosc. Radiat. Transfer*; **60** (1998), 277.
- [RD55] H. H. Walter, J. Landgraf and O. P. Hasekamp; Linearization of a pseudo-spherical vector radiative transfer model. *J. Quant. Spectrosc. Radiat. Transfer*; **85** (2004), 251.
- [RD56] H. H. Walter and J. Landgraf; Towards linearization of atmospheric radiative transfer in spherical geometry. *J. Quant. Spectrosc. Radiat. Transfer*; **95** (2005), 175.
- [RD57] H. H. Walter, J. Landgraf, F. Spada *et al.*; Linearization of a radiative transfer model in spherical geometry. *J. Geophys. Res.*; **111** (2006), 10.1029/2005JD007014.
- [RD58] E. A. Ustinov; Adjoint sensitivity analysis of radiative transfer equation: Temperature and gas mixing ratio weighting functions for remote sensing of scattering atmospheres in thermal IR. *J. Quant. Spectrosc. Radiat. Transfer*; **68** (2001), 195.
- [RD59] J. Landgraf, O.P. Hasekamp and T. Trautmann; Linearization of radiative transfer with respect to surface properties. *J. Quant. Spectrosc. Radiat. Transfer*; **72** (2002), 327.
- [RD60] G. I. Bell and S. Glasstone; *Nuclear Reactor Theory* (Van Nostrand Reinhold Company, New York, 1970).
- [RD61] J. Lewins; *Importance, the adjoint function* (Pergamon Press, Oxford, England, 1965).
- [RD62] E. A. Ustinov; Inverse problem of photometric observation of solar radiation reflected by an optically dense planetary atmosphere. Mathematical methods and weighting functions of linearized inverse problem. *Cosmic Res.*; **29** (1991), 519.
- [RD63] P.L. Phillips; A technique for the numerical solution of certain integral equations of the first kind. *J. Ass. Comput. Mat.*; **9** (1962), 84.
- [RD64] A.N. Tikhonov; On the solution of incorrectly stated problems and a method of regularization. *Dokl. Acad. Nauk SSSR*; **151** (1963), 501.
- [RD65] T. Borsdorff, O. P. Hasekamp, A. Wassmann *et al.*; Insights into Tikhonov regularization: application to trace gas column retrieval and the efficient calculation of total column averaging kernels. *Atmospheric Measurement Techniques*; **7** (2014), 523; 10.5194/amt-7-523-2014.
- [RD66] C.D. Rodgers; *Inverse Methods for Atmospheres: Theory and Practice*; volume 2 (World Scientific, 2000).
- [RD67] R. van Deelen, O. P. Hasekamp and J. Landgraf; Accurate modeling of spectral fine-structure in Earth radiance spectra measured with the Global Ozone Monitoring Experiment. *Appl. Opt.*; **46** (2007), 243; 10.1364/AO.46.000243.
- [RD68] A. Wassmann, T. Borsdorff, J. M. J. aan de Brugh *et al.*; The direct fitting approach for total ozone column retrievals: a sensitivity study on GOME-2/MetOp-A measurements. *Atmospheric Measurement Techniques*; **8** (2015), 4429; 10.5194/amt-8-4429-2015.
- [RD69] D. Schepers, J.M.J. aan de Brugh, Ph. Hahne *et al.*; LINTRAN v2.0: A linearised vector radiative transfer model for efficient simulation of satellite-born nadir-viewing reflection measurements of cloudy atmospheres. *Journal of Quantitative Spectroscopy and Radiative Transfer*; **149** (2014), 347 ; <https://doi.org/10.1016/j.jqsrt.2014.08.019>. URL <http://www.sciencedirect.com/science/article/pii/S002240731400363X>.
- [RD70] M. Hess and M. Wiegner; COP: a data library of optical properties of hexagonal ice crystals. *Appl. Opt.*; **33** (1994) (33), 7740; 10.1364/AO.33.007740.

 	<p style="text-align: center;">S5L2PP Carbon Monoxide ATBD</p>	<p>Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 12/74</p>
--	--	---

- [RD71] M. Hess, R. B. A. Koelemeijer and P. Stammes; Scattering matrices of imperfect hexagonal ice crystals. *Journal of Quantitative Spectroscopy and Radiative Transfer*; **60** (1998), 301 ; DOI: 10.1016/S0022-4073(98)00007-7.
- [RD72] A Marshak, A. Davis, W. Wiscombe *et al.*; The Verisimilitude of the Independent Pixel Approximation Used in Cloud Remote Sensing. *Remote Sens. Environ.*; **52** (1995), 71.
- [RD73] U.S. Standard Atmosphere, 1976.
source: National Oceanic and Atmospheric Administration; **ref:** NOAA-S/T76-1562.
- [RD74] Observation Techniques and Mission Concepts for Atmospheric Chemistry (CAMELOT).
source: European Space Agency; **ref:** 20533/07NL/HE.
- [RD75] R.B.A. Koelemeijer, P. Stammes, J.W. Hovenier *et al.*; A fast method for retrieval of cloud parameters using the Oxygen A-band measurements from the Global Ozone Monitoring Experiment. *J. Geophys. Res.*; **106** (2001), doi:10.1029/2000JD900657.
- [RD76] B van Diedenhoven, O.P. Hasekamp and J. Landgraf; Retrieval of cloud parameters from satellite-based reflectance measurements in the ultraviolet and the oxygen A-band. *J. Geophys. Res.*; **112** (2007), D15208, doi:10.1029/2006JD008155.
- [RD77] M. Krol, S. Houweling, B. Bregman *et al.*; The two-way nested global chemistry-transport zoom model TM5: algorithm and applications. *Atmos. Chem. Phys.*; **5(2)** (2005), 417.
- [RD78] J. F. Meirink, P. Bergamaschi and M. C. Krol; Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: Method and comparison with synthesis inversion. *Atmos. Chem. Phys.*; **8** (2008), 6341.
- [RD79] P. Bergamaschi, M. Krol, F. Dentener *et al.*; Inverse modelling of national and european CH₄ emissions using the atmospheric zoom model TM5. *Atmos. Chem. Phys.*; **5** (2005), 2431.
- [RD80] P. Bergamaschi, C. Frankenberg, J.-F. Meirink *et al.*; Inverse modeling of global and regional CH₄ emissions using SCIAMACHY satellite retrievals. *J. Geophys. Res.*; **114** (2009), doi:10.1029/2009JD012287.
- [RD81] S. Houweling, M. Krol, P. Bergamaschi *et al.*; A multi-year methane inversion using SCIAMACHY, accounting for systematic errors using TCCON measurements. *Atmos. Chem. Phys.*; **14** (2014), 3991; 10.5194/acp-14-3991-2014.
- [RD82] A. Butz, O. P. Hasekamp, C. Frankenberg *et al.*; Retrievals of atmospheric CO₂ from simulated space-borne measurements of backscattered near-infrared sunlight: accounting for aerosol effects. *Appl. Opt.*; **48** (2009), 3322; 10.1364/AO.48.003322.
- [RD83] A. Butz, O. P. Hasekamp, C. Frankenberg *et al.*; CH₄ retrievals from space-based solar backscatter measurements: Performance evaluation against simulated aerosol and cirrus loaded scenes. *J. Geophys. Res.*; **115** (2010), D24302; 10.1029/2010JD014514.
- [RD84] A. Butz, S. Guerlet, O. Hasekamp *et al.*; Toward accurate CO₂ and CH₄ observations from GOSAT. *Geophys. Res. Lett.*; **38** (2011), doi:10.1029/2011GL047888.
- [RD85] D. Schepers, S. Guerlet, A. Butz *et al.*; Methane retrievals from Greenhouse Gases Observing Satellite (GOSAT) shortwave infrared measurements: Performance comparison of proxy and physics retrieval algorithms. *J. Geophys. Res.*; **117** (2012), D10307; 10.1029/2012JD017549.
- [RD86] Comprehensive Error Characterisation Report: SRON full-physics retrieval algorithm for XCH₄.
source: ESA Climate Change Initiative (CCI); **ref:** CECRv1–SRFP-XCH₄.
- [RD87] R. Locatelli, P. Bousquet, M. Saunois *et al.*; Sensitivity of the recent methane budget to LMDz sub-grid-scale physical parameterizations. *Atmos. Chem. Phys.*; **15** (2015), 9765; 10.5194/acp-15-9765-2015.
- [RD88] D. Salstein, Ponte R.M. and K. Cady-Pereira; Uncertainties in atmospheric surface pressure fields from global analyses. *J. Geophys. Res.*; **113** (2007) (D14).
- [RD89] T.G. et al. Farr; The Shuttle Radar Topography Mission. *Reviews of Geophysics*; **45** (2007), doi:10.1029/2005RG000183.

 	<p style="text-align: center;">S5L2PP Carbon Monoxide ATBD</p>	<p>Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 13/74</p>
--	--	---

- [RD90] D. Wunch, G. C. Toon, P. O. Wennberg *et al.*; Calibration of the Total Carbon Column Observing Network using aircraft profile data. *Atmospheric Measurement Techniques*; (2010), 1351; 10.5194/amt-3-1351-2010.
- [RD91] C. P. Rinsland, E. Mahieu, R. Zander *et al.*; Free tropospheric CO, C₂H₆, and HCN above central Europe: Recent measurements from the Jungfraujoch station including the detection of elevated columns during 1998. *J. Geophys. Res.*; **105** (2000), 24235; 10.1029/2000JD900371.
- [RD92] R. Sussmann and T. Borsdorff; Technical Note: Interference errors in infrared remote sounding of the atmosphere. *Atmospheric Chemistry & Physics*; **7** (2007), 3537.
- [RD93] A. Karon, C. Sweeney, P. Tans *et al.*; AirCore: An Innovative Atmospheric Sampling System. *J. Atm. Ocean Tech.*; **27** (2010), 1839; doi:10.1175/2010JTECHA1448.1.
- [RD94] SURFsara GRID documentation.
url: URL <http://doc.grid.surfsara.nl/en/latest/>.

 	<p style="text-align: center;">S5L2PP Carbon Monoxide ATBD</p>	<p>Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 14/74</p>
--	--	---

3 Terms, Definitions and Abbreviated Terms

Terms, definitions and abbreviated terms that are used in development program for the Sentinel 5 data processors are described in [RD2]. Terms, definitions and abbreviated terms that are specific for this document can be found below.

3.1 Terms and Definitions

There are no document specific terms and definitions.

 	<p style="text-align: center;">S5L2PP Carbon Monoxide ATBD</p>	<p>Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 15/74</p>
--	--	---

3.2 Acronyms and Abbreviations

ADEOS	Advanced Earth Observing System
AIRS	Atmospheric Infrared Sounder
AOT	Aerosol Optical Thickness
AQUA	A NASA Earth Science satellite mission focussing on the Earth's water cycle
CTM	Chemical Transport Model
CAMS	Copernicus Atmosphere Monitoring Service
DFS	Degree of Freedom for Signal
S-LINTRAN	Scalar linearised Radiative Transfer Program for a Multi-Layered Plane-Parallel Medium
ECMWF	European Centre for Medium-Range Weather Forecasts
ERI	European Research Institute
ESRL	Earth System Research Laboratory
FTS	Fourier Transform Spectrometer
FTIR	Fourier Transform Infrared
FRESCO	Fast Retrieval Scheme for Clouds from the Oxygen A band
FWHM	Full Width Half Maximum
GCM	General Circulation Model
GMES	Global Monitoring for Environment and Security
GMTED2010	Global Multi-resolution Terrain Elevation Data 2010
GNIP	Global Network for Isotopes in Precipitation
GOSAT	Greenhouse gases Observing Satellite
IAGOS	In-service Aircraft for a Global Observing System
IASI	Infrared Atmospheric Sounding Interferometer
IMAP	Iterative Maximum A Posteriori
IMG	Interferometric Monitor for Greenhouse gases
IMLM	Iterative Maximum Likelihood Method
IRWG	Infrared Working Group
ISRF	Instrument Spectral Response Function
L1	Level-1
L2	Level-2
LER	Lambert-equivalent Reflectivity
LOS	Line of Sight
MACC	Monitoring Atmospheric Composition and Climate
MAPS	Measurement of Air Pollution from Satellites
METimage	Multispectral imaging radiometer for meteorological applications
MODIS	Moderate Resolution Imaging Spectroradiometer
MOPITT	Measurements of Pollution in the Troposphere
MOZAIC	Measurement of Ozone and Water Vapour on Airbus in-service Aircraft
NDACC	Network for the Detection of Atmospheric Composition Change
NRT	Near Real Time
PIFM	Practical Improved Flux Method
RemoTeC	Remote Sensing of Greenhouse Gases for Carbon Cycle Modelling
RMS	Root Mean Square

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 16/74
--	--------------------------------	--

S5	Sentinel-5
S5P	Sentinel-5 Precursor
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Chartography
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
VZA	Viewing Zenith Angle
WFM-DOAS	Weighting Function Modified-Differential Optical Absorption Spectroscopy

 	<p style="text-align: center;">S5L2PP Carbon Monoxide ATBD</p>	<p>Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 17/74</p>
--	--	---

4 Introduction to the carbon monoxide algorithm SICOR

Carbon monoxide (CO) is an atmospheric trace gas with a column mixing ratio typically in the range 50-300 ppb. It is one of the highest priority chemical species required by the Monitoring Atmospheric Composition and Climate (MACC) project of the Global Monitoring for Environment and Security (GMES) program [RD3]. Measurements of its global abundance improve our understanding of tropospheric chemistry and atmospheric long range transport [RD4, RD5, RD6, RD7] and in certain urban areas, it is a major atmospheric pollutant. 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 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 [RD8], and with the IMG (Interferometric Monitor for Greenhouse gases) instrument onboard ADEOS (Advanced Earth Observing System) in 1996 and 1997 [RD9]. Since 2000, long-term global data sets of CO are provided by the MOPITT (Measurements of Pollution in the Troposphere) instrument (e.g. [RD10]), which measures 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 that measure the infrared brightness of the Earth's surface and the atmosphere: (1) AIRS (Atmospheric Infrared Sounder [RD11]) launched in 2002 onboard the Aqua satellite, (2) TES (Tropospheric Emission Spectrometer [RD12] launched in 2004 onboard the Aura satellite and (3) IASI (Infrared Atmospheric Sounding Interferometer [RD13]) 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 [RD14].

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 absorbs light 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 clear sky 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, [RD15]) in the year 2002 on ESA's Envisat satellite, a continuous time series of global CO SWIR measurements is available. Moreover, first results were reported using MOPITT measurements in the SWIR [RD16] and a combination of SWIR and the 4.7 μm CO absorption band [RD17]. In the year 2017, the SWIR measurements of CO are continued by the Sentinel 5 Precursor mission (S5P [RD18]) with the TROPOMI instrument as single payload and by the Sentinel 5 mission [RD19, RD20] to be launched in 2022 as a successor of S5P. These missions will provide the opportunity to extend the unique long-term global data set of CO using the same type of measurement. Moreover, a major objective of these missions is to provide observations for air quality model processes in support of operational services, including air quality forecasting and protocol monitoring. This requires the improved radiometric performances and higher spatial sampling of the S5P and S5 instruments and the development of fast and accurate retrieval algorithms.

4.1 Heritage

Several algorithms are used to retrieve CO column from SCIAMACHY SWIR measurements, including the Weighting Function Modified-Differential Optical Absorption Spectroscopy (WFM-DOAS) approach ([RD21] and references therein) the Iterative Maximum A Posteriori (IMAP) approach [RD22], and the Iterative Maximum Likelihood Method (IMLM) approach ([RD23], 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. Buchwitz et al. [RD24] and Gloudemans et al. [RD23] use a priori methane information to

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002
		Version : 3.1 Page
		Date : 17 May 2019 18/74

characterize the light path through the atmosphere. Both approaches are successfully used to detect enhanced CO concentrations at strongly polluted areas, the seasonal variability of global atmospheric CO and its long range transport. The retrievals have been compared with chemical transport models (CTM, e.g. [RD25, RD23]), with MOPITT retrievals (e.g. [RD21, RD26]) and ground-based FTIR (Fourier Transform Infrared) measurements [RD27, RD28]. Since CO is a weak absorber and SCIAMACHY exhibits a low signal-to-noise ratio in the $2.3\ \mu\text{m}$ region, the retrieval of single-sounding CO column density is subject to large retrieval noise, typically on the order of 10-100 % [RD29]. Hence, to compare SCIAMACHY retrievals with ground-based measurements, the averaging of multiple SCIAMACHY measurements is needed. De Laat et al. [RD28] demonstrated that, at the northern latitudes and for a typical sampling area of $8^\circ \times 8^\circ$, 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./cm². Recently, Borsdorff et al. [RD30, RD31] presented a SCIAMACHY full-mission CO data product using the SICOR-S5P algorithm. This algorithm is designed to infer CO total column densities from S5P SWIR clear sky observations over land and cloudy observations over both land and oceans. The 10 year data set is validated extensively using TCCON and NDACC ground based measurements of the CO total column and MOZAIC/IAGOS aircraft measurements at airport sites, showing good agreement between the SCIAMACHY soundings and the validation measurements. The data product comprises the total column estimate, its precision and the so-called column averaging kernel, which provides essential information to validate SWIR CO retrievals with other independent retrievals or CTM simulations.

Using SCIAMACHY and S5P algorithm heritage, the SICOR-S5 algorithm infers the total column density of CO from S5 SWIR clear sky and cloudy measurements. The approach is a so-called physics-based retrieval, which employs online radiative transfer to simulate S5 measurements. 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 H₂O, CO and CH₄, with the optical depth of CO generally much smaller than those of H₂O and CH₄.

Clouds and aerosols affect the sensitivity of the measurement to CO in several manners: due to multiple light scattering by cloud droplets, the path length of the observed light is enhanced. 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 significantly enhance the light path. The 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 the two-stream radiative transfer solver 2S-LINTRAN, 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 linearization of the model with respect to scattering and absorption properties in the model atmosphere. In summary, the presented SICOR-S5 algorithm builds on strong SCIAMACHY and S5P algorithm heritage and enables the exploitation of S5 SWIR measurement for operational data processing.

4.2 Requirements

To improve our present knowledge on CO on a global scale, the following requirements are formulated as baseline for the Sentinel-5 prototype algorithm development [AD4]:

S5-L2-PRO-300 The uncertainty in the total CO column density shall be smaller than $4 \cdot 10^{17}$ molec./cm².

S5-L2-PRO-310 The bias in the total CO column density shall be smaller than 5 % for total CO column densities larger than $1.6 \cdot 10^{18}$ molec./cm².

S5-L2-PRO-320 The stability in the CO total column density shall be smaller than $8 \cdot 10^{16}$ molec./cm² per decade.

For the SICOR-S5 algorithm implementation, the CO retrieval bias, i.e. the systematic error of the retrieved CO column, and the CO precision can be quantified for an ensemble of simulated measurements. To deduce requirements on these diagnostics, we assume that the product uncertainty is derived from the precision and the retrieval bias using statistical error propagation. Considering an atmospheric background scene with a typical CO column density of $1.6 \cdot 10^{18}$ molec./cm² and for a polluted scenario with a CO column density of about $4.0 \cdot 10^{18}$ molec./cm², the uncertainty requirement corresponds to a precision requirement of $3.9 \cdot 10^{17}$ and $3.5 \cdot 10^{17}$ molec./cm². Thus, the

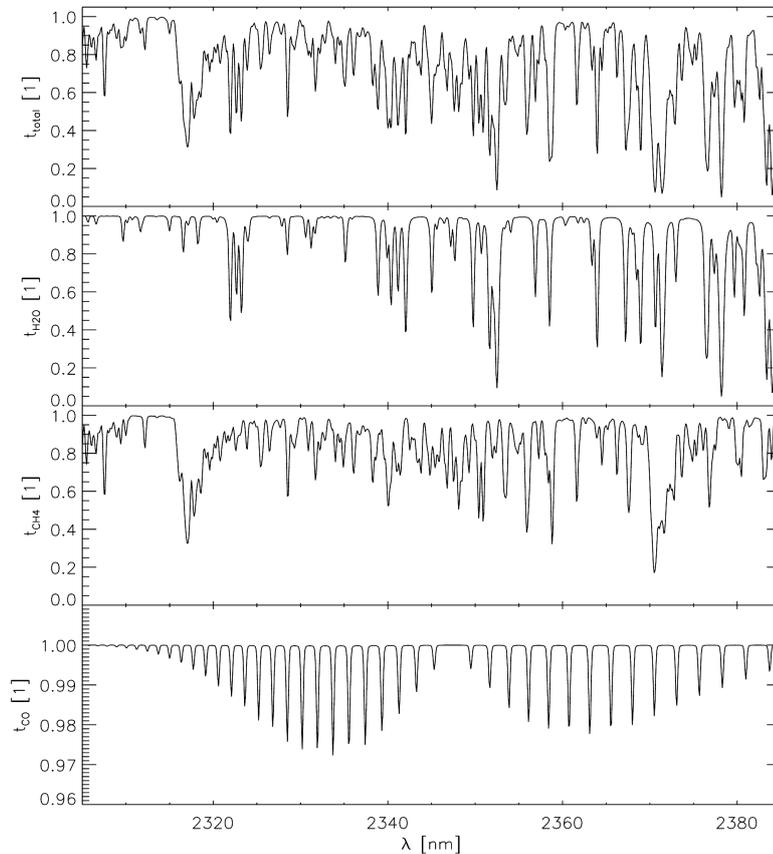


Figure 1: Spectral transmittance along the light path of the solar beam reflected by the Earth surface into the instrument viewing direction for the 2.3 μm band. Simulations are performed for viewing zenith angle (VZA) of 0° , and a solar zenith angle (SZA) of 30° , and by assuming a US standard atmospheric profile. From top to bottom, the figure shows the total transmittance, the individual transmittances due to H_2O , CH_4 , and CO , respectively. Note the different y-axis scale for CO transmittance.

required precision and uncertainty are very similar and so to simplify matters, we interpret requirement S5-L2-PRO-300 as a precision requirement. Moreover, the stability of the CO product is determined by the stability of the instrument in space and cannot be considered as a design driver for the algorithm definition. Therefore, we omit the stability requirement as a diagnostic tool to investigate algorithm performance. So we remain with the following requirements for the SICOR algorithm baseline:

1. The precision of the total CO column density shall be smaller than $4 \cdot 10^{17}$ molec./ cm^2 .
2. The bias in the total CO column density shall be smaller than 5 % for total CO column densities larger than $1.6 \cdot 10^{18}$ molec./ cm^2 .

Based on the S5P algorithm heritage, we conclude that the SICOR algorithm is compliant with these requirements. No limitations on solar and viewing geometry are foreseen.

 	<p>S5L2PP Carbon Monoxide ATBD</p>	<p>Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 20/74</p>
--	--	---

5 Instrument Overview

A description of the Sentinel 5 instrument and performance can be found in [RD32].

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Date : 17 May 2019	Page 21/74
--	---------------------------------------	---	---------------

Table 1: Spectral fit windows of SICOR based on the S5 band definition [AD2].

fit window	spectral range	spectral resolution	sampling ratio	S5 band
Non-scattering	2315–2324 nm	0.25 nm	3	SWIR-3 (band 5)
Physics-based	2324–2328 nm	0.25 nm	3	SWIR-3 (band 5)

6 Detailed Algorithm Description

The SICOR algorithm is designed to retrieve CO total column density from SWIR-3 radiance measurements around 2.3 μm , the so-called SWIR-3 band¹. It is a further development of the retrieval method by Vidot et al. [RD33] and Landgraf et al. [RD34, RD35] and is well suited for the operational level 2 data processing of the Sentinel 5 mission. The overall structure of the algorithm is summarized in Fig. 2. SICOR requires inputs of different categories:

- The primary input of the algorithm is the Sentinel 5 level 1B (S5-L1B) product comprising SWIR-3 measured Earth radiance and solar irradiance spectra including noise estimate, solar and viewing geometry, and information on geo-location and the ground scene elevation.
- To model atmospheric absorption, lookup tables of the atmospheric absorbers CO, CH₄, H₂O, and HDO are required.
- The measurement simulations need the instrument spectral response function (ISRF) to degrade the spectral simulation to the sensor spectral resolution. The ISRF is part of the calibration key data (CKD).
- To initialize SICOR, chemical transport model (CTM) estimates of the atmospheric profiles of CO, H₂O, CH₄ temperature, and pressure profiles are needed including the geo-potential height.
- METimage data on cloud height and cloud fraction are used as first guess to initiate the inversion and to define quality flags of the CO data product.
- SICOR requires several configuration parameters, which are provided by the software settings.

After a first screening of the L1B data based on e.g. the spectral signal-to-noise (SNR) performance, two main processing steps are performed. First, observations with high and optically thick clouds are discarded. These measurements provide only little information on tropospheric CO and so would be of minor benefit. The filter relies on the vertically integrated amount of methane which is retrieved from the dedicated SWIR-3 fit window 2315–2324 nm using a non-scattering radiative transfer model. The difference between the retrieved column and a priori CH₄ knowledge indicates the lightpath shortening or enhancement due to atmospheric scattering by clouds and aerosols. If the difference ΔCH_4 exceeds a certain threshold, observations are strongly contaminated by clouds and are rejected. In a second step, the SICOR physics-based retrieval approach is used to infer CO columns from the adjacent spectral window, 2324–2338 nm. Here, the methane absorption features provide the information on atmospheric scattering by clouds and aerosols, which is modeled in the retrieval by a scattering layer of triangular height distribution with fixed geometrical thickness. Its optical depth and height are retrieved together with the atmospheric CO, H₂O and HDO abundances, surface albedo and a spectral calibration of the radiance spectrum. Obviously this step of the retrieval relies on accurate a priori knowledge of CH₄, which has to be provided within an accuracy of $\pm 3\%$ prior to the retrieval. The different retrieval windows are summarized in Tab. 1. Finally, the algorithm product consists of the retrieved CO column, its column averaging kernel and a random error estimate. A detailed description of input and output parameters is given in Sect. 9.2.

As shown in Fig. 1, CO is the weakest of the absorbers in the SWIR-3 spectral window. Therefore, the signal-to-noise ratio of an individual CO retrieval is rather low. To mitigate this weakness, as many retrievals as possible have to be performed so that noise contributions can be averaged out. This means that the data filters must be quite loose, also accepting soundings over partially cloud-covered scenes or over stratiform clouds. These optical depths are much higher than those of clear skies with just aerosols or cirrus. Therefore, our algorithm must be suitable to handle scenes with relatively high optical thickness. This forces the use of a small spectral window, so that the assumed microphysical

¹ The band nomenclature is adapted from [AD2] with the 1590-1675 nm SWIR-1 band and 2305-2385 nm SWIR-3 band. The SWIR-2 band covers the 2.0 μm range but was discarded later in the project.

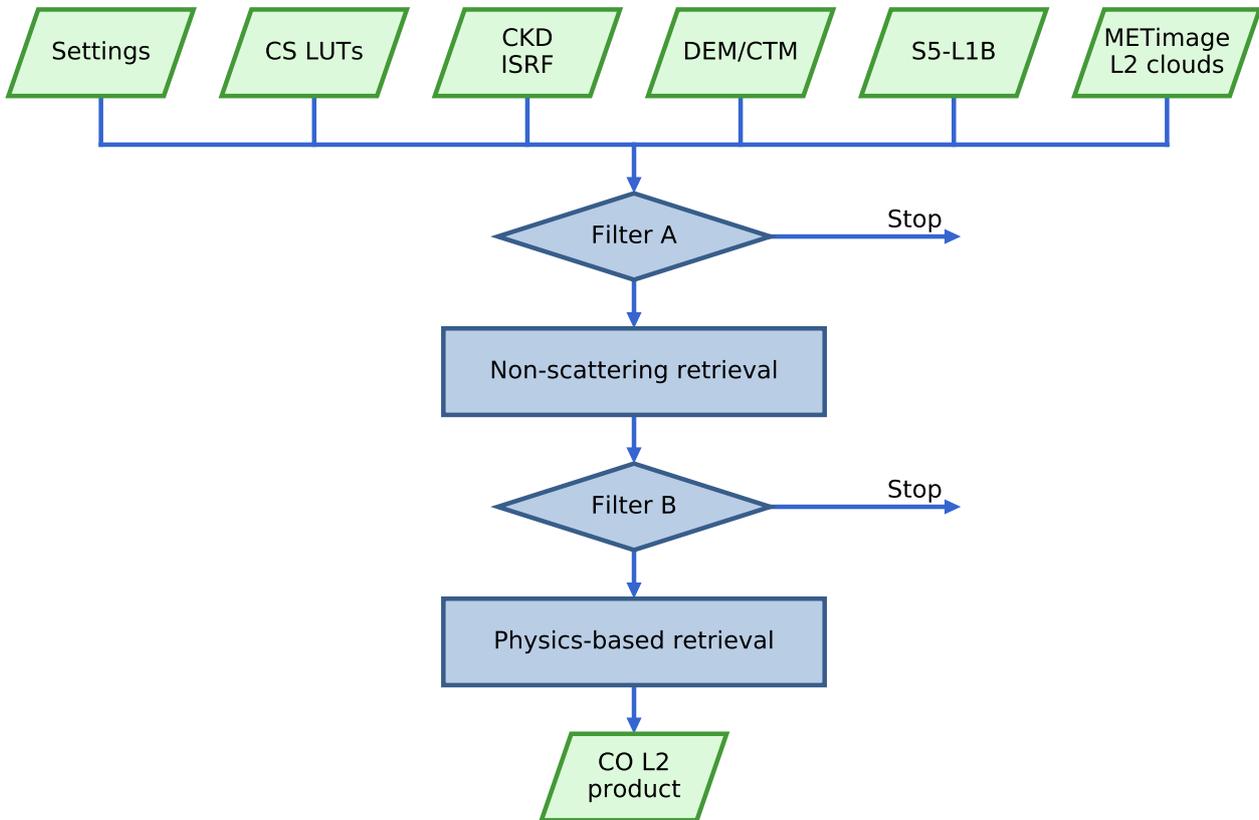


Figure 2: Overall processing flowchart of the CO retrieval algorithm.

properties of the cloud droplets are less relevant. Furthermore, several speed optimizations are required to perform all the CO retrievals in the required time.

6.1 Input

All input data are shown in the green boxes in the top of Fig. 2. We distinguish between static and dynamic input, where the static input (e.g. retrieval input settings, cross section lookup tables and calibration key data) is required to initialize the CO processor and needs to be read in only once when the processing starts. The calculation settings will be provided to the algorithm as a character string that can be interpreted by the algorithm. For the contents of these settings, we refer to B. The dynamic input consists e.g. of SWIR-3 level 1b radiance data. Moreover, the input comprises S5 METimage cloud data [RD36] and the atmospheric input generated from (static) DEM (GMTED2010, [RD37]) and dynamic CTM (ECMWF, CAMS) data. All dynamic input is assumed to be co-located for each S5 ground pixel by the processor frame.

6.1.1 Absorption cross sections

The absorption cross-sections of CH₄, CO, H₂O, and HDO are pre-calculated from the latest spectroscopic databases [RD38, RD39] assuming Voigt line shapes and stored in a lookup-table as a function of pressure, temperature, and wavenumber. For water vapor, the updated spectroscopic line list of Scheepmaker et al., 2013, [RD40] is employed. As H₂O concentrations are relatively high, self-broadening effects are taken into account by using an effective pressure as outlined in [RD41]. This is done the same as for GOSAT, OCO-2 and S5P.

Because calculation time must be minimized, the spectral grid on which the model calculations are performed must be rather coarse, possibly too coarse to properly resolve all absorption lines, especially at low pressures. When sampling

	S5L2PP	Reference : SRON-ESA-S5L2PP-ATBD-002
	Carbon Monoxide ATBD	Version : 3.1 Page
		Date : 17 May 2019 23/74

Table 2: The S5 L2 CLA data product used by SICOR. For more detail see Tab. 6

METimage product	symbol	averaging approach
METimage liquid cloud fraction	f_{liq}	SRF
METimage liquid cloud top pressure	p_{liq}	SRF
METimage liquid cloud optical depth	r_{liq}	SRF

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 are defined for each spectral sampling point of the (coarse) internal model grid by an average with a triangular weight function around the sample wavenumber k_i ,

$$\sigma_i = \sqrt[m]{\frac{\int T_i(k)\sigma^m(k)dk}{\int T_i(k)dk}}, \quad (1)$$

where the triangular weight function $T_i(k)$ is defined as

$$T_i(k) = \begin{cases} k - k_{i-1} & \text{for } k_{i-1} < k < k_i \\ k_{i+1} - k & \text{for } k_i < k < k_{i+1} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

assuming an equidistant wavenumber grid (k_i) in the model. This approach introduces a tuning parameter m . If $m = 1$, Eq. (1) represents the arithmetic mean, which overestimates the absorption because of the non-linear relationship between cross section and modeled radiance. For the limit $m \rightarrow 0$ it describes the geometric mean, which appears to be too low. As baseline, m is taken to be 0.85, which is obtained from a test with a standard cloudless atmosphere.

6.1.2 METimage

SICOR requires the S5 METimage cloud products summarized in Tab. 2. These only include data for liquid clouds averaged over the spatial response function (SRF) of the instrument. The rest of the S5 L2 CLA product is not used for the CO algorithm. Here, the cloud top height provides the first guess cloud information to initialize the retrieval and in combination with cloud fraction is used to derive the confidence level of the data product (TBC).

6.2 Standard retrieval setup

For both retrieval approaches, the non-scattering and the physics-based retrievals, a forward model \mathbf{F} is needed that describes the measurement as a function of the state of the atmosphere, namely

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \mathbf{e}_y. \quad (3)$$

Here, vector \mathbf{y} has the spectral measurements as its components, state vector \mathbf{x} represents the parameters to be retrieved, \mathbf{b} describes parameters other than the state vector that influences the measurement, and \mathbf{e}_y is the error of the measurement.

The retrieval process retrieves a state $\hat{\mathbf{x}}$, by minimizing the following cost function:

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \|\mathbf{S}_y^{-1/2}(\mathbf{F}(\mathbf{x}) - \mathbf{y})\|^2 \quad (4)$$

where \mathbf{S}_y is the error covariance matrix of the measurement, for which we assume uncorrelated noise estimates that are provided as part of the L1B input.

	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 24/74
---	--------------------------------	--

For the retrievals performed by the SICOR algorithm, the forward model is non-linear in the state vector \mathbf{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 \mathbf{x}_0 ,

$$\mathbf{F}(\mathbf{x}, \mathbf{b}) = \mathbf{F}(\mathbf{x}_0, \mathbf{b}) + \mathbf{K}(\mathbf{x}_0, \mathbf{b})(\mathbf{x} - \mathbf{x}_0) \quad (5)$$

where

$$\mathbf{K} = \frac{\partial \mathbf{F}}{\partial \mathbf{x}}. \quad (6)$$

The solution of the minimization problem with the linearized forward model is

$$\hat{\mathbf{x}} = \mathbf{G}\tilde{\mathbf{y}} \quad (7)$$

with the gain matrix

$$\mathbf{G} = (\mathbf{K}^T \mathbf{S}_y^{-1} \mathbf{K})^{-1} \mathbf{K}^T \mathbf{S}_y^{-1} \quad (8)$$

and $\tilde{\mathbf{y}} = \mathbf{y} - \mathbf{F}(\mathbf{x}_0) + \mathbf{K}\mathbf{x}_0$.

The calculation of \mathbf{G} involves a matrix inversion. Based on SCIAMACHY and Sentinel-5 Precursor heritage and numerical experience, we perform this matrix inversion using Cholevsky decomposition, which is roughly twice as efficient as the LU decomposition [RD42] with good numerical stability. Cholevsky decomposition employs an optimization that only works for a symmetric matrix, which we have in Eq. (8).

An estimate of the instrument noise propagation on the retrieved state can be calculated from the measurement error covariance matrix with

$$\mathbf{S}_x = \mathbf{G}\mathbf{S}_y\mathbf{G}^T, \quad (9)$$

resulting in an error covariance matrix for the retrieved state.

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. We consider the difference in χ^2 between two consecutive iteration steps. This χ^2 is a measure of the goodness of the fit and is officially defined as

$$\chi_i^2 = \frac{(\mathbf{y} - \mathbf{F}(\mathbf{x}_i))^T \mathbf{S}_y^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x}_i))}{N_y - \text{DFS}} \quad (10)$$

with N_y is the number of elements in the measurement vector and DFS is the degrees of freedom for signal, usually the number of parameters in the state vector. However, in more advanced retrievals, DFS can be lower and even change per iteration. For the convergence criterion, we leave out DFS from Eq. (10) because a changing DFS between successive iterations could theoretically cause numeric problems in the convergence criterion. As $\text{DFS} \ll N_y$, the error in χ^2 when ignoring DFS is small and can easily be mitigated by choosing the threshold $\Delta\chi^2$ for convergence accordingly. At convergence, the correct χ^2 from the full Eq. (10) will be calculated to diagnose how good the fit is.

The threshold value of $\Delta\chi^2$ for convergence can only be determined in a reliable manner using real measurements during the mission consolidation phase, where 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 retrieval is considered a failure.

The likelihood of convergence is directly linked to the validity of the linearized model around \mathbf{x}_i at state $\hat{\mathbf{x}}$. This depends on the non-linearity of the model and on how much $\hat{\mathbf{x}}$ deviates from \mathbf{x}_i . To increase the chance of convergence, we tackle the large differences between $\hat{\mathbf{x}}$ and \mathbf{x}_i by introducing a step size reduction factor Λ (also explained by Butz et al. [RD43]). Reducing the step size increases the validity of the linearized model on which the inversion is based. The state vector retrieved during iteration i is the weighted mean of the state vector retrieved during iteration $i - 1$ and the calculated state during iteration i , where the reduction factor Λ determines the weights:

$$\mathbf{x}_i = \frac{1}{1 + \Lambda} \hat{\mathbf{x}}_i + \frac{\Lambda}{1 + \Lambda} \mathbf{x}_{i-1}. \quad (11)$$

	S5L2PP	Reference : SRON-ESA-S5L2PP-ATBD-002
	Carbon Monoxide ATBD	Version : 3.1 Page
		Date : 17 May 2019 25/74

Each inversion step is evaluated by a forward simulation. A step size is accepted if either the χ^2 decreases with respect to the previous iteration, or it increases slightly. When a step is rejected, Δ is increased and the new state is calculated with Eq. (11). When a step is accepted, Δ is reduced and the next iteration is performed.

The underlying idea of this approach is that the step is most strongly reduced at the first iterations, where \mathbf{x}_0 is further away from the minimum of the cost function and the proposed step size of a linear inversion is generally large.

6.3 Initial data filtering A

Overall, we assume that only radiance and irradiance data are provided, which passes at the calling framework level data quality control. Appendix C gives an overview on the use of L1b flagging information for data control prior to the RemoTeC retrieval. Data with any warning are processed in the current baseline setting but later during the S5 commission phase one may consider to use warning flags of the L1B data product to potentially filter the input data. Moreover, the SICOR algorithm relies on two consecutive data filtering steps, indicated by A and B in Fig. 2, which requires dynamic algorithm input such as L1B data including quality flags and intermediate SICOR quantities. For the first data filtering step, we evaluate the quality of the S5 radiance measurement in the SWIR-3 spectral range. In first instances, this check uses quality flags of the S5-L1B data product. Subsequently, we filter out scenes which are too dark to provide a useful spectrum, which typically occurs for clear sky observations over water. The quality check is performed using the maximum signal-to-noise ratio (SNR), which is typically below 50 for the scenes to be excluded. Additionally, forward model errors increases significantly for low sun and so corresponding S5 observations are not considered for further processing. Hence, the initial screening A of the S5 observations comprises the following three filter steps:

A1 Filter out too dark scenes rejecting data with a too low SNR in the SWIR-3 spectral range,

$$\max[I_{SW3}] > T_{A1} \quad (12)$$

A2 For low sun, forward model errors increases substantially. Therefore, we propose to consider only measurements with sufficiently small solar zenith angle (SZA)

$$SZA < T_{A2} \quad (13)$$

The threshold values T_{A1} and T_{A2} will be determined first during the commissioning phase using real data.

We are reluctant to filter data based on the external METimage cloud product collocated on the S5 data (the so-called S5 L2 CLA product) [RD36], to reduce the dependency of the CO product on secondary input. Instead, we based our cloud filtering (data filtering B) on the retrieval of the methane total column assuming a non-scattering atmosphere. The numerical effort of this retrieval is minor and is described in more detail in the next section.

6.4 Non-scattering methane retrieval

Assuming accurate a priori methane knowledge, the presence of optically thick clouds can be filtered using methane retrieved from the SWIR-3 spectral range for a non-scattering model atmosphere. To illustrate the idea of such a cloud filter, Fig. 3 shows the amount CH_4 above a cloud with a cloud top height z_{clid} . Above 2, 5 and 10 km about 80, 50 and 20 % of the total amount of methane is located. For numerical implementation of this filter, we employ a non-scattering retrieval from SWIR-3 measurements in the spectral range 2315–2324 nm with strong methane absorption (see Fig. 1).

6.4.1 Non-scattering forward model

The algorithm relies on a radiative transmission model ignoring any atmospheric scattering. Here, sunlight is reflected at the Earth surface into the satellite line of sight and is attenuated by atmospheric absorption along its path. Using this approximation, the simulated radiance at the top-of-atmosphere (TOA) $I^{\text{TOA}}(\lambda)$ is given by:

$$I^{\text{TOA}}(\lambda) = F_0(\lambda) A_s(\lambda) \frac{\mu_0}{\pi} e^{-\tau_{\text{tot}}(\lambda)/\mu_0}, \quad (14)$$

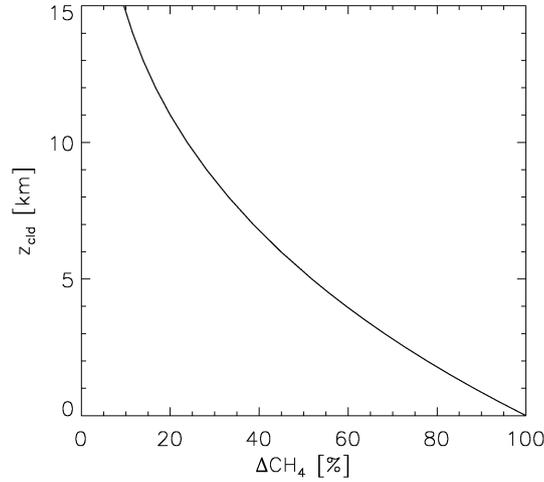


Figure 3: Relative CH₄ column above a cloud top height z_{cld} with respect to the total column amount, using the US standard model atmosphere.

where A_s is the surface albedo, $\mu_0 = \cos(\Theta_0)$ with the solar zenith angle Θ_0 . A correction for μ_0 is applied to account for the sphericity of the Earth according to Kasten and Young (1989) [RD44]. F_0 is the solar irradiance, which can be obtained from a reference solar spectrum or from the S5 irradiance measurements (see Sect. 6.7) And

$$\tilde{\mu} = \frac{\mu_0 \mu_v}{\mu_0 + \mu_v} \quad (15)$$

with $\mu_v = \cos(\Theta_v)$ and viewing zenith angle Θ_v . The total optical thickness τ_{tot} is given by

$$\tau_{\text{tot}}(\lambda) = \sum_{n,k} \sigma_{n,k}(\lambda) s_{n,k}, \quad (16)$$

where k indicates the layer index in the model atmosphere and index n represents the relevant absorbers CO, CH₄, H₂O and HDO, $s_{n,k}$ is the sub-column of absorber n in model layer k and $\sigma_{n,k}(\lambda)$ is the corresponding absorption cross section interpolated at the temperature and pressure of model layer k .

The derivative of the TOA radiance with respect to any subcolumn $s_{n,k}$ of absorber n in layer k is

$$\frac{\partial I^{\text{TOA}}}{\partial s_{n,k}} = -\frac{I^{\text{TOA}}}{\tilde{\mu}} \sigma_{n,k} \quad (17)$$

Finally, the derivative of I^{TOA} with respect to surface albedo A_s is

$$\frac{\partial I^{\text{TOA}}}{\partial A_s} = \frac{I^{\text{TOA}}}{A_s} \quad (18)$$

To account for the spectral instrument response, the TOA radiance and its derivatives have to be convolved with the instrument spectral response function (ISRF)

$$F_i = S * I^{\text{TOA}} = \int S(\lambda_i + \delta_I, \lambda) I^{\text{TOA}}(\lambda) d\lambda, \quad (19)$$

and corresponding expressions for any derivative of I^{TOA} with respect to atmospheric parameters. Here, S is the ISRF, λ_i is the wavelength assigned to the spectral pixel i of the measurement and δ_I is a spectral shift. We calculate the derivative of the forward simulation with respect to a spectral shift from Eq. (19) by

$$\frac{\partial F_i}{\partial \delta_I} = S' * I^{\text{TOA}} \quad (20)$$

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Date : 17 May 2019	Page 27/74

where $S' = \frac{\partial S}{\partial \delta_I}$ is the spectral derivative of the ISRF.

6.4.2 Non-scattering inversion

In the non-scattering retrieval, we assume that the relative profile

$$s_{n,k}^{\text{rel}} = \frac{s_{n,k}}{C_n} \quad (21)$$

of absorber n is known a priori and does not change during the retrieval, where

$$C_n = \sum_k s_{n,k} \quad (22)$$

is the column density of this absorber. So,

$$\frac{\partial I^{\text{TOA}}}{\partial C_n} = -\frac{I^{\text{TOA}}}{\bar{\mu}} \sum_k \sigma_{n,k} s_{n,k}^{\text{rel}} \quad (23)$$

The surface albedo is assumed to have a linear spectral dependence:

$$A_s(\lambda) = A_0 + A_1 (\lambda - \lambda_0) \quad (24)$$

with an arbitrary reference wavelength λ_0 . This λ_0 is chosen inside the retrieval window with the sole purpose of making parameter A_0 a representative value for the albedo in that window.

The derivatives of the TOA radiance with respect to A_0 and A_1 are obtained with straight forward chain rules.

$$\frac{\partial I^{\text{TOA}}}{\partial A_0} = \frac{I^{\text{TOA}}}{A_s} \quad (25)$$

$$\frac{\partial I^{\text{TOA}}}{\partial A_1} = \frac{I^{\text{TOA}}}{A_s} (\lambda - \lambda_0) \quad (26)$$

With these assumptions, the state vector for this retrieval is defined as

$$\mathbf{x} = (C_{\text{CO}}, C_{\text{CH}_4}, C_{\text{H}_2\text{O}}, C_{\text{HDO}}, A_0, A_1, \delta_I) \quad (27)$$

A basic inversion is executed as described in Sect. 6.2. The output of this retrieval is the C_{CH_4} element of the retrieved state vector.

6.5 Cloud filter B

Scenes covered by optically thick clouds can be identified by comparing the retrieved CH_4 from the non-scattering retrieval with an accurate a-priori estimate. Methane can be predicted by forecast models with an uncertainty of 2–3% using state-of-the-art chemical transport models [RD35]. Therefore we propose filter B to be:

B1 Spectra pass the filter if

$$\left| \frac{C_{\text{CH}_4}^{\text{non-scat}}}{C_{\text{CH}_4}^{\text{apr}}} - 1 \right| < T_{B1} \quad (28)$$

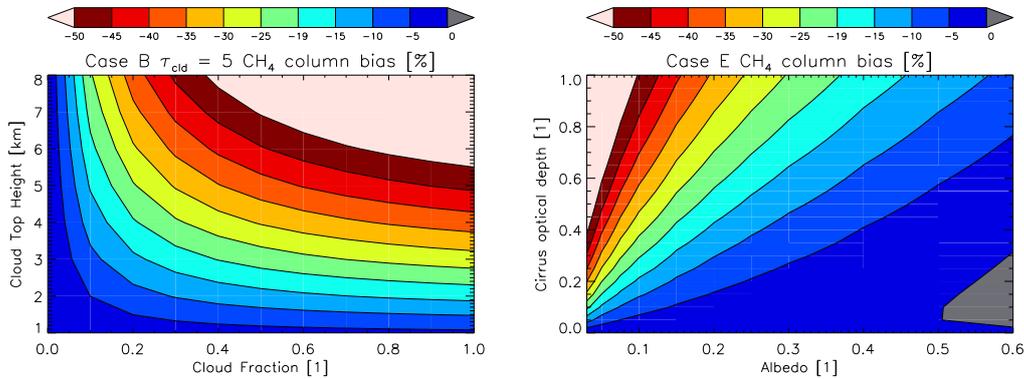


Figure 4: CH₄ error of a non-scattering retrieval from the SWIR-3 2315–2324 nm spectral window for a cloud with optical thickness of 5 as function of cloud height and cloud fraction (left panel) and for a cirrus cloud at 10 km height as function of surface albedo and cirrus optical thickness (right panel).

With a threshold T_{B1} that is significantly larger than the uncertainty of the a-priori knowledge of methane. As an initial guess for T_{B1} we use 25 %, but the value has to be evaluated during the S5 commissioning phase.

Figure 4 shows the cloud filter for a cloud with an optical depth of 5 as a function of cloud fraction and cloud top height and for a cirrus at 10 km altitude as a function of surface albedo and cirrus optical depth. For a cloud, the methane error increases with cloud coverage and cloud height. In case of a cirrus at 10 km height, the non-scattering CH₄ column can be used to identify cirrus contamination with optical depth > 0.5 for low and moderate surface reflection. Obviously for bright surfaces, the lightpath shortening due the reflection of light by optically thin cirrus can be compensated by an enhancement of the lightpath because of multiple scattering between the cirrus cloud and the surface. These atmospheric circumstances cannot be screened by our cloud filter and this cirrus effect on the atmospheric light path is treated effectively by the subsequent physics-based retrieval. 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 [RD35], this simple cloud filter is a powerful tool to screen measurement with respect to the presence of high and optically thick clouds.

Figure 5 shows the probability density function (PDF) of the non-scattering CH₄ retrieval error ΔCH_4 and its cumulative distribution (CPDF) for one year of GOSAT data (2010). As references we use collocated CH₄ columns from TM5 model simulations. The maximum of the PDF around small retrieval error indicates scenes that 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 path length enhancement. The largest light path enhancement occur on land pixels, where the surface albedo is high enough to facilitate photon trapping. Overall, the figure clearly demonstrates the ability to use the difference ΔCH_4 between the non-scattering retrieval and a priori knowledge for cloud screening of the measurements.

6.6 Physics-based retrieval

The physics-based retrieval of CO requires a forward model \mathbf{F} that describes the measurement as a function of the atmospheric state including an appropriate description of atmospheric scattering by clouds, cirrus and aerosols on the measurements between 2324 nm and 2338 nm. The fit window compromises optimal CO sensitivity, little interference with water vapor, and moderate methane absorption to retrieve scattering parameters. The window is small enough so that the assumed cloud model has only little effect on the retrieval.

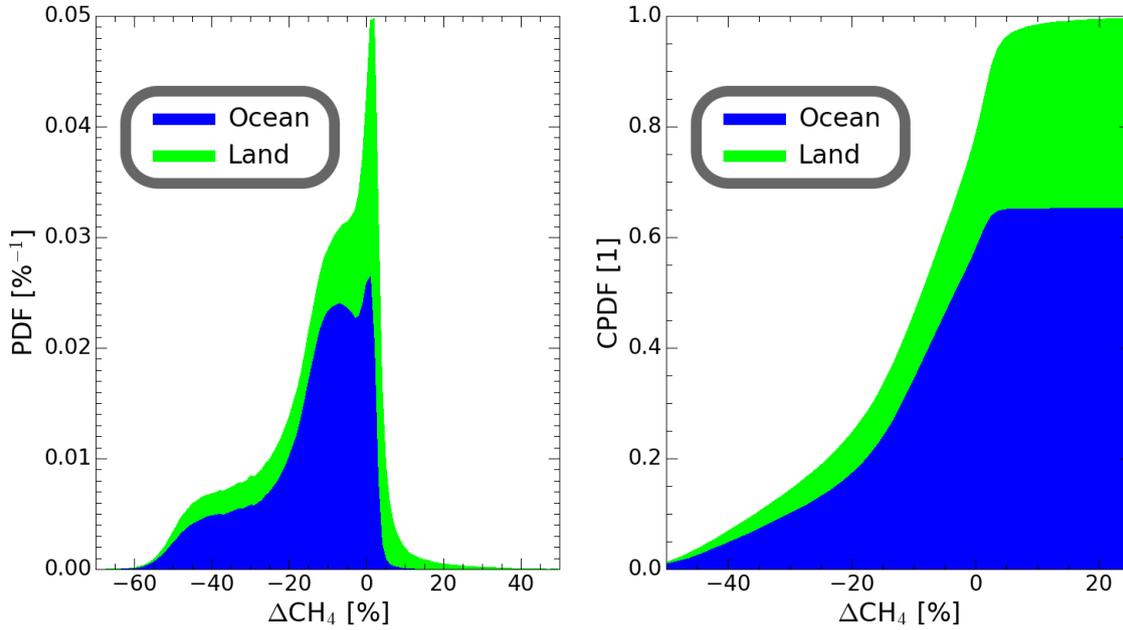


Figure 5: Probability density function (left panel) and cumulative probability density function (right panel) of the difference ΔCH_4 of one year of GOSAT non-scattering retrievals (2010) compared with the corresponding TM5 model simulations. The figure differentiates between the contribution of ocean and land pixels (blue and green areas). The dataset comprises $2.4 \cdot 10^6$ GOSAT measurements in total, under which $1.6 \cdot 10^6$ ocean pixels and $8 \cdot 10^5$ land pixels. All retrievals are performed using RemoTeC V2.1.

6.6.1 Forward model setup

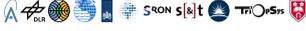
State-of-the-art radiative transfer models account for multiple scattering in multiple propagation directions (streams) including the polarization of light. The computational effort of such simulations is too high for the operational retrieval of CO within the performance requirements. So, we use a numerically efficient two-stream radiative transfer model.

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 (diffuse) radiation field is described by two, one upward and one downward, propagation directions of the radiance field. Due to the little numerical cost, these methods are commonly used to describe radiative transfer in global circulation models and weather forecast models. For the Sentinel-5 CO column retrieval, it is used to account for atmospheric scattering by clouds and aerosols that passed the cloud filter.

The error because of the approximation of atmospheric scattering is mitigated by retrieving effective scattering parameters. With these effective scattering parameters, the simplified two-stream model simulates the correct light path for the retrieval window. This is only possible for a small retrieval window, which we have for the CO retrieval. As a result, the retrieved scattering parameters themselves should be interpreted with care.

Atmospheric scattering by clouds and aerosols is represented by a triangular cloud of effective scatterers with a fixed geometric width. The microphysical properties of these effective scatterers is of minor relevance. This has been verified by performing retrievals on synthetic measurements, where the single-scattering albedo and the phase function for a highly absorption dust aerosol (type HA DUST 14 of Torres et al. [RD45]) is used in the retrievals, although the measurements were simulated with water and ice clouds. Those retrievals gave desirable results. 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 a fixed Ångström parameter:

$$\tau(\lambda) = \tau(\lambda_0) \left(\frac{\lambda}{\lambda_0} \right)^{-\alpha} . \quad (29)$$

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002
		Version : 3.1 Page
		Date : 17 May 2019 30/74

Here, λ_0 is a fixed reference wavelength and α is the Ångström parameter. The phase function for single scattering is approximated with the Henyey-Greenstein phase function, depending only on the asymmetry parameter.

Moreover, we ignore Rayleigh scattering, which contributes less than 0.15 % to the total signal [RD46] in our spectral window. With this simplification, scattering is limited to a small fraction of the atmosphere, which enables an important speed optimization. Without any loss of accuracy, adjacent non-scattering atmospheric layers can be combined for the RTM calculation. By combining all layers above the scattering layer and eventually the non-scattering layers between the scattering layer and the surface, the number of layers is reduced from 50 (baseline) to generally less than 10. The resulting derivatives with respect to optical properties of layers can easily be remapped to the original vertical grid.

The two-stream method that we use is TS-LINTRAN. This model is based on the practical improved flux method (PIFM) of Zdunkowski et al. [RD47]. In TS-LINTRAN, this model is extended by applying the forward-adjoint perturbation theory to derive analytical expressions for the derivatives of the measurement simulation with respect to total optical depth, single scattering albedo and surface albedo.

6.6.2 The two-stream radiative transfer model

For the numerical simulation, we assume a vertically inhomogeneous atmosphere described by N homogeneous layers. Each layer is characterized by its optical depth τ_n , the single-scattering albedo ω_n and the phase function P_n , with layer index $n = 1, \dots, N$.

Generally for an arbitrary layer n , the fluxes at the layer interfaces $n-1$ and n are constrained by internal boundary conditions as

$$\begin{bmatrix} S_n \\ F_n^\downarrow \\ F_{n-1}^\uparrow \end{bmatrix} = \begin{bmatrix} a_{1,n} & 0 & 0 \\ a_{2,n} & a_{4,n} & a_{5,n} \\ a_{3,n} & a_{5,n} & a_{4,n} \end{bmatrix} \begin{bmatrix} S_{n-1} \\ F_{n-1}^\downarrow \\ F_n^\uparrow \end{bmatrix} \quad (30)$$

with $n = 1, \dots, N$ and coefficients $a_{1,n}$, $a_{2,n}$, $a_{3,n}$, $a_{4,n}$ and $a_{5,n}$ as given in the appendix A. Here, interface 0 describes the top of the model atmosphere and interface N indicates the surface level. Furthermore, S_n is the direct solar irradiance, F_n^\downarrow and F_n^\uparrow are the downward and upward diffuse fluxes, all defined per layer interface. Additionally, the external boundary conditions are given as

$$S_0 = \mu_0 F_0 \quad (31)$$

$$F_0^\downarrow = 0 \quad (32)$$

$$F_N^\uparrow = A_s (F_N^\downarrow + S_N). \quad (33)$$

where A_s is the surface albedo and $\mu_0 = \cos(\Theta_0)$ with the solar zenith angle Θ_0 . As example, we show the radiative TS-LINTRAN model for a three-layer model atmosphere to simplify matters. The approach can be generalized in a straight forward 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

$$\mathbf{MF} = \mathbf{C} \quad (34)$$

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Date : 17 May 2019	Page 31/74

with the sparse matrix

$$\mathbf{M} = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-a_{3,1} & -a_{5,1} & 1 & 0 & 0 & -a_{4,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-a_{1,1} & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-a_{2,1} & -a_{4,1} & 0 & 0 & 1 & -a_{5,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -a_{3,2} & -a_{5,2} & 1 & 0 & 0 & -a_{4,2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -a_{1,2} & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -a_{2,2} & -a_{4,2} & 0 & 0 & 1 & -a_{5,2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -a_{3,3} & -a_{5,3} & 1 & 0 & 0 & -a_{4,3} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -a_{1,3} & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -a_{2,3} & -a_{4,3} & 0 & 0 & 1 & -a_{5,3} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -A_s & -A_s & 1 & 0
\end{pmatrix}, \quad (35)$$

the flux vector

$$\mathbf{F} = \left[S_0, F_0^\downarrow, F_0^\uparrow, S_1, F_1^\downarrow, F_1^\uparrow, S_2, F_2^\downarrow, F_2^\uparrow, S_3, F_3^\downarrow, F_3^\uparrow \right]^T \quad (36)$$

and the right hand side

$$\mathbf{C} = \left[\mu_0 F_0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 \right]^T. \quad (37)$$

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.(34) 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

$$I^{\text{TOA}}(\mu_v) = \frac{F_N^\uparrow}{\pi} e^{-\tau_{\text{tot}}/\mu_v} + \frac{1}{\mu_v} \int_0^{\tau_{\text{tot}}} d\tau J(\tau, \mu_v) e^{-\tau/\mu_v}, \quad (38)$$

where $\mu_v = \cos(\Theta_v)$ with the viewing zenith angle Θ_v , τ indicates optical depth, and τ_{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

$$J(\tau, \mu_v) = \frac{\omega(\tau)}{2} \int_{-1}^{+1} d\mu' P(\tau; \mu_v, \mu') I(\tau, \mu') + \frac{\omega(\tau)}{4\pi} P(\tau; \cos \Theta_v) F_0 e^{-\tau/\mu_0}. \quad (39)$$

Here, the scattering phase function $P(\tau; \cos \Theta_v)$ is a function of scattering angle Θ_v between solar beam and instrument line of sight and optical depth τ and $P(\tau; \mu, \mu')$ is its azimuthal average with the cosine μ and μ' of the zenith angle of the incoming and outgoing direction. Next, we employ the relationship between the irradiances (flux) and mean radiances,

$$F^{\downarrow\uparrow} = 2\pi \int_0^{\pm 1} d\mu \mu I(\tau, \mu) = 2\pi \mu_{\downarrow\uparrow} I_{\downarrow\uparrow} \quad (40)$$

where $I_{\downarrow\uparrow}$ is the mean radiance in the upper and lower hemisphere and with the mean inclination $\mu_{\downarrow\uparrow} \in [0, \pm 1]$. 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 β ,

$$\beta(\tau, \mu) = \frac{1}{2} \int_{-1}^0 d\mu' P(\tau; \mu, \mu') = 1 - \frac{1}{2} \int_0^1 d\mu' P(\tau; \mu, \mu') \quad \text{for } \mu > 0, \quad (41)$$

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 32/74
--	--	--

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

$$\bar{\beta}(\tau) = \frac{1}{2} \int_0^1 d\mu \int_{-1}^0 d\mu' P(\tau; \mu, \mu') \quad (42)$$

(see e.g. Wiscombe and Grams, 1976, [RD48]). The expansion of the phase function in Legendre polynomials P_l ,

$$P(\tau; \mu, \mu') = \sum_{l=0}^{\infty} \chi_l(\tau) P_l(\mu) P_l(\mu') \quad (43)$$

with coefficients χ_l , provides a corresponding expansion of the back scattering ratios,

$$\beta(\tau, \mu) = \sum_{l=0}^{\infty} \frac{\chi_l(\tau)}{2} P_l(\mu) \int_{-1}^0 d\mu' P_l(\mu') \quad (44)$$

and

$$\bar{\beta}(\tau) = \sum_{l=0}^{\infty} \frac{\chi_l(\tau)}{2} \int_0^1 d\mu P_l(\mu) \int_{-1}^0 d\mu' P_l(\mu'). \quad (45)$$

Using Eq. (40) and (45), we can rewrite the source function as

$$J(\tau, \mu_v) = \frac{\omega}{2\pi} \{ U_{\uparrow} [1 - \bar{\beta}(\tau)] F^{\uparrow}(\tau) + U_{\downarrow} \bar{\beta}(\tau) F^{\downarrow}(\tau) \} + \frac{\omega}{4\pi\mu_0} P(\cos(\Theta_v)) S_0 e^{-\tau/\mu_0}. \quad (46)$$

Finally, to evaluate the integration in Eq. (38) of the scattering source function over optical depth, the upward and downward fluxes F^{\uparrow} and F^{\downarrow} within the scattering layer i are approximated by their mean values,

$$F^{\downarrow\uparrow}(\tau) = \frac{F_{i-1}^{\downarrow\uparrow} + F_i^{\downarrow\uparrow}}{2} \quad \text{for } \tau_{i-1} < \tau < \tau_i. \quad (47)$$

Thus, we obtain

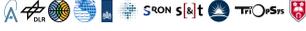
$$\begin{aligned} \frac{1}{\mu_v} \int_0^{\tau_{\text{tot}}} d\tau J(\tau, \mu_v) e^{-\tau/\mu_v} &= \sum_{n=1}^N \frac{\omega_n}{4\pi\mu_0\mu_v} \tilde{\mu} P_n t_n(\tilde{\mu}) S_{n-1} \\ &+ U_{\uparrow} \sum_{n=1}^N \frac{\omega_n}{4\pi\mu_v} (1 - \beta_n) t_n(\mu_v) [F_{n-1}^{\uparrow} + F_n^{\uparrow}] + U_{\downarrow} \sum_{n=1}^N \frac{\omega_n}{4\pi\mu_v} \beta_n t_n(\mu_v) [F_{n-1}^{\downarrow} + F_n^{\downarrow}] \end{aligned} \quad (48)$$

with $\tilde{\mu} = \frac{\mu_0\mu_v}{\mu_0 + \mu_v}$, $P_n = P(\tau, \cos(\Theta_v))$, $\omega_n = \omega(\tau)$, $\beta_n = \bar{\beta}(\tau)$ for $\tau_n < \tau < \tau_{n+1}$ and the auxiliary function

$$t_n(\mu) = e^{-\frac{\tau_{n-1}}{\mu_v}} \left(1 - e^{-\frac{\Delta\tau_n}{\mu}} \right) \quad \text{for } n = 1, \dots, N. \quad (49)$$

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

$$I^{\text{TOA}} = \mathbf{R}^T \mathbf{F} \quad (50)$$

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002
		Version : 3.1 Page
		Date : 17 May 2019 33/74

with the so-called response vector

$$\mathbf{R} = (u_0, v_0, w_0, \dots, u_N, v_N, w_N) \quad (51)$$

and its coefficients

$$u_n = \frac{\omega_{n+1}}{4\pi\mu_0\mu_v} \tilde{\mu} P_{n+1} t_{n+1}(\tilde{\mu}) \quad (52)$$

$$v_n = \frac{U^\downarrow}{4\pi\mu_v} [\omega_n \beta_n t_n(\mu_v) + \omega_{n+1} \beta_{n+1} t_{n+1}(\mu_v)] \quad (53)$$

$$w_n = \frac{U^\uparrow}{4\pi\mu_v} [\omega_n (1 - \beta_n) t_n(\mu_v) + \omega_{n+1} (1 - \beta_{n+1}) t_{n+1}(\mu_v)] \quad (54)$$

All indices on the right hand side for P , t , ω and β represent layers $(1, \dots, N)$ and the indices on the left hand side for u , v and w represent interface $(0, \dots, N)$. Any contributions from non-existing layers $(0$ or $N + 1)$ in Eq. (54) can just be ignored. The contribution of surface reflection is added to the response to diffuse upward radiation at the surface, redefining w_N .

$$w_N = \frac{U^\uparrow}{4\pi\mu_v} [\omega_N (1 - \beta_N) t_N(\mu_v)] + \frac{1}{\pi} e^{-\tau_N/\mu_v} \quad (55)$$

Equation (50) is an essential step towards the linearization of the model with respect to absorption and scattering properties of the model atmosphere, which we discuss below.

6.6.3 The forward-adjoint perturbation theory applied to the two-stream method

Due to multiple scattering, the functional dependence of I^{TOA} on optical depth, single-scattering albedo and on the scattering phase function characteristics is not obvious (see Eqs. (34) and (50)). In this context, the forward-adjoint perturbation theory provides a method, which is frequently used to linearize multiple scattering radiative transfer simulations with respect to absorption and scattering properties of the model atmosphere (e.g. [RD49, RD50, RD51, RD52, RD53, RD54, RD55, RD56, RD57]). Here, we summarize the main features of the perturbation theory, which is subsequently used to motivate our linearization approach of our two-stream model.

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

$$\mathbf{L}I = Q, \quad (56)$$

where I is the intensity, Q is the solar source and \mathbf{L} is the radiative transfer operator. The scalar observable can be obtained from the intensity with

$$I^{\text{TOA}} = \langle R, I \rangle, \quad (57)$$

The adjoint radiative transfer equation is formulated as

$$\mathbf{L}^\dagger I^\dagger = R, \quad (58)$$

Here \mathbf{L}^\dagger is the adjoint radiative transfer operator, which is related to \mathbf{L} according to forward-adjoint perturbation theory [RD50, RD58, RD59]. As a consequence,

$$I^{\text{TOA}} = \langle I^\dagger, Q \rangle, \quad (59)$$

so the adjoint intensity field represents the importance of sources within the atmosphere for a given observation I^{TOA} [RD60, RD61].

With the solutions of the forward and the adjoint transfer equations I and I^\dagger for an atmospheric state \mathbf{x}_0 , it is possible to calculate the derivatives of I^{TOA} with respect to an optical parameter x at x_0 by

$$\frac{\partial I^{\text{TOA}}}{\partial x} = -\langle I^\dagger, \mathbf{L}'I \rangle + \langle I^\dagger, Q' \rangle + \langle R', I \rangle \quad (60)$$

	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 34/74
---	--------------------------------	--

with the partial derivative

$$\mathbf{L}' = \lim_{\Delta x \rightarrow 0} \frac{\Delta \mathbf{L}}{\Delta x}, \quad (61)$$

and $R' = \partial R / \partial x$ and $Q' = \partial Q / \partial x$ (e.g. [RD62, RD55]).

In multi-stream radiative transport models, the adjoint radiative transport equation is solved using the pseudo-forward method. This method flips the directions of I^\dagger , so that the same operator can be used for \mathbf{L}^\dagger as for \mathbf{L} [RD52]. The symmetry relationships of light propagation allow such a method. Numerical approximations that are done to represent the intensity in a finite number of streams and on a finite vertical grid lack these symmetry relationships. So, using the pseudo-forward method introduces approximations in the adjoint equation that are inconsistent with those in the forward model. In two-stream methods, these approximations are significant and the inconsistency between the forward and adjoint equations can cause convergence problems in the inversion. Therefore, we avoid the pseudo-forward method by applying forward-adjoint perturbation theory on the matrix equation, Eq. (34) instead of the radiative transport equation, Eq. (56).

$$\mathbf{M}^\dagger \mathbf{F}^\dagger = \mathbf{R}, \quad (62)$$

The adjoint matrix \mathbf{M}^\dagger is given by its transposed, \mathbf{M}^T . Analogously to Eq. (60), the derivative of the TOA radiance with respect to an optical parameter x can be calculated by

$$\frac{\partial I^{\text{TOA}}}{\partial x} = -\langle \mathbf{F}^\dagger, \mathbf{M}' \mathbf{F} \rangle + \langle \mathbf{F}^\dagger, \mathbf{C}' \rangle + \langle \mathbf{R}', \mathbf{F} \rangle \quad (63)$$

with the derivatives $\mathbf{M}' = \frac{\partial}{\partial x} \mathbf{M}$, $\mathbf{C}' = \frac{\partial}{\partial x} \mathbf{C}$, and $\mathbf{R}' = \frac{\partial}{\partial x} \mathbf{R}$. With \mathbf{C} given in Eq. (37), the derivative \mathbf{C}' vanishes and so Eq. (63) simplifies to

$$\frac{\partial I^{\text{TOA}}}{\partial x} = -\langle \mathbf{F}^\dagger, \mathbf{M}' \mathbf{F} \rangle + \langle \mathbf{R}', \mathbf{F} \rangle \quad (64)$$

In general, x represents any forward model parameter and in our case, it comprises the optical depth $\Delta\tau_n$, the single-scattering albedo ω_n and the surface albedo A_s . Equation (64) can be numerically implemented in a straight forward manner, which is described in detail in Appendix A.

6.6.4 Physics-based CO column inversion

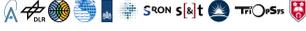
The SWIR-3 measurements are sensitive to the total amount of CO along the path of the measured light. Within the bounds of the measurement error, no information can be retrieved on the relative vertical distribution of CO. In the presence of clouds, the measurement loses sensitivity to the amount of 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. 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 dominated by noise contribution. Thus, the inversion requires regularization.

For this purpose, we employ the Tikhonov regularization technique [RD63, RD64] 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,

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \left\{ \|\mathbf{S}_y^{-1/2} (\mathbf{F}(\mathbf{x}) - \mathbf{y})\|^2 + \|\mathbf{W}\mathbf{x}\|^2 \right\}. \quad (65)$$

Compared to Eq. 4, Eq. 65 has additional term, where the state \mathbf{x} is mapped with matrix \mathbf{W} to a set of values that will be constrained in addition.

The state vector is slightly changed compared to the non-scattering retrieval from Sect. 6.4.2. It contains a full CO profile instead of one parameter for the CO column. In addition, effective cloud properties, the cloud optical depth and

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002	
		Version : 3.1	Page
		Date : 17 May 2019	35/74

the cloud center height, are added to the state vector. However, the methane column is removed. In this inversion, the methane column is fixed to the a-priori (from the CTM) to retrieve the effective cloud properties. So the state vector is defined as

$$\mathbf{x} = (\mathbf{x}_{\text{CO}}, c_{\text{H}_2\text{O}}, c_{\text{HDO}}, z_{\text{cld}}, \tau_{\text{cld}}, A_0, A_1, \delta_I). \quad (66)$$

where the CO profile is expressed as relative value compared to the reference profile.

$$\mathbf{x}_{\text{CO}} = \frac{s_{\text{CO}}}{s_{\text{CO}}^{\text{ref}}}, \quad (67)$$

The regularization \mathbf{W} constrains the first derivative of \mathbf{x}_{CO} with the discrete first derivative matrix \mathbf{L}_1 .

$$\mathbf{W} = \begin{bmatrix} \gamma \mathbf{L}_1 & \mathbf{0} \end{bmatrix} \quad (68)$$

with

$$\mathbf{L}_1 = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix} \quad (69)$$

The choice of the regularization parameter γ is of crucial importance for the inversion. The smaller γ is taken, the more information is extracted from the measurement. When extracting too much information from the measurement, also pieces of information are retrieved for which the instrument noise propagation is high, resulting into a noisy end result. With higher γ , only pieces of information with low noise propagation are retrieved resulting into a lower noise contribution, although some information may be missing. Information that is missing for this reason is referred to as the null-space error or smoothing error. As discussed in the beginning of this section, no profile information of CO can be retrieved from the measurement, leaving only one degree of freedom, the total column. A degree of freedom of one can be achieved with this regularization scheme by choosing $\gamma \rightarrow \infty$ (e.g. [RD65]).

Compared to the standard retrieval of Sect. 6.2, the gain matrix changes because of the existence of \mathbf{W} to

$$\mathbf{G} = (\mathbf{K}^T \mathbf{S}_y^{-1} \mathbf{K} + \gamma^2 \mathbf{W}^T \mathbf{W})^{-1} \mathbf{K}^T \mathbf{S}_y^{-1} \quad (70)$$

Due to the regularization, the retrieved state $\hat{\mathbf{x}}$ is a smoothed version of the true state \mathbf{x}_{true} , where the smoothing is characterized by the averaging kernel,

$$\mathbf{A} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}_{\text{true}}} = \mathbf{G} \mathbf{K}. \quad (71)$$

So, the retrieved state vector $\hat{\mathbf{x}}$ can be written as

$$\hat{\mathbf{x}} = \mathbf{A} \mathbf{x}_{\text{true}} + \mathbf{e}_x, \quad (72)$$

where $\mathbf{e}_x = \mathbf{G} \mathbf{e}_y$ represents the error on the retrieved CO profile caused by the error \mathbf{e}_y .

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

$$\hat{c}_{\text{CO}} = \mathbf{C} \hat{\mathbf{x}} = \sum_k \hat{s}_{k,\text{CO}}, \quad (73)$$

	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 36/74
---	--------------------------------	--

where $\hat{s}_{k,CO}$ is acquired from Eq. (67) together with the CO profile parameters from $\hat{\mathbf{x}}$, \hat{c}_{CO} is the total CO column and \mathbf{C} is the corresponding operator. Using this formulation, we can characterize the effect of regularization on the CO column by:

$$\hat{c}_{CO} = \mathbf{CA}\mathbf{x}_{true} + \mathbf{e}_c, \quad (74)$$

where \mathbf{CA} is the sensitivity of the retrieved CO column to changes in the true state. Because only the CO profile is regularized, only changes in the CO-related part of the true state influence the retrieved CO column, so we can define the column averaging kernel as the unitless sensitivity of the CO retrieval to CO at different altitudes, using Eqs. (67) and (74).

$$\begin{aligned} \mathbf{A}_c &= \frac{\partial \hat{c}_{CO}}{\partial s_{CO,true}} \\ &= \mathbf{CA} \frac{\partial \mathbf{x}_{true}}{\partial s_{CO,true}} \\ &= \mathbf{CA} \begin{bmatrix} \text{diag}\left(\frac{1}{s_{CO}^{rel}}\right) \\ \mathbf{0} \end{bmatrix} \end{aligned} \quad (75)$$

This is the column averaging kernel, an important diagnostic tool for inverse modellers, who need the relationship between the retrieved CO columns from S5 and the sub-columns in their chemical transport models.

The retrieval noise on the retrieved column is given by the standard deviation

$$\sigma_c = \sqrt{\mathbf{CS}_x\mathbf{C}^T}. \quad (76)$$

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

Borsdorff et al. [RD65] showed that for the definition of the CO state vector \mathbf{x}_{CO} in Eq. (67) and for a regularization parameter $\gamma \rightarrow \infty$, the solution of the Tikhonov minimization problem (65) is identical to a simulation least squares approach in Eq. (4) where the state vector contains the total CO column instead of the vertically resolved CO profile. The formalism of a regularized profile retrieval, however, enables the calculation of the column averaging kernel of CO, which we desire. Transforming from a state with CO profile to a state with a CO column is defined as

$$\mathbf{x}_{col} = \mathbf{Z}\mathbf{x} \quad (77)$$

with

$$\mathbf{Z} = \begin{bmatrix} \mathbf{C} \\ \mathbf{0} \quad \mathbf{I} \end{bmatrix}. \quad (78)$$

This transformation uses operator \mathbf{C} to acquire the CO column from the state and leaves the non-CO parameters as they are. The Jacobian and gain matrices for the column retrieval are related to those of the profile retrieval by

$$\mathbf{K}_{col} = \mathbf{K}\mathbf{Z}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \quad (79)$$

$$\mathbf{G}_{col} = \mathbf{Z}\mathbf{G} \quad (80)$$

To calculate the column averaging kernel, the intermediate result $\mathbf{CA} = \mathbf{C}\mathbf{G}\mathbf{K}$ needs the profile Jacobian, but does not need the profile gain matrix.

$$\begin{aligned} \mathbf{C}\mathbf{G}\mathbf{K} &= \mathbf{Z}^{\text{CO}}\mathbf{G}\mathbf{K} \\ &= \mathbf{G}_{col}^{\text{CO}}\mathbf{K}. \end{aligned} \quad (81)$$

In the numeric implementation in SICOR, the retrieval is performed as an unconstrained column retrieval using the Jacobian \mathbf{K}_{col} from Eq. (79) and calculating the gain \mathbf{G}_{col} by performing the inversion with \mathbf{x}_{col} as state vector. But the profile Jacobian \mathbf{K} is calculated in addition, so that the column averaging kernel can be calculated using Eq. (81).

	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002
		Version : 3.1 Page Date : 17 May 2019 37/74

Using the unconstrained column retrieval saves computation time, because the state vector is smaller. And it is more stable, because we do not have to cope with numerically challenging values of $\gamma \rightarrow \infty$.

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

$$e_n = c_{\text{CO,true}} - \hat{c}_{\text{CO}} = (\mathbf{I} - \mathbf{A}_c) s_{\text{CO,true}} \quad (82)$$

This is the so-called null space contribution of a profile $s_{\text{CO,true}}$ and is also known as smoothing error of the retrieval [RD66]. [RD65] have shown that the null space contribution of any scaled reference profile $s_{\text{CO,true}} = X s_{\text{CO}}^{\text{ref}}$ is zero. 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, Eqs. (81) and (82) have to be accounted for a proper interpretation of the retrieval product.

The S5 L2 CO product also includes a QA-value. This is a number ranging from zero to one indicating the reliability of the product. the QA value is calculated as follows:

- If the retrieval did not converge, the QA-value is set to 0.
- If the retrieval converged and the retrieved optical thickness is less than 0.5 and the retrieved scattering height excluding null-space filling is less than 500 m, the scene is considered clear-sky and the QA-value is set to 1.
- Otherwise, if the optical thickness is larger than 0.5 and the retrieved scattering height (without null-space filling) is less than 5 km, the retrieved behaved as expected for a common moderately cloudy scene and the QA-value is set to 0.7.
- If the retrieved optical thickness is less than 0.5, but the retrieved scattering height (without null-space filling) is above 500 m, the retrieved results are considered abnormal and the QA-value is set to 0.4.
- If the retrieved scattering height is above 5 km, the scene is considered a high-cloud scene and the QA-value is also set to 0.4.

6.6.5 Inversion stability measures

The solution of our minimization problem is vulnerable to state variables to which the measurement is occasionally insensitive. For example, in case of a measurement scene fully overcast by an optically thick cloud, the measurement has no sensitivity to the surface albedo. In this case, one or more eigenvalues of $\mathbf{K}^T \mathbf{S}_y^{-1} \mathbf{K}$ approach zero, leading to extremely large values in \mathbf{G} . Therefore, a side constraint similar to that of Eq. (65) is desirable.

Instead of taking a \mathbf{W} based on the first derivative matrix, we take a regularization matrix based on the identity matrix. In order to have the least effect on our main product, the CO column, we restrict regularization to those state variables to which the modeled reflectance can be insensitive. Those are the albedo parameters and the scattering layer height. So,

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \left\{ \|\mathbf{S}_y^{-1/2} (\mathbf{F}(\mathbf{x}) - \mathbf{y})\|^2 + \gamma^2 \|\mathbf{W}\mathbf{x}\|^2 \right\} \quad (83)$$

with state vector \mathbf{x} defined as

$$\mathbf{x} = (c_{\text{CO}}, c_{\text{H}_2\text{O}}, c_{\text{HDO}}, z_{\text{cld}}, \tau_{\text{cld}}, A_0, A_1, \delta_I) . \quad (84)$$

and \mathbf{W} is defined as

$$\begin{aligned} W_{ij} &= \gamma \quad \text{for } i = j \in (A_0, A_1, z_{\text{cld}}) \\ &= 0 \quad \text{otherwise} . \end{aligned} \quad (85)$$

	S5L2PP	Reference : SRON-ESA-S5L2PP-ATBD-002
	Carbon Monoxide ATBD	Version : 3.1 Page
		Date : 17 May 2019 38/74

The gain matrix is defined similarly to Eq. (70), only with another \mathbf{W} :

$$\mathbf{G} = (\mathbf{K}^T \mathbf{S}_y^{-1} \mathbf{K} + \mathbf{W}^T \mathbf{W})^{-1} \mathbf{K}^T \mathbf{S}_y^{-1}. \quad (86)$$

The averaging kernel $\mathbf{A} = \mathbf{G}\mathbf{K}$ differs from the identity matrix for $\gamma \neq 0$, leading to a null-space error. Because of the zero elements in \mathbf{W} for the carbon monoxide column c_{CO} , the null-space contribution for the total carbon monoxide column can be neglected.

The underlying formalism of calculating the column averaging kernel for CO using the idea of an infinitely regularized CO profile retrieval remains the same.

Another source of instability can be the state vector itself. In a scattering retrieval, unphysical atmospheric states exist within the state space (e.g. a negative cloud optical depth). Running the two-stream model with such state leads to numeric problems or nonsense output, so unphysical states should be avoided. The problem is that the inversion is based on a linearized forward model, which is defined for the full state space. This can lead to an inversion that tends to unphysical values.

To avoid the risk of unphysical values of the state vector, two mechanisms are added to the algorithm: After each iteration, we add the null-space contribution $(\mathbf{I} - \mathbf{A})\mathbf{x}_{\text{a priori}}$ with an a priori estimate of the state vector $\mathbf{x}_{\text{a priori}}$ during the iteration. This contribution is not added to the final result. If a state vector still contains unphysical values for a given iteration, the iteration step is scaled accordingly so that the state is just legal (e.g. zero optical depth). Subsequently, that parameter is fixed and left out of the state vector for the upcoming few (typically 3) iterations. This exclusion prevents a repetitive trigger of this event with paralyzing effect on the inversion. After those few iterations, the variable is returned to the inversion, allowing it to tend to physical (positive) values again.

This means an ad hoc regularization of the inversion for a few iterations. It affects only the search path of the minimization 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 parameters are excluded from the inversion. Two reasons may cause this result: First, the true value of the state vector is just legal (e.g. almost zero optical depth) 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 straight forward and depends on the particular error source. For these cases, the retrieval output is error flagged.

6.7 Deconvolution solar spectrum

In Sect. 6.4.1, it was mentioned that as an alternative of using an a-priori model solar spectrum, the solar spectrum can be deduced from S5 irradiance measurements. We assume that there is only one irradiance measurement available during the execution of the processor, so the deconvolution of the solar spectrum is executed at initialization.

Analogous to Eq. (19), we can simulate the solar measurement $\mathbf{F}_{0,\text{meas}}$ by

$$\mathbf{F}_{0,\text{meas}}(\lambda_i) = \mathbf{S}_{\text{sun}} \mathbf{F}_0(\lambda) \quad (87)$$

where \mathbf{S}_{sun} is the ISRF matrix of the solar measurement analogue to Eq. (19). Before executing the deconvolution, the solar measurement is spectrally shifted to ensure that the Fraunhofer lines of the solar measurement are aligned with the model solar spectrum. For this purpose, we apply a line search algorithm, where the S5 irradiance measurements are compared with the model solar spectrum convolved with the ISRF,

$$\delta_S = \max_{\delta_S} \rho_{\mathbf{F}_{0,\text{meas}}(\lambda+\delta_S), \mathbf{F}_0^*(\lambda)} \quad (88)$$

where $\rho_{X,Y}$ is the Pearson's correlation coefficient.

After correcting the spectral shift, the deconvolution is executed by inverting Eq. 87, where $\mathbf{F}_{0,\text{meas}}$ is the measurement (\mathbf{y}), the deconvolved solar spectrum \mathbf{F}_0 represents the state to be retrieved (\mathbf{x}) and the ISRF matrix \mathbf{S}_{sun} is the Jacobian (\mathbf{K}). This is a linear problem, so no iterative approach is required. However, this problem is underdetermined because the measurement vector contains fewer values than the state to be retrieved.

	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002
		Version : 3.1 Page
		Date : 17 May 2019 39/74

Table 3: Composition of of state vector \mathbf{x} . LER means Lambert-equivalent reflectivity.

State vector element	Non-scattering	Physics-based	A-priori value	Regularization (only in physics-based)
CO total column	yes	yes	CTM/DEM	0
CH ₄ total column	yes	no	CTM/DEM	0
H ₂ O total column	yes	yes	CTM/DEM	0
HDO total column	yes	yes	CTM/DEM	0
Lambertian surface albedo	yes	yes	LER	L0
Linear spectral dependence of Lambertian surface albedo	yes	yes	0	L0
Cloud height	no	yes	METimage	L0
Cloud optical depth	no	yes	1	0
Spectral offset	yes	yes	0	0

Van Deelen et al. (2007) [RD67] showed that the least squares minimum length solution, which minimizes the length of the solution vector as a side constraint, is of sufficient accuracy to simulate earth radance measurements of the GOME mission. Following this approach, we obtain

$$\mathbf{F}_0 = \mathbf{S}_{\text{sun}}^T (\mathbf{S}_{\text{sun}} \mathbf{S}_{\text{sun}}^T)^{-1} \mathbf{F}_{0,\text{meas}}, \quad (89)$$

which contains significant noise contributions. However, this noise is in the null-space of the ISRF and will affect the simulation only little after the final convolution of the earth radance spectrum with its corresponding ISRF. Moreover, to mitigate edge effects, we execute the deconvolution on a slightly extended window (~ 3 nm extra on both sides).

Assuming that the earth radance and the solar irradiance are affected similarly by instrument and calibration errors, the use of \mathbf{F}_0 instead of an a priori solar reference spectrum has a clear advantage to reduce spectral fit residuals. For further details and for an overview of the benefits of this method see [RD67] and [RD68].

6.8 State vector, ancillary parameters and a priori knowledge

Table 3 summarizes the state vector definition for the cloud filter based on the CH₄ non-scattering retrieval and the physics based retrieval of CO. The Lambert-equivalent reflectivity, mentioned as a-priori surface albedo, is the surface albedo that matches the highest measured reflectance at zero atmospheric absorption or scattering. The algorithm requires a priori knowledge on the following atmospheric parameters:

1. CO mixing ratio profile.
2. CH₄ mixing ratio profile.
3. Temperature profile.
4. Specific humidity profile.
5. Surface elevation.
6. Surface pressure.
7. Liquid cloud top pressure.

All atmospheric profiles must be provided as a function of atmospheric pressure on the same vertical and horizontal grid. Here, surface elevation and surface pressure are needed to resample the profile information on the retrieval grid. Because the retrieval method does not rely on any statistical regularization method, no estimate of the statistical variability (e.g. covariance) of the a priori information is needed.

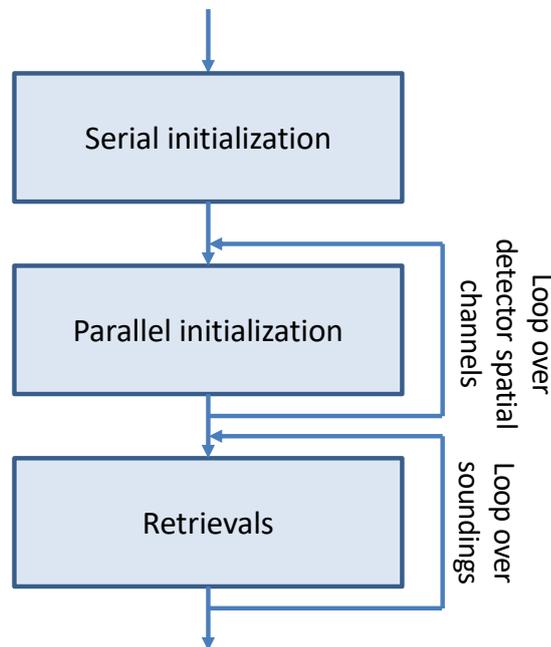


Figure 6: SICOR algorithm architecture including the algorithm initialization and the pixel processing. Here, two parallelization options are foreseen: first, the initialization of the algorithm with respect to the detector spatial index of the measurements and second, the actual retrievals for the pixels to be processed within a data granule.

6.9 Detailed Algorithm Setup

This section focuses on the detailed algorithm structure of SICOR. The SICOR software implementation separates the initialization from the actual ground-pixel based retrieval to allow performance optimization within an overall processing framework. Here, we distinguish between the initialization, whose computation can be parallelized, from a serial initialization. Finally, the actual retrieval is performed per ground scene in a separate software module. Parallelization of this processing step is essential for performance optimization. Figure 6 depicts the software break down in three main modules, which are discussed in more detail below.

6.9.1 Serial Initialization

The serial initialization provides the algorithm input, which has to be allocated once per processing call. The numerical effort is minor and so parallelization of the software is not considered. The module requires static input data, indicated by the solid green data boxes in Fig. 7. It comprises specific retrieval settings, lookup tables for molecular absorption cross sections, the ISRF as part of the instrument calibration key data, a solar reference spectrum and the L1B solar irradiances. Furthermore, the module defines the internal line-by-line spectral grid and the optical properties of the scattering layer. All data are accessible for the remaining software via pointers to the internal memory, which are indicated by the green dashed data boxes in the figure. The static input of this algorithm component is summarized in Tab. 4. The module returns the static output listed in Tab. 5, which is applicable for all retrievals performed during this processor call.

6.9.2 Parallel initialization

Part of the initialization includes independent processing steps for the different viewing geometries within the across-track swath and so can be easily computed in parallel. For example, the calibration key data of the ISRF has to be adapted for the spectral and spatial pixel depending on the swath position. Here, the swath position of a ground

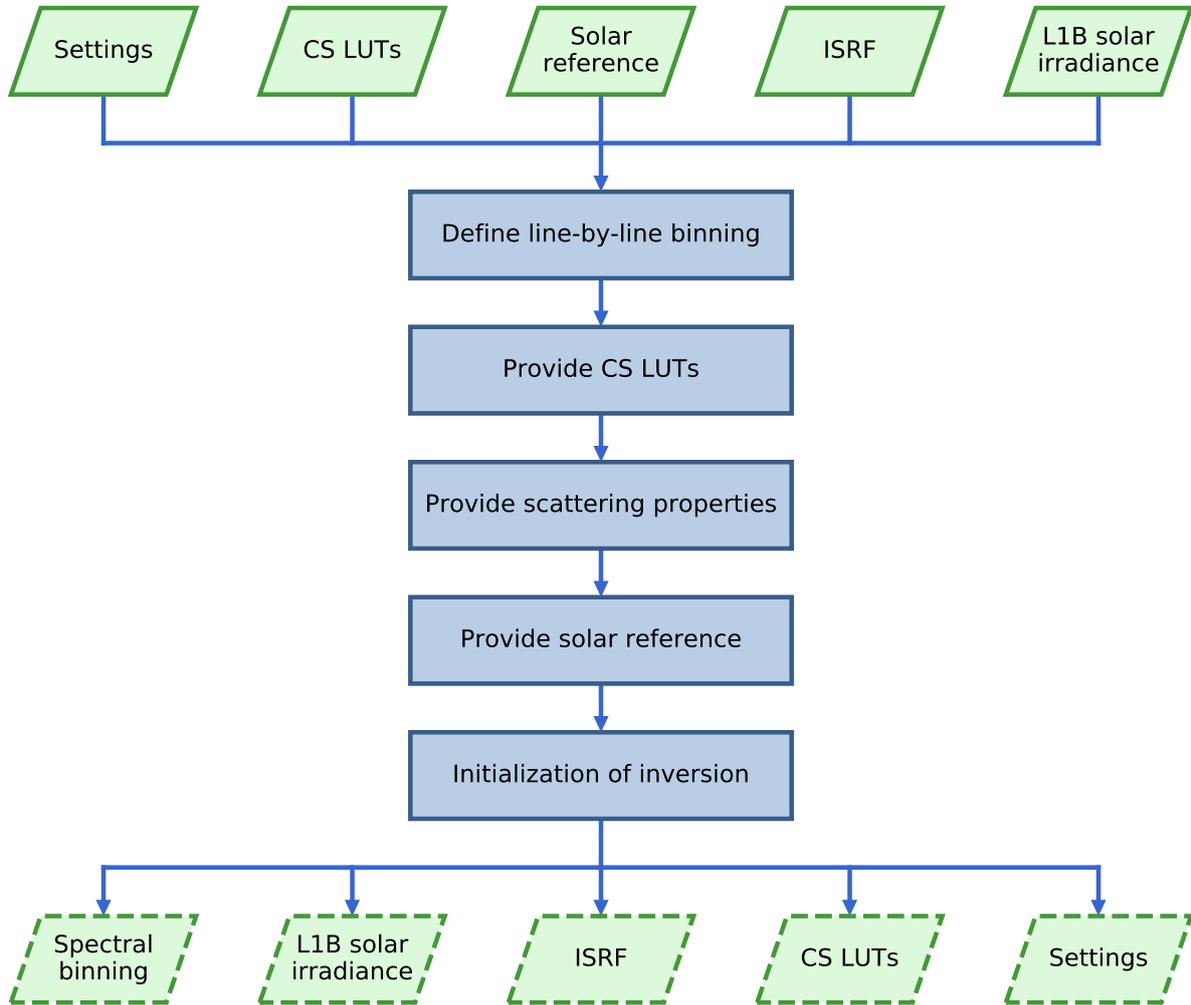


Figure 7: Serial SICOR initialization, which is executed once per processor call. The solid green boxes indicate static input, the dashed green boxes indicate initialization results accessible for the remaining SICOR software.

Table 4: Static input of the serial algorithm initialization.

Parameter	Symbol	Source
Algorithm settings	-	S5L2PP
Absorption cross sections CO, CH ₄ , H ₂ O and HDO	$\sigma_{\text{CO}}, \sigma_{\text{CH}_4}, \sigma_{\text{H}_2\text{O}}, \sigma_{\text{HDO}}$	LUT
Reference solar spectrum	$F_{0,\text{ref}}$	LUT
Instrument spectral response function	S	CKD
Irradiance spectrum SWIR-3	F_0	L1B
CAMS hybrid pressure coefficients A, B	$A_{\text{hyb}}, B_{\text{hyb}}$	CAMS

Table 5: Output data of the static initialization.

Parameter	Symbol	Destination
Model atmosphere grid	ATM	S5 L2 CO product

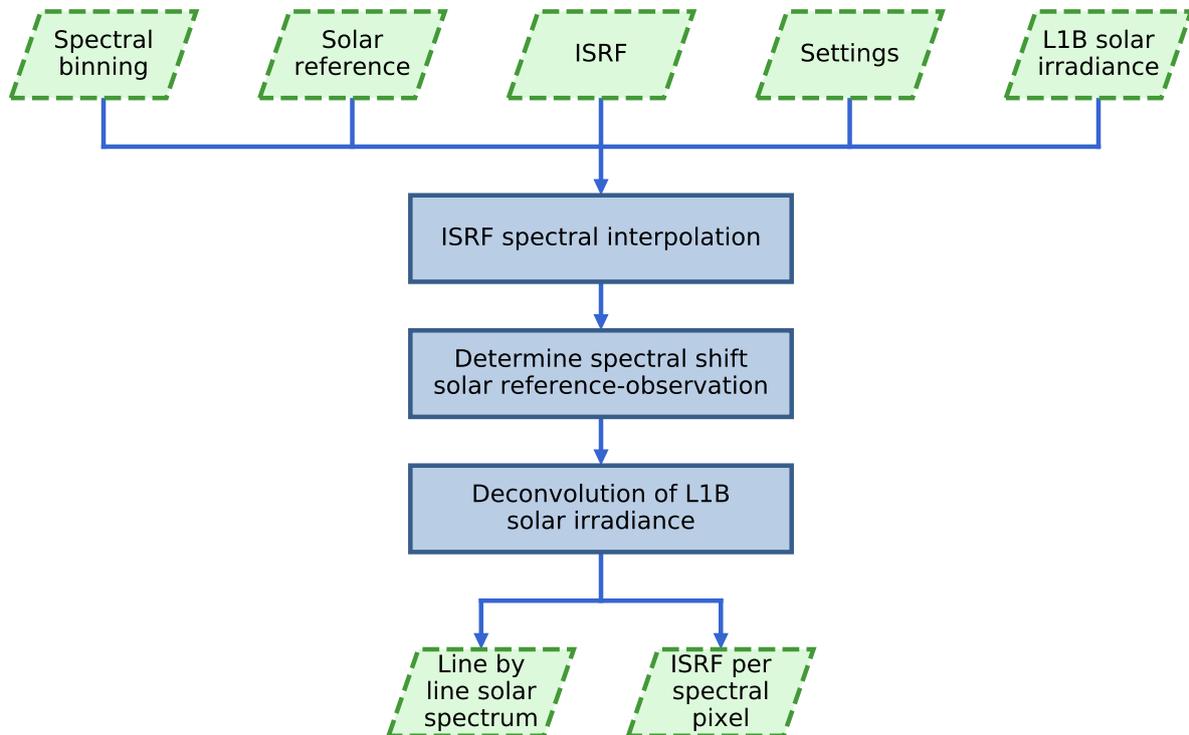


Figure 8: Parallel initialization: The ISRF is interpolated for each spectral pixel. Subsequently, the spectral shift between solar reference spectrum and observation is determined. This is needed as input for the subsequent deconvolution of the solar spectrum.

pixel is an instrument specific property that can be indicated by the viewing geometry or the detector spatial channel index. Similarly, due to instrument related issues, the measured solar spectrum may vary with swath position. Therefore, the deconvolution of the solar spectrum (Sect. 6.7) can be parallelized over detector spatial channel index. The parallel initialization of SICOR is outlined in Fig. 8. The module does not receive any input from the calling framework nor does it provide output to it. Data transfer from the serial initialization or to the retrieval part of SICOR is managed by pointers to internal memory.

6.9.3 Retrievals

The interpretation of S5 SWIR-3 measurements to infer CO total column information is performed per individual ground pixel and so can be computed in parallel for all observations of the data granule to be processed. The parallelization of this processing step is essential to achieve a satisfying processing time. This processing step includes two sequential retrievals, which are based on the same software tool. The corresponding software architecture of the processing step is depicted in Fig. 9, reflecting the algorithm design of Fig. 2. Several input data stem from the initialization and are provided via pointers to the internal memory, like retrieval settings, the line-by-line solar irradiances, the corresponding spectral grid and the ISRF. The dynamic input is summarized in Tab. 6 and includes the L1B radiances, the surface elevation from the digital elevation map (DEM), the atmospheric trace gas profiles from the chemical transport model (CTM) and cloud information from METimage. The model atmosphere is constructed using the atmospheric input from the DEM and the CTM.

After the data filtering A, discussed in Sect. 6.3, the non-scattering retrieval is performed for the spectral window specified in Tab. 1. Subsequently, filter B rejects observations with high and optically thick clouds for further processing and for the remaining data, the physics based CO retrieval is performed. Finally, the output, which is summarized in Tab. 7, is made available to the framework.

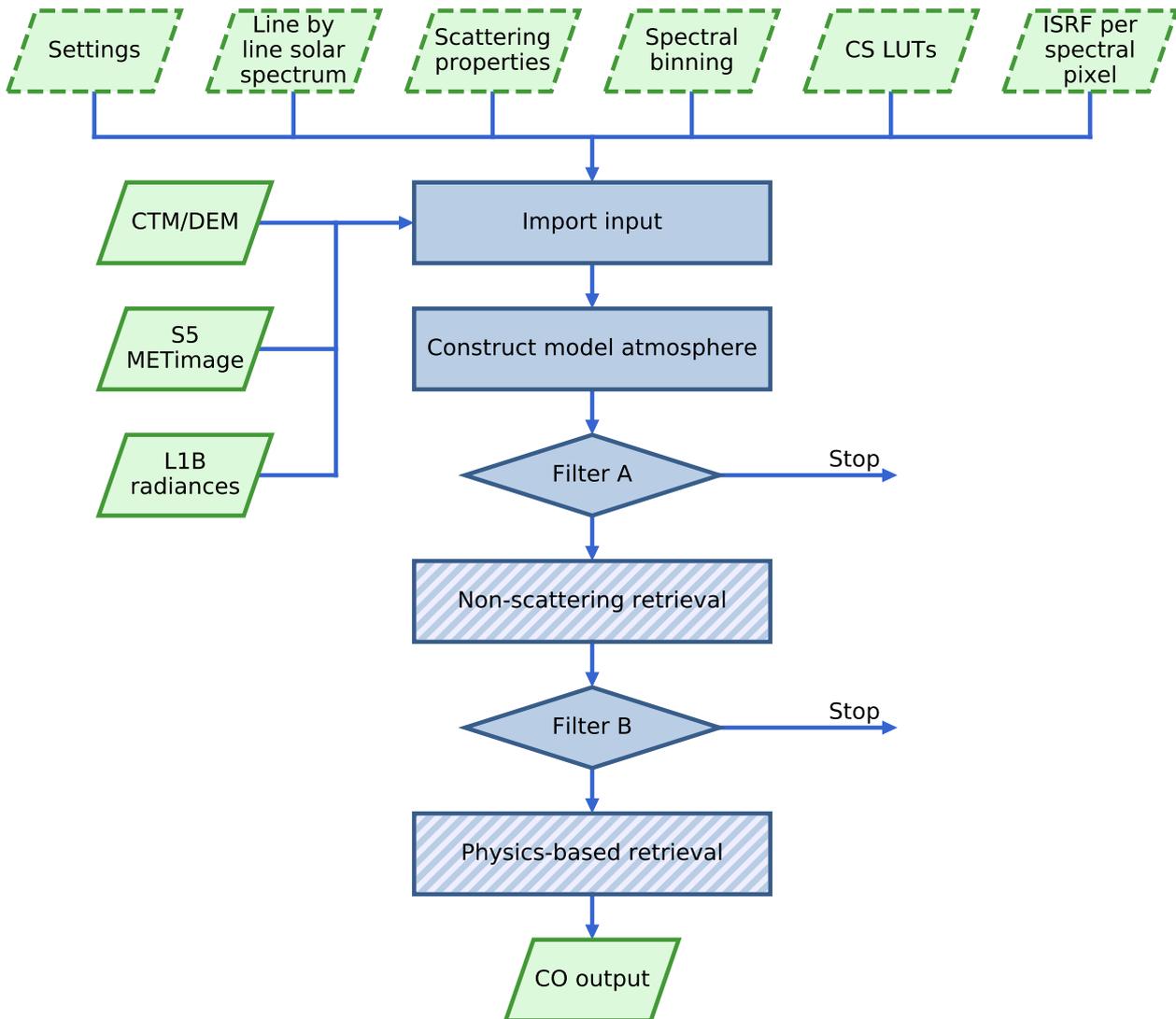


Figure 9: Software architecture of the S5 CO retrieval tool. The shaded boxes indicate two specific calls of the retrieval model (non-scattering retrieval and physics-based retrieval), which is explained in more detail in Fig. 10.

Both the non-scattering and physics based retrievals are performed by the same software tool, initialized with different settings. The software architecture of this component is depicted in Fig. 10 and first defines the measurement vector for the specific spectral fit window. The cross sections of CO, CH₄, H₂O and HDO are calculated on the model atmosphere. Before the inversion starts, all parameters to be retrieved are set to their first guesses as shown in Tab. 3. Each iteration step of the inversion starts with simulations of line-by-line radiance spectra at the top of the model atmosphere in the viewing direction of the instrument, including all required derivatives. The spectral convolution with the ISRF completes the forward simulations and the model is evaluated. In the inversion process, a new state vector is constructed and the next iteration starts. During the evaluation of the model, convergence is checked and the step size management is updated by checking the χ^2 of the spectral fit residuals. In case of convergence, results are provided to the calling routine.

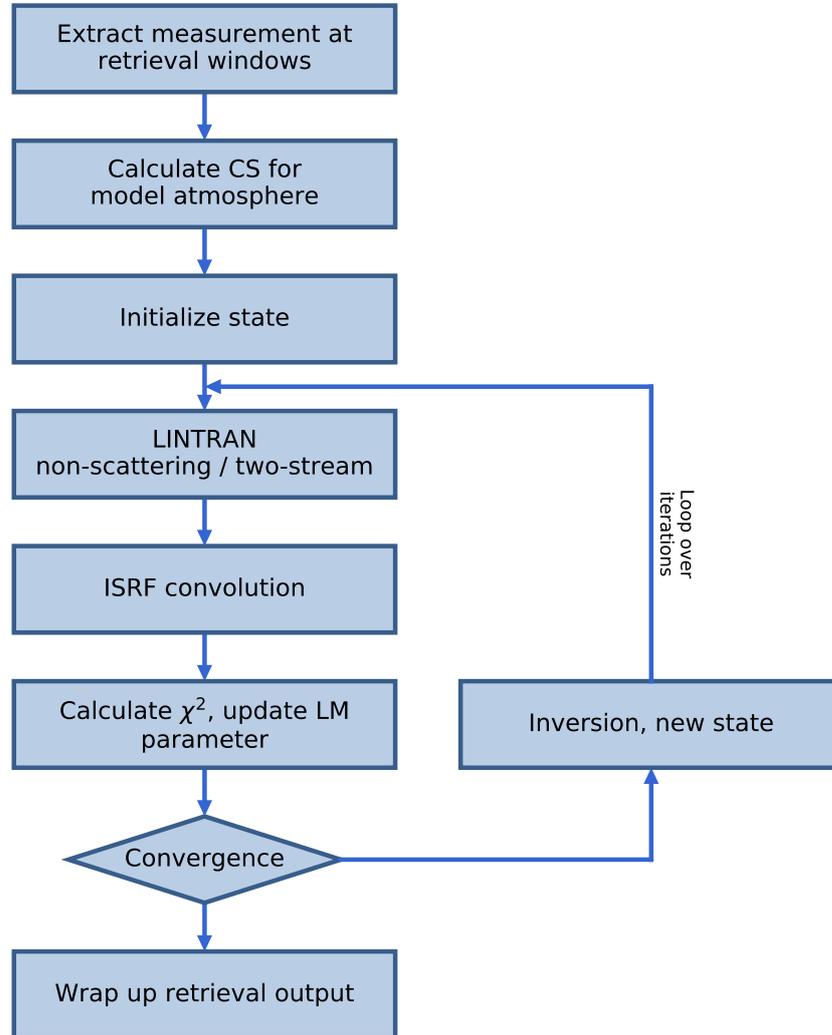


Figure 10: Software architecture of the non-scattering and physics-based retrieval.

Table 6: Dynamic input to the CO retrieval algorithm. Metimage input as specified in Tab. 2

Parameter	Symbol	Source
Radiance spectrum SWIR-3	I	L1B
Solar geomtry	ϑ_0, φ_0	L1B
Viewing geomtry	ϑ_v, φ_v	L1B
Surface elevation of pixel	z_{surf}	DEM
CAMS surface pressure	$p_{\text{CAMS, surf}}$	CAMS
CAMS CO and CH ₄ mixing ratio profiles	$\rho_{\text{CO}}, \rho_{\text{CH}_4}$	CAMS
CAMS surface elevation	$z_{\text{surf, CAMS}}$	CAMS
CAMS specific humidity profile	q	CAMS
CAMS temperature profile	T	CAMS

Table 7: Dynamic output data

Parameter	Symbol	Destination
CO total column	C_{CO}	S5 L2 CO product
CO column noise estimate	e_{CO}	S5 L2 CO product
CO column averaging kernel	A_c	S5 L2 CO product
CH ₄ column non-scattering	$C_{CH_4, noscat}$	S5 L2 CO product
H ₂ O total column	C_{H_2O}	S5 L2 CO product
HDO total column	C_{HDO}	S5 L2 CO product
Cloud optical depth and height	τ_{cld}, z_{cld}	S5 L2 CO product
SWIR-3 surface albedo	A_s	S5 L2 CO product
SWIR-3 spectral	δ_I	S5 L2 CO product
Number of iterations	N_{iter}	S5 L2 CO product
Reduced chi square of the spectral fit	χ^2	S5 L2 CO product

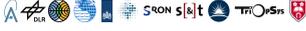
 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002
		Version : 3.1 Page
		Date : 17 May 2019 46/74

Table 8: Microphysical properties of water and ice clouds: $n(r)$ represents the size distribution type, r_{eff} and v_{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].

	water clouds	ice clouds
$n(r)$	gamma	$(r/r_1)^{-3.85}$
r_{eff} [μm]	20	-
v_{eff}	0.10	-
n_r	1.28	1.26
n_i	$4.7 \cdot 10^{-4}$	$2.87 \cdot 10^{-4}$

Table 9: S5 instrument noise model

	A	B	N
NIR 750 nm	$6.47 \cdot 10^{-9}$	187	36
SWIR-1 1630 nm	$2.27 \cdot 10^{-8}$	193	9
SWIR-3 2344 nm	$7.00 \cdot 10^{-8}$	212	3

7 Error Analysis

For individual CO observations, the Sentinel-5 mission envisages a product precision $< 4 \cdot 10^{17}$ molec./cm² and a product bias $< 5\%$ for total CO column densities larger than $1.6 \cdot 10^{18}$. This section presents the results of Landgraf et al., 2016 [RD35] discussing the CO retrieval sensitivity of our algorithm to forward model errors and a set of key atmospheric and instrument parameters and compare these errors to the envisaged product uncertainties. To estimate the retrieval accuracy, we have generated synthetic measurements for generic test cases using the S-LINTRAN radiative transfer model [RD69]. The model is a scalar plane-parallel radiative transfer model that fully accounts for multiple elastic light scattering by clouds and air molecules and the reflection of light at the Earth surface. The optical properties of clouds are calculated using Mie theory assuming the microphysical cloud properties given in Tab. 8. For ice clouds, the ray tracing model by Hess et al., 1994 and 1998 [RD70, RD71] is used. Finally, we describe cirrus and clouds by their top and base heights, and cloud optical thickness at 2315 nm. We assume that cirrus fully overcasts the observed scene, whereas broken cloud coverage is addressed by the independent pixel approximation [RD72]. Moreover, we assume the US standard atmosphere [RD73] for the profiles of dry air density, pressure, water and CO. The CH₄ profile is taken from the European background scenario of [RD74]. In all cases, the model atmosphere comprises a CO column of $2.38 \cdot 10^{18}$ molec./cm² and so for this specific scenario, the CO product requirements are $< 17\%$ precision and $< 5\%$ bias.

The radiance spectra are perturbed by measurement noise from the ESA's Sentinel 5 noise model for Earth shine observations,

$$\text{SNR}_{\text{EARTH}} = \sqrt{N} \frac{A I}{\sqrt{A I + B^2}} \quad (90)$$

(pers. comm. H. Weber, ESA, May 12, 2017) where A and B are instrument specific parameters and N represents the spectral binning factor. The Earth radiance I is given in photons/cm² s sr nm. Parameter A depends on wavelength, however for simplicity reason we consider only one representative value per band, which is summarized in Table 9.

To evaluate the performance of our algorithm, we initialize the retrieval with an atmospheric state, which differs from the truth. For example, first guess surface albedo is given by the maximum Lambert-equivalent reflectivity (LER) value of the measurement defined by

$$\text{LER}(\lambda) = \frac{I^{\text{TOA}}(\lambda)\pi}{\mu_o F_o(\lambda)}, \quad (91)$$

Moreover, we assume a triangular scattering layer centered at 5 kilometers altitude, with a fixed geometric thickness of 2.5 km and an optical depth of 1. The retrieval performance is characterized by the retrieval noise σ_{CO} of the CO

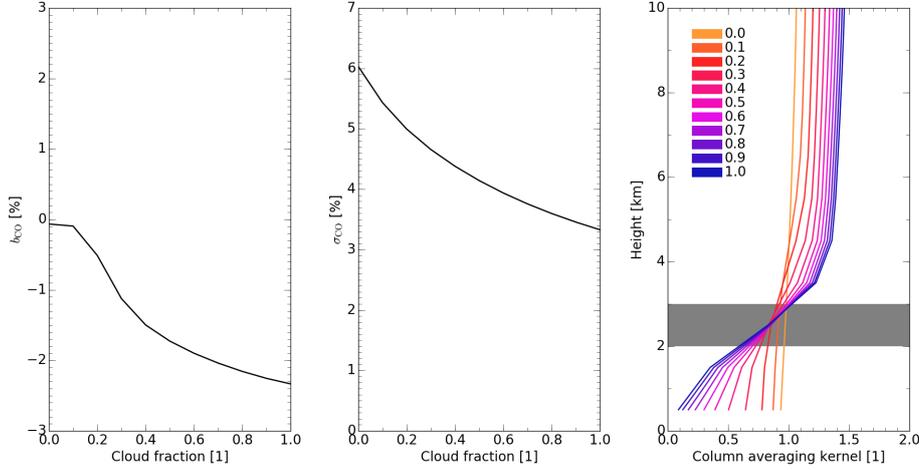


Figure 11: Example of the S5 CO data product and its performance as a function of cloud fraction f_{cld} . The SWIR-3 measurements are simulated for a scene partially covered by a cloud between 2 and 3 km with optical depth $\tau_{\text{cld}} = 5$, a surface albedo $A_s = 0.05$, a solar zenith angle of 50° and a viewing zenith angle of 40° . Left panel: CO retrieval bias b_{CO} . Middle panel: retrieval noise σ_{CO} . Right panel: column averaging kernel for different cloud fractions as indicated in the legend. The grey area indicates the position of the cloud.

column (see Eq. (76)) and the CO bias defined by

$$\Delta\text{CO} = C_{\text{CO}} - \mathbf{G}_{\text{CO}}\mathbf{e}_y - \mathbf{A}_c \rho_{\text{CO,true}}, \quad (92)$$

where \mathbf{A}_c is the CO column averaging kernel from Eq. (74), $\rho_{\text{CO,true}}$ is the true CO profile from the test ensemble, C_{CO} represents the retrieved CO column, and $\mathbf{G}_{\text{CO}}\mathbf{e}_y$ estimates the noise of the CO column using the CO relevant contribution of the gain matrix \mathbf{G}_{CO} .

Figure 11 shows an example of the CO retrieval performance for simulated measurements with increasing cloud coverage over land and a dark land surface with an albedo $A_s = 0.05$. It depicts the retrieval bias b_{CO} , the retrieval noise σ_{CO} , and the column averaging kernel. The retrieval biases increases to 2.3% with increasing cloud fraction because deficits of our cloud model become more relevant with increasing cloud coverage. At the same time, the retrieval noise of the CO column decreases due to the gain in the measurement signal. The change of the retrieval sensitivity with cloud coverage is clearly illustrated by the column averaging kernels shown in the right panel of Fig. 11. When the cloud fraction is greater than zero, the column averaging kernel starts to increase above the cloud and at the same time decreases below the cloud and so reflects the effect of cloud shielding on the retrieved column utilizing the profile scaling approach [RD65].

Similar results were already presented by Vidot et al., 2012 [RD33], who used a previous version of the SICOR algorithm. In their study, clouds were accounted for in the retrieval by an elevated Lambertian reflector. This approach appeared to be appropriate to describe the effect of optically thick clouds and boundary layer aerosols in the retrieval and similar small retrieval biases are achieved with the latest version of SICOR described here. However, in case of an optically thin scattering layer due to an elevated dust layer, optically thin clouds and cirrus above a bright surface, the previous version of SICOR [RD33] could not account for any path enhancement of the observed light due to light trapping between the scattering layer and the surface. In the study of Vidot et al., 2012 [RD33], this shortcoming became clear when assessing the retrieval accuracy for optically thin cirrus above bright surfaces. This is the main reason why the two-stream radiative transfer solver is used in the current algorithm, which approximates both transmission and reflection of a cloud and so allows for photon trapping between optically thin clouds and a bright surface. In the following, our analysis focuses on these new aspects of our algorithm.

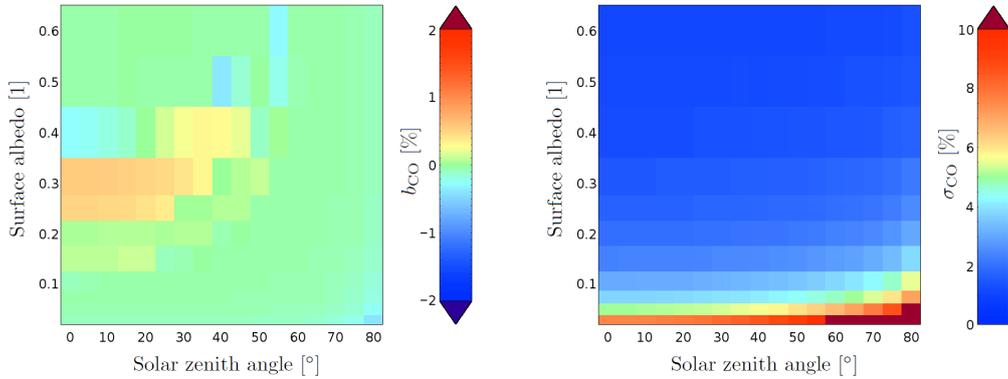


Figure 12: Retrieval bias b_{CO} (left panel) and retrieval noise σ_{CO} (right panel) for the clear sky conditions (without aerosol, clouds and cirrus) and for a viewing zenith angle (VZA) of 0 degree as a function of solar zenith angle (SZA) and surface albedo A_s .

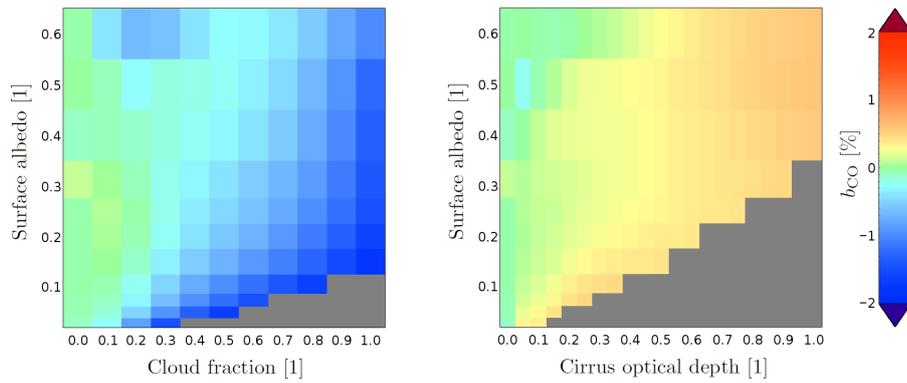


Figure 13: Left panel: Retrieval bias in case of a cloud atmosphere. The CO bias is shown as a function of surface albedo A_s and cloud fraction f for a cloud between 4 and 5 km altitude with optical depth $\tau_{scat} = 2$ and a VZA of 0 degree. Right panel: CO retrieval bias for measurements in presence of optically thin cirrus, which overcasts the entire scene, as a function of surface albedo and cirrus optical depth that defined at 2300 nm. The grey area indicates measurement simulations, which were rejected by the cloud filter.

7.1 Forward model errors

The forward model of our retrieval introduces errors due to the accuracy of the two-stream model, the neglect of atmospheric Rayleigh scattering and the description of clouds and aerosols by a single triangular scattering layer. To elicit the impact of these approximations, we consider three generic measurement ensembles for a clear sky atmosphere and for a cloudy atmosphere with optically thin clouds and cirrus.

Figure 12 shows the CO retrieval bias and the corresponding retrieval noise for simulated clear sky measurements including atmospheric Rayleigh scattering with a variable surface albedo and a variable solar zenith angle. Overall, the retrieval bias is small with $-0.5\% \leq b_{CO} \leq 0.5\%$. The retrieval noise increases from values $< 1\%$ at high sun and for bright surfaces to $\approx 11\%$ for low sun (SZA = 70°) and low albedo ($A_s = 0.03$). This increase is governed by the signal strength and so by the signal-to-noise ratio of the measurement.

To investigate the effect of photon trapping between clouds and the surface, Fig. 13 depicts the CO bias for a cloud between 4 and 5 km altitude with a small optical depth $\tau_{scat} = 2$ as a function of surface albedo and cloud coverage. Here, the CO bias reaches 1.5% with increasing cloud coverage. For a cirrus layer between 9 and 10 km of varying optical depth as function of the surface albedo, the light trapping effect at high surface albedo results in a CO biases $b_{CO} \leq 0.5\%$. Similar small biases are found for an elevated dust layer and optically thin clouds (not shown).

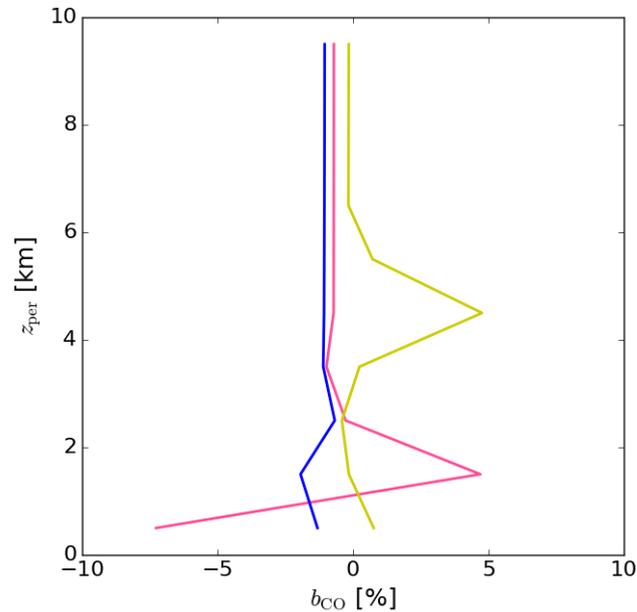


Figure 14: Retrieval bias b_{CO} in cloudy atmospheres in case of strongly enhanced CO concentrations. Measurement simulations are performed for a surface albedo $A_s = 0.05$, SZA and VZA of 50 and 0 degree and for overcast sky with a cloud at 1–2 km altitude with an optical depth of $\tau_{\text{cld}} = 2$ (pink) and $\tau_{\text{cld}} = 5$ (blue). Additionally, we consider a case of partially cloud cover with cloud fraction $f_{\text{cld}} = 0.1$ at 4–5 km altitude with $\tau_{\text{cld}} = 2$ (yellow). The CO profile represents the US standard atmosphere with a perturbation at the indicated altitude z_{per} enhancing the total amount of CO by 50 %.

Moreover, we investigated the implications of the retrieved cloud parameters being effective cloud parameters. These parameters differ from the truth because of the limited information available from the satellite measurements. Here, the retrieval forward model has to describe clouds in a simplified manner with a few free parameters and all remaining cloud properties have to be fixed a priori (see e.g. [RD75, RD76]). In our case, the cloud model includes several simplifications, e.g. a horizontally homogenous cloud with the triangular height distribution in optical depth and a two stream radiative transfer model to describe the cloud radiative properties. Considering the measured radiometric signal as a mean of a photon ensemble with different light paths through the atmosphere, the retrieval adjusts the cloud parameters and the simulated light paths such that the methane absorption features can be fitted by the forward model. This may include erroneous light paths, which effects average out in the simulated measurement for the particular height distribution of methane. However for another trace gas with a different vertical profile, such as CO, the relevance of the individual photons for the observed signal may differ and so the simulated light paths introduce spectral errors in the simulated CO absorption features. Subsequently adjusting the trace gas concentrations in the retrieval, CO biases are introduced for cloudy atmospheres.

Obviously, this retrieval error depends on the particular CO profile and the altitude at which the simulated light path deviates from its truth. So to characterize this inherent bias of our retrieval approach, we simulate SWIR measurements for a cloudy atmosphere adding CO abundance in a 1 km thick, vertically homogenous layer with varying layer top height z_{per} . Here, the CO enhancement increases the CO total column by 50 %. Figure 14 shows the CO biases as a function of z_{per} for scenes covered with low clouds at 1–2 km altitude with optical thicknesses of 2 and 5, and a cloud at 4–5 km covering 10 % of the scene with a cloud optical thickness of 2. In each case, the simulated measurement passes the cloud filter of Sect. 6.5. We clearly see a positive retrieval bias up to 5 % for enhanced CO concentration at the altitude of the optically thin cloud, whereas a negative bias of 7 % is found for low clouds in combination with a near-surface CO enhancements. The latter error is relevant for burning events localized in the tropospheric boundary layer. Above the cloud, the error sensitivity is only small, indicating that the light path at this

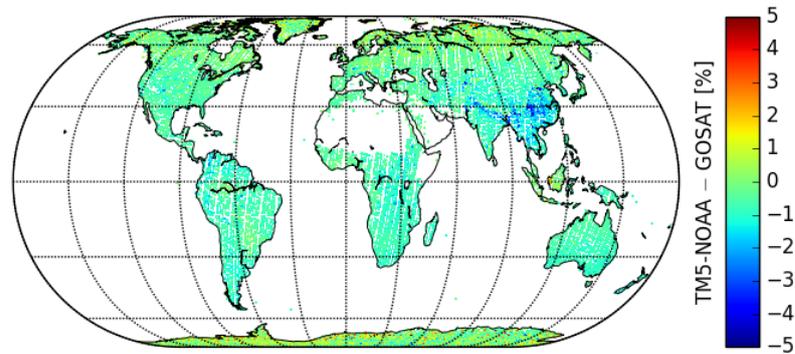


Figure 15: Difference between CH₄ total column dry air mixing ratios from TM5-NOAA simulations and GOSAT retrievals for the period June 2009 to December 2012.

altitude range is well described by our simplified radiative transfer model. Furthermore, for the optically thicker clouds the error sensitivity is below 2%, as expected for a primarily reflecting cloud.

7.2 Atmospheric parameters

An important element of the CO retrieval approach is the use of methane a priori information to determine effective cloud properties from the SWIR measurements as discussed above. The SICOR retrieval relies on simulated CH₄ fields from CAMS. To estimate typical uncertainties of the model forecast, we discuss the TM5 model [RD77] in more detail, which have been used in several studies (e.g. [RD78, RD79, RD80]). Via the inverse modeling technique the sources and sinks of CH₄ in the TM5 model are optimized by minimizing the residual differences between model and measurements from the NOAA-ESRL global monitoring network and deviations from the a priori surface flux distribution [RD81]. In the following, we refer to these model runs as the TM5-NOAA simulations.

To test the overall accuracy of the model simulations, we compare one year of CH₄ model fields with collocated GOSAT observations [RD82, RD83, RD84, RD85]. Here, the GOSAT CH₄ product is extensively validated with TCCON ground measurements with an overall root-mean-square (RMS) difference of 15 ppb and a station-to-station bias of 3.5 ppb [RD86]. Within these boundaries, the GOSAT XCH₄ retrieval can be used to estimate the model accuracy. To this end Fig. 15 shows the difference between GOSAT and TM5-NOAA simulated XCH₄. Over China, the largest biases of up to 3% occur because of inconsistencies in the underlying emission scenario in combination with a limited regional coverage of the NOAA-ESRL ground-based measurements. Overall biases are smaller with an RMS difference between GOSAT and TM5-NOAA, amounting to 20 ppb and increasing towards southern latitudes. This latitudinal bias in TM5, relative to GOSAT, is found also in other models (see e.g. [RD87]) and is currently under further investigation. Comparisons of the modeled CH₄ columns with collocated TCCON measurements are largely consistent with these findings with an RMS difference between 8 and 22 ppb depending on the TCCON site.

Inherent to this analysis is the assumption that the NOAA-ESRL measurements are available timely to perform model simulation as input to the retrieval. This timeliness of the simulation needs further consideration. Commonly, inverse modeling derived estimates lag behind real-time by approximately one year. This is mostly due to the availability of

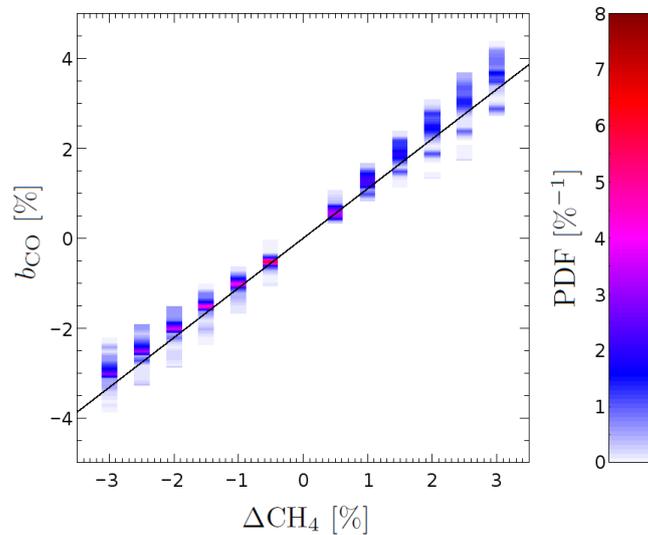


Figure 16: CO bias due to a priori errors in CH₄ for the clear sky measurement ensemble of Fig. 12. For each CH₄ error, the CO bias probability function is shown. The CO error sensitivity is estimated by a linear regression through all data points (solid line).

various types of inputs that are required, including meteorological fields, a priori emission estimates, and measurements. Due to that, we propose a modeling procedure that uses the inversion-optimized TM5 estimates of the dry air mole column mixing ratio of methane XCH₄ of the previous year. Obviously, the largest error source is the variability in XCH₄ caused by the year-to-year variations in meteorology and the inter-annual variability of the methane sources and sinks. We estimate the size of the error from results of a multi-year inversion for the period 2003-2010, calculating how XCH₄ on a given day of the year (15th of January, April, July and October) varied between the years. Largest variations are found over South East Asia, due to large regional sources of methane, but also errors in the meteorology of the northern and southern hemispheric storm tracks are present. On average, the 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. Acknowledging these limitations in our approach, an uncertainty of 3 % of our methane a priori knowledge seems a reasonable margin that should be achievable for most conditions encountered throughout the global domain.

For the generic clear sky measurement ensembles, Fig. 16 shows the probability distribution function (PDF) of the CO biases as a function of the methane model error between $\pm 3\%$. A linear regression through the data points indicates a nearly one-to-one error correspondence with 1.11 % CO bias due to 1 % error in the methane model columns. Table 10 provides the corresponding bias sensitivity for the cloudy and cirrus measurement ensembles in Fig. 13. Aggregating these results, we conclude that the CO retrieval bias due to the uncertainty of the TM5-NOAA model input typically does not exceed 3%.

Additionally to the CH₄ a priori error, an erroneous surface pressure affects the inferred CO column both through a wrong conversion of the methane mixing ratio XCH₄ into the total column density of methane and via an erroneous spectroscopy because of the pressure broadening of individual absorption lines. For the operational retrieval, we use pressure information from the European Centre for Medium-Range Weather Forecast (ECMWF) with a typical accuracy of 2–3 hPa [RD88]. Subsequently, ECMWF surface pressure is interpolated on the particular S5 pixel by means of the digital elevation map of [RD89] and [RD37] accounting for the topography of the terrain. For pressure uncertainties in the range ≤ 3 hPa, we obtain an error sensitivity of 0.11–0.13 % CO column error per 1 hPa surface pressure error for the clear sky and cloudy scenarios of our generic measurement ensemble. Furthermore, we evaluated the impact of uncertainties in the atmospheric temperature forecast of ECMWF, which has been estimated at a few Kelvin. Table 10 lists the CO retrieval sensitivities with respect to an offset of the atmospheric temperature profile in the range ± 3 K, which vary between 0.17 and 0.23 % CO column error per 1 K temperature offset. Thus for the CO column product, we expect the corresponding retrieval biases due to inaccuracies in the atmospheric parameters to be well within 1%.

	S5L2PP	Reference : SRON-ESA-S5L2PP-ATBD-002	
	Carbon Monoxide ATBD	Version : 3.1	Page
		Date : 17 May 2019	52/74

Table 10: CO column retrieval sensitivity in % with respect to knowledge uncertainties of a set of atmospheric and instrument parameters for the generic clear sky, cloud and cirrus ensemble: (1) CH₄ a priori uncertainty of TM5-NOAA runs, (2) ECMWF surface pressure uncertainty, (3) ECMWF temperature profile offset, (5) FWHM uncertainty of the ISRF, (6) spectral calibration error δs and (7) the radiometric offset I_{offset} and a multiplicative radiometric error I_{scal} of the level 1 data product.

#	parameter	clear sky	cloud	cirrus
1	CH ₄ a priori [%/%]	1.11	1.18	1.21
2	pressure [%/hPa]	0.11	0.13	0.13
3	temperature [%/K]	0.23	0.17	0.20
4	FWHM [%/%]	0.51	0.40	0.43
5	δs [%/10 pm]	0.88	0.87	0.87
6	I_{offset} [%/%]	-0.63	-0.47	-0.46
7	I_{scal} [%/%]	0.01	0.01	0.02

7.3 Instrument effects

Finally, we studied the CO retrieval sensitivity with respect to a set of instrument related parameters. First, the Earth radiance spectrum may be subject to a radiometric offset I_{offset} , expressed relative to the radiance level at the reference wavelength of 2315 nm, or a spectrally constant multiplicative error δI_{scal} . Instrumental reasons for these errors can be manifold, e.g. uncorrected stray light, detector and read-out electronics performance and an erroneous pre-flight instrument calibration. For the generic ensembles, we derived an error sensitivity of -0.47 to -0.63 % CO column error per percent radiometric offset and 0.01 to 0.02 % per percent multiplicative radiometric error. The main reason for this robust CO retrieval performance with respect to this type of radiometric errors is the selected spectral window with relatively weak atmospheric absorption. Here, these spectral biases can be mitigated efficiently by the retrieval of an effective surface albedo and cloud properties.

To study an erroneous spectral calibration of the measurement, we assumed a correct instrument calibration λ_i of spectral detector i and an erroneous calibration

$$\lambda'_i = \lambda_i + \frac{\lambda_i - \lambda_m}{\lambda_r - \lambda_m} \delta s. \quad (93)$$

Here, $\lambda_r = 2385$ nm indicates the longwave edge of the SWIR band and $\lambda_m = 2345$ nm is the spectral centre. So, δs characterizes the spectral calibration errors at the edges of the SWIR spectral range whereas in the centre λ_m the calibration error vanishes. The corresponding spectral squeeze for the CO fit windows (2315–2338 nm) is about one third of δs . The error sensitivity of the CO column product is about 0.9 % per $\delta s = 10$ pm. Due to the required knowledge of the centre of all SWIR channels of < 2 pm [AD2], this CO error sensitivity is not critical for a compliant instrument. Moreover, the CO retrieval has no error sensitivity to an overall offset of the spectral calibration because this parameter is adjusted by the retrieval.

Errors in the instrument spectral response function can be manifold and are hard to quantify in a general manner. In this study, we restricted ourself to an erroneous full width at half maximum (FWHM) of the instrument spectral response function (ISRF), which may occur e.g. because of pre-flight instrument calibration errors or because of fluctuations of the instrument temperature. Table 10 shows the ISRF retrieval sensitivity of about 0.5 % CO error for a 1 % FWHM uncertainty of the ISRF, which is within the knowledge requirement for the S5 instrument calibration [AD2].

7.4 Sentinel 5 orbit ensemble

TBD

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002
		Version : 3.1 Page
		Date : 17 May 2019 53/74

8 Validation

The validation of the CO column product depends on whether or not information on the vertical profile of CO is available. If so, we can refer to Eq. (73) where the independent measurement of the CO profile \mathbf{x}_{est} is considered as an estimate of the true profile \mathbf{x}_{true} . Thus, the retrieval can be verified considering the difference

$$\delta c_{\text{val}} = \hat{c} - \mathbf{A}_c \mathbf{x}_{\text{est}} . \quad (94)$$

This approach is beneficial for the validation because no model information is used to fill-up the null space contribution of the retrieval. Unfortunately, only little measurements of the vertical profile of CO will be available and so other means are needed to validate the CO product. The two measurement networks TCCON and NDACC-IRWG of ground-based solar Fourier-transform spectrometers exist, which provide regular measurements of the vertical integrated column amounts of a number of trace gases, including CO, for several sites. The column product c_{est} of these networks is an estimate of the true column and so we can only compare it with the retrieved column \hat{c} after filling up the null-space. For this purpose, we will use CO profiles \mathbf{x}_{ctm} simulated with a state-of-the-art chemical transport model. Thus the error term to be validated is given by

$$\delta c_{\text{val}} = \hat{c} + (\mathbf{I} - \mathbf{A}_c) \mathbf{x}_{\text{ctm}} - c_{\text{est}} . \quad (95)$$

Generally, to evaluate δc_{val} in Eqs. (94) and (95), the statistical error, pseudo-statistical errors and systematic errors on predefined spatial and temporal scales have to be considered. This requires sufficient validation measurements and hence, in this perspective, the ground based networks are considered to provide the primary data sets for validation. In the following paragraphs, we describe shortly the different data sets that, to our knowledge, are presently available for the CO validation.

TCCON

The Total Carbon Column Observing Network (TCCON, <https://tcon-wiki.caltech.edu>) is a network of ground-based Fourier Transform Spectrometers recording direct solar spectra in the near-infrared spectral region in the spectral range between 0.7 and 2,5 μm . From these spectra, accurate and precise column-averaged abundance of CO_2 , CH_4 , N_2O , HF, CO, H_2O , and HDO are retrieved [RD90]. By default, CO is determined from a spectral window around 2.3 μm with a precision of 0.5 ppb and an accuracy of 3 ppb. The TCCON began in 2004 with the installation of the first instrument in Park Falls, Wisconsin, USA, and has since grown to 20 operational instruments worldwide.

TCCON data is currently guaranteed to be made available within one year after measurement. Only for a few stations earlier delivery is guaranteed. It should be realised, however, that one year of data is needed to assess the quality of the satellite data. In practice, this would mean that the S5 data product can be earliest validated after two years into the mission, which is clearly not adequate. Therefore, we recommend that for an adequate and timely validation of S5P CO, TCCON data should be made available on shorter timescales, e.g. 3 months.

IRGW

The Infrared Working Group (IRWG, <http://www.acd.ucar.edu/irwg/>) represents a similar network of infrared solar Fourier-transform spectrometers that is part of the Network for the Detection of Atmospheric Composition Change (NDACC). It is a multi-national collection of over twenty high-resolution spectrometers that regularly record the atmospheric absorption spectrum from sites distributed from pole to pole. Solar absorption spectra are used to retrieve concentrations of a number of the gaseous atmospheric components, including: O_3 , HNO_3 , HCl, HF, CO, N_2O , CH_4 , HCN, C_2H_6 , H_2O , HDO and ClONO_2 . In contrast to the TCCON network, direct solar measurements are performed at longer wavelengths of the solar spectrum, e.g. for the retrieval of CO, three standard fit windows are employed between 4.8-4.9 μm [RD91, RD92]. The accuracy of the CO total column of this data source is about 5 % (tbc). It should be noted that in terms of geo-location the NDACC and TCCON network have a large overlap.

 	<p style="text-align: center;">S5L2PP Carbon Monoxide ATBD</p>	<p>Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 54/74</p>
--	--	---

IAGOS MOZAIC

As part of the European Research Infrastructure IAGOS-ERI (<http://www.iagos.fr/web/>), the MOZAIC program provides airborne in-situ measurements for O₃, H₂O, CO, and total nitrogen oxides (NO_y) since August 1994. Measurements are geo-localised (latitude, longitude and pressure) and come with meteorological observations (wind direction and force, temperature). Data acquisition is automatically performed during round-trip international flights (ascent, descent and cruise phases) from Europe to America, Africa, Middle East, and Asia. During ascent and descent, trace gas height information is recorded which is of particular interest to validate the S5P CO column product. However, care should be taken in case of strong spatial gradients of CO as the TROPOMI and MOZAIC spatial sampling will not be exactly coaligned.

AIRCORE

The AirCore ([RD93] and <http://www.esrl.noaa.gov/gmd/ccgg/aircraft/aircore.html>) is an innovative atmospheric sampling system that consists of a long coil of stainless steel tubing. It is a recent development with great potential for the validation of satellite observations. AirCore ascends on a helium balloon and fills with surrounding atmosphere during a parachute-controlled descent, collecting a sample from balloon burst (up to 30 km) down to ground level. An AirCore sample can be analysed in the laboratory for concentrations of trace atmospheric gases. The length of the tubing and short time to analysis minimises mixing inside the tubing, so that each AirCore sample provides up to 100 measurements of CO₂, CH₄, and CO from top altitude to ground level. These measurements will be calibrated on the World Meteorological Organization scales (expected within 0.05% for CO₂ and CH₄, 5% for CO) and has the potential to provide a ground-truth standard for comparison with total column measurements from either ground-based Fourier Transfer Spectrometers or satellites.

	S5L2PP	Reference : SRON-ESA-S5L2PP-ATBD-002
	Carbon Monoxide ATBD	Version : 3.1 Page
		Date : 17 May 2019 55/74

Table 11: Calculation settings and computation time for the S5 orbit ensemble.

	Methane cloud filter	SICOR CO full- physics retrieval
Internal sampling (cm ⁻¹)	0.05	0.03
Convergence $\Delta\chi^2$	0.5	0.5
Minimum iterations	5	10
Maximum iterations	15	15
Time per spectrum (s)	0.03147	0.16875
Time with gfortran (s)	0.03734 (+18.65%)	0.24695 (+46.34%)

9 Feasibility

The SICOR algorithm is already implemented in several processing frameworks. Recently, Borsdorff et al. [RD30, RD31] have used SICOR to process the full-mission SCIAMACHY CO data product. Moreover, it is implemented in the Sentinel 5 Precursor operational framework as well as in the SRON scientific data processing framework using a GRID processing infrastructure, which is a distributed infrastructure of compute clusters and storage systems. During the preparation phase of the S5P mission, both framework implementations were successfully tested, where processing of 300 S5-P test orbits on 1000 cores of the SURFsara GRID [RD94] showed good scalability of the software for the employed infrastructure. Here MPI is used for parallelization on single processing nodes. Overall, the SICOR heritage demonstrates the readiness level and maturity of the algorithm and software implementation. The SICOR software is written in Fortran 90 with a well-defined interface to a framework written in C.

9.1 Estimated Computational Effort

To evaluate the computation cost of the SICOR-S5 algorithm, we assume data flow of 700000 spectra per orbit, lasting 6090 seconds. To estimate the computation time of the retrievals, we consider the test ensemble as described in Sec. 7.4. The local computer is a HP Z230 SFF D1P35AV with a Intel(R) Xeon(R) CPU E3-1226 v3 @ 3.30GHz processor and a cache size of 8192 KB. The hardware computer benchmark for floating point performance is about SPECfp2006 = 60, which is estimated from the specification of the similar processor Intel(R) Xeon(R) CPU E3-1230 v3 with SPECfp2006 = 72.9 scaled by a factor 0.83 to account for the performance difference between both processors (see also <https://www.spec.org/cpu2006/results/res2014q2/cpu2006-20140407-29279.pdf>). Executables were generated with the Intel FORTRAN compiler with optimization flag O3. Table 11 provides also the computation time using the GNU gfortran compiler with optimization flag O3. Overall, the mean computation time of the non-scattering methane filter is 0.03147 seconds and the mean computation time of the SICOR CO scattering retrieval is 0.16875 seconds (excluding file I/O).

For the test ensemble, 28 % of the observations were filtered out by the methane cloud filter with a threshold of $\delta\text{CH}_4 < 25\%$ and so 72 % of the data requires full processing. In contrast, for one year of GOSAT observation, which is already considered in Fig. 5, 80.6 % of the non-scattering retrievals passes the cloud filter (79.7 % for ocean pixels and 82.4 % for land pixels). Figure 17 shows the corresponding global distribution of the fraction of data to be processed. This discrepancy might be explained by the fact that the test ensemble represents an atmospheric snapshot and so cannot be used for an overall performance estimate for Sentinel-5. Thus to estimate of the computational effort, we assume that a scattering retrieval has to be performed for 80.6 % of all data. This means a mean computational time of $0.03147 + 0.806 \cdot 0.16875 = 0.1675$ seconds per spectrum and so, to handle the data stream by a continuous data processing, in total 20 cores of the described hardware are needed. Concerning the software implementation, we expect no significant improvement of the algorithm runtime performance in the near future.

9.2 Inputs

The SICOR input and output is already discussed in detail in Sec. 6.9 and is summarized in the Tables 12 to 15. We distinguish between static and dynamic fields, where static fields do not change in value and/or dimension during the

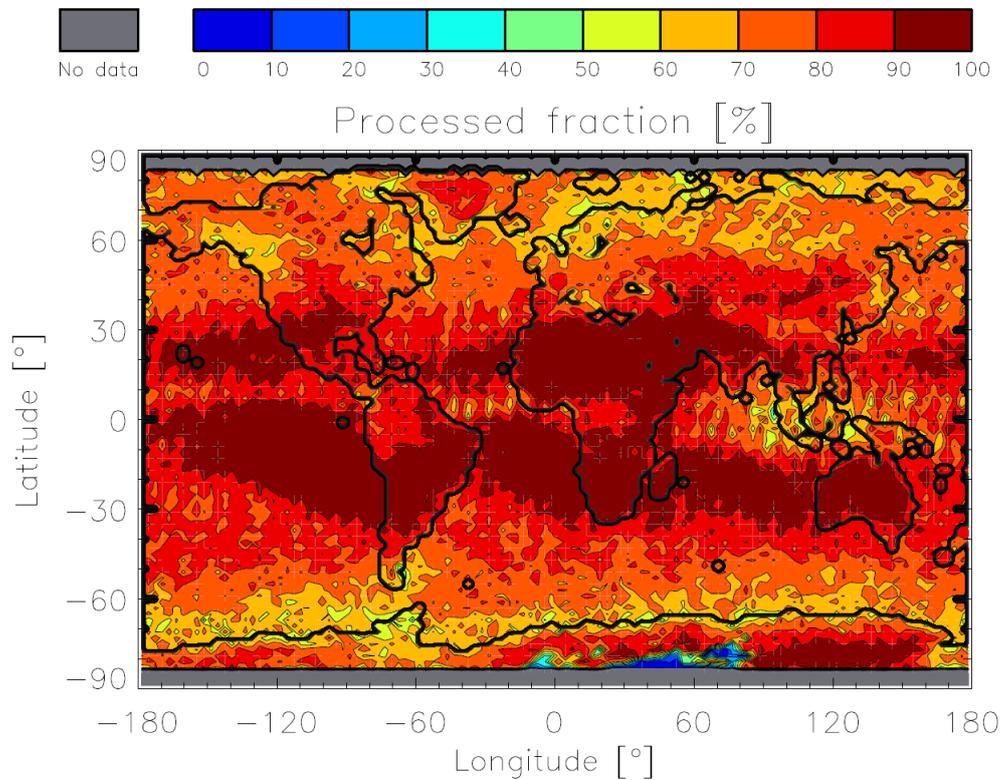


Figure 17: Fraction of GOSAT non-scattering retrievals which are accepted by the cloud filter $\Delta\text{CH}_4 < 25\%$. The analysis is based on one year (2010) of GOSAT non-scattering retrievals (RemoTeC V2.1).

processing of a data granule, whereas dynamic fields have different values and/or dimension for the individual ground pixel processing. It is important to realize that the SICOR-S5 implementation relies only on SWIR-3 level 1B data and so inter-band coregistration is not applicable here. Hence, we assumed that no spatial resampling of the measurement is applied to the L1B data before processing. The algorithm specific settings are provided via a cumulative string, which is unraveled by the RemoTeC software. Appendix B gives an overview of all relevant setting parameters.

9.2.1 Static Inputs

Table 12: Static input data of SICOR-S5.

Parameter	Symbol	Dimension	Physical Units	Source	M/O ^a	C/F/U ^b
Length of algorithm settings	CFG- <i>N</i>	1	-	S5 L2 CH4 CFG	M	U
Algorithm settings	CFG- <i>S</i>	CFG- <i>N</i>	-	S5 L2 CH4 CFG	M	U
Level of output	FWK- <i>L</i>	1	-	S5 L2 CH4 CFG	O	U
Number of swath positions	N_{sp}	1	-	S5 L2 AUX ISRF	M	U
Number of wavelength pixels SWIR-3	N_{SW3}	1	-	S5 L1B irradiance product	M	U
Irradiance wavelength SWIR-3	IR- λ_{SW3}	$N_{sp} \times N_{SW3}$	nm	S5 L1B irradiance product	M	U
Irradiance spectrum SWIR-3	IR- I_{SW3}	$N_{sp} \times N_{SW3}$	mol/(m ² s nm)	S5 L1B irradiance product	M	U
Irradiance noise spectrum SWIR-3	IR- e_{SW3}	$N_{sp} \times N_{SW3}$	mol/(m ² s nm)	S5 L1B irradiance product	M	U
Pixel mask of irradiance spectrum SWIR-3	IR- p_{SW3}	$N_{sp} \times N_{SW3}$	-	S5 L1B irradiance product	M	U
Number of wavelength difference in ISRF definition SWIR-3	ISRF- $N_{\delta\lambda,SW3}$	1	-	S5 L2 AUX ISRF	M	U
Wavelength difference in ISRF definition SWIR-3	ISRF- $\delta\lambda_{SW3}$	ISRF- $N_{\delta\lambda,SW3}$	nm	S5 L2 AUX ISRF	M	U
Instrument spectral response function SWIR-3	ISRF- S_{SW3}	$N_{sp} \times$ ISRF- $N_{\lambda,SW3}$ \times ISRF- $N_{\delta\lambda,SW3}$	nm ⁻¹	S5 L2 AUX ISRF	M	U
Number of wavelengths reference solar spectrum	SUN- <i>N</i>	1	-	S5 L2 CH4 LUT	M	U
Wavelengths reference solar spectrum	SUN- λ	SUN- <i>N</i>	nm	S5 L2 CH4 LUT	M	U
Reference solar spectrum	SUN- <i>I</i>	SUN- <i>N</i>	mol/(m ² s nm)	S5 L2 CH4 LUT	M	U
Number of hybrid layers of meteorological input	ATM- N_{met}	1	-	S5 L2 AUX product (CAMS)	M	U
Hybrid pressure coefficient A at layer interface for meteorological input	ATM- A_{met}	ATM- $N_{met} + 1$	Pa	S5 L2 AUX product (CAMS)	M	U
Hybrid pressure coefficient B at layer interface for meteorological input	ATM- B_{met}	ATM- $N_{met} + 1$	1	S5 L2 AUX product (CAMS)	M	U
Number of hybrid layers of CH ₄ input	ATM- N_{CH_4}	1	-	S5 L2 AUX product (CAMS)	M	U
Hybrid pressure coefficient A at layer interface for CH ₄ input	ATM- A_{CH_4}	ATM- $N_{CH_4} + 1$	Pa	S5 L2 AUX product (CAMS)	M	U

Continued on next page

Table 12 – Static input - Continued from previous page

Parameter	Symbol	Dimension	Physical Unit	Source	M/O ^a	C/F/U ^b
Hybrid pressure coefficient B at layer interface for CH ₄ input	ATM- B_{CH_4}	ATM- $N_{CH_4} + 1$	1	S5 L2 AUX product (CAMS)	M	U
Number of hybrid layers of CO input	ATM- N_{CO}	1	-	S5 L2 AUX product (CAMS)	M	U
Hybrid pressure coefficient A at layer interface for CO input	ATM- A_{CO}	ATM- $N_{CO} + 1$	Pa	S5 L2 AUX product (CAMS)	M	U
Hybrid pressure coefficient B at layer interface for CO input	ATM- B_{CO}	ATM- $N_{CO} + 1$	1	S5 L2 AUX product (CAMS)	M	U
Number of pressures cross sections	XS- N_p	1	-	S5 L2 CH4 LUT	M	U
Number of temperatures cross sections	XS- N_T	1	-	S5 L2 CH4 LUT	M	U
Pressures cross sections	XS- p	XS- N_p	Pa	S5 L2 CH4 LUT	M	U
Temperatures cross sections	XS- T	XS- $N_T \times XS-N_p$	K	S5 L2 CH4 LUT	M	U
Number of SWIR-3 wavenumbers cross sections	XS- $N_{v_{SW3}}$	1	-	S5 L2 CH4 LUT	M	U
SWIR-3 wavenumbers cross sections	XS- v_{SW3}	XS- $N_{v_{SW3}}$	cm ⁻¹	S5 L2 CH4 LUT	M	U
Cross sections HDO SWIR-3	XS- $\sigma_{HDO,SW3}$	XS- $N_p \times XS-N_T \times XS-N_{v_{SW3}}$	cm ²	S5 L2 CH4 LUT	M	U
Cross sections H ₂ O SWIR-3	XS- $\sigma_{H_2O,SW3}$	XS- $N_p \times XS-N_T \times XS-N_{v_{SW3}}$	cm ²	S5 L2 CH4 LUT	M	U
Cross sections CH ₄ SWIR-3	XS- $\sigma_{CH_4,SW3}$	XS- $N_p \times XS-N_T \times XS-N_{v_{SW3}}$	cm ²	S5 L2 CH4 LUT	M	U
Cross sections CO SWIR-3	XS- $\sigma_{CO,SW3}$	XS- $N_p \times XS-N_T \times XS-N_{v_{SW3}}$	cm ²	S5 L2 CH4 LUT	M	U

^a M mandatory input, O optional input.

^b C copy to output, F use for filtering U use for processing.

^c L1B-IRR means Level 1b irradiance product.

The static input contains the algorithm configuration parameters as character string. This contains the contents of a settings file, which will be interpreted by the algorithm. The content of these settings is explained in Appendix B.

9.2.2 Dynamic Inputs

Table 13: Dynamic input of SICOR.

Parameter	Symbol	Dimension	Physical Unit	Source	M/O ^a	C/F/U ^b
Framework processing quality flags	FWK-PQF _{in}	1	-	S5 L2 CH4 CFG	O	C
Swath position index of the sounding	i_{sp}	1	-	S5 L1B radiance product (band 5)	M	U
Radiance wavelength SWIR-3	RA- λ_{SW3}	IR- N_{SW3}	nm	S5 L1B radiance product (band 5)	M	U
Radiance spectrum SWIR-3	RA- I_{SW3}	IR- N_{SW3}	mol/(m ² s sr nm)	S5 L1B radiance product (band 5)	M	U
Radiance noise spectrum SWIR-3	RA- e_{SW3}	IR- N_{SW3}	mol/(m ² s sr nm)	S5 L1B radiance product (band 5)	M	U
Pixel mask of radiance spectrum SWIR-3	RA- p_{SW3}	IR- N_{SW3}	-	S5 L1B radiance product (band 5)	M	U
Solar zenith angle	GEO- θ_0	1	°	S5 L1B radiance product (band 5)	M	U
Solar azimuth angle	GEO- φ_0	1	°	S5 L1B radiance product (band 5)	M	U
Viewing zenith angle	GEO- θ_v	1	°	S5 L1B radiance product (band 5)	M	U
Viewing azimuth angle	GEO- φ_v	1	°	S5 L1B radiance product (band 5)	M	U
Latitude pixel center	GEO- Θ_{center}	1	°	S5 L1B radiance product (band 5)	M	C
Longitude pixel center	GEO- Φ_{center}	1	°	S5 L1B radiance product (band 5)	M	C
Latitude pixel corners	GEO- Θ_{corner}	4	°	S5 L1B radiance product (band 5)	M	C
Longitude pixel corners	GEO- Φ_{corner}	4	°	S5 L1B radiance product (band 5)	M	C
Time of measurement	GEO- t	7	TMDHMSM ^c	S5 L1B radiance product	M	C
Mean surface elevation of pixel	GEO- z_{surf}	1	m	S5 L2 AUX product (GMTED2010)	M	U
Standard deviation of surface elevation	GEO- $\sigma_{z_{surf}}$	1	m	S5 L2 AUX product (GMTED2010)	M	C
Surface classification	GEO- f_{surf}	1	-	S5 L2 AUX product (GMTED2010)	M	C
Surface pressure meteorological model	ATM- $p_{surf,met}$	1	Pa	S5 L2 AUX product (CAMS)	M	U
Surface elevation model	ATM- z_{surf}	1	m	S5 L2 AUX product (ECMWF)	M	U
Specific humidity profile	ATM- q	ATM- N_{met}	kg/kg	S5 L2 AUX product (ECMWF)	M	U
Temperature profile	ATM- T	ATM- N_{met}	K	S5 L2 AUX product (ECMWF)	M	U
CH ₄ mixing ratio profile	ATM- r_{CH_4}	ATM- N_{CH_4}	1	S5 L2 AUX product (CAMS)	M	U
CO mixing ratio profile	ATM- r_{CO}	ATM- N_{CO}	1	S5 L2 AUX product (CAMS)	M	U
Liquid cloud fraction	MET- f_{liq}	1	1	S5 L2 CLA	O	U
Liquid cloud optical depth	MET- r_{liq}	1	1	S5 L2 CLA	O	U
Liquid cloud top pressure	MET- p_{liq}	1	Pa	S5 L2 CLA	O	U

^a M mandatory input, O optional input.

^b C copy to output, F use for filtering U use for processing.

^c Year, month, day, hour, minute, second, millisecond.

SICOR-S5 only relies on the SWIR-3 L1B product as measurement input and so no spatial coregistration is required between different spectral bands. There, SICOR request the SWIR-3 L1B radiance product not corrected for spatial miss-registration between channels. The METimage data are classified as optional input. They are used to initialize the inversion and to support cloud filtering. In case the data are not available, the retrieval will be initialized by fixed a priori values and the cloud filtering will purely rely on the non-scattering CH₄ retrieval as described in Sect. 6.5.

 S5L2PP	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Date : 17 May 2019	Page 60/74
--	---------------------------------------	---	---------------

	S5L2PP	Reference : SRON-ESA-S5L2PP-ATBD-002
	Carbon Monoxide ATBD	Version : 3.1 Page
		Date : 17 May 2019 61/74

9.3 Outputs

9.3.1 Static output

Table 14: Output data of the static initialization.

Parameter	Symbol	Dimension	Physical Unit	Range	Destination
String length of SICOR version	VER- N	1	-	1 – 10	S5 L2 CO product
SICOR version	VER- S	VER- N	-	NA	S5 L2 CO product
Number of vertical layers	MDL- N	1	-	1 – 40	S5 L2 CO product
SWIR-3 spectral shift sun	PH- $\delta\lambda_{s,SW3}$	N_{sp}	nm	$\pm 0 - 0.02$	S5 L2 CO product

9.3.2 Dynamic output

Table 15: Dynamic output data of the data processor per ground pixel.

Parameter	Symbol	Dimension	Physical Unit	Range	Destination
Pressure grid	MDL- p	MDL- N	Pa	$0 - 1.1 \cdot 10^5$	S5 L2 CO product
altitude grid	MDL- z	MDL- N	m	$0 - 1.0 \cdot 10^5$	S5 L2 CO product
Dry air column	AP- C_{air}	1	mol / m ²	$3 \cdot 10^5 - 4 \cdot 10^5$	S5 L2 CO product
A-priori CH ₄ profile	AP- c_{CH_4}	MDL- N	mol / m ²	0 – 0.1	S5 L2 CO product
A-priori CO profile	AP- c_{CO}	MDL- N	mol / m ²	0 – 0.01	S5 L2 CO product
Non-scattering CH ₄ column	NS- C_{CH_4}	1	mol / m ²	0 – 1	S5 L2 CO product
CO column	PH- C_{CO}	1	mol / m ²	0 – 0.1	S5 L2 CO product
CO column precision	PH- e_{CO}	1	mol / m ²	0 – 0.1	S5 L2 CO product
CO column averaging kernel	PH- a_{CO}	MDL- N	1	0 – 3	S5 L2 CO product
H ₂ O column	PH- C_{H_2O}	1	mol / m ²	$0 - 1 \cdot 10^4$	S5 L2 CO product
HDO column	PH- C_{HDO}	1	mol / m ²	$0 - 1 \cdot 10^2$	S5 L2 CO product
SWIR-3 surface albedo	PH- A_{SW3}	1	1	0 – 1	S5 L2 CO product
Effective cloud optical thickness SWIR-3	PH- $\tau_{cl,SW3}$	1	1	0 – 5	S5 L2 CO product
Effective cloud layer height	PH- z_{cl}	1	m	0 – 15000	S5 L2 CO product
SWIR-3 spectral shift	PH- $\delta\lambda_{e,SW3}$	1	nm	$\pm 0 - 0.02$	S5 L2 CO product
Number of iterations	PH- N_{iter}	1	-	0 – 30	S5 L2 CO product
Total reduced χ^2 of the spectral fit	PH- χ^2	1	-	tbd	S5 L2 CO product
Total degrees of freedom for signal	PH-DFS	1	-	tbd	S5 L2 CO product
Number of spectral pixels in SWIR-3 fitting window	PH- N_{SW3}	1	-	$\leq IR-N_{SW3}$	S5 L2 CO product
Processing quality flag	FWK-PQF	1	-	any	S5 L2 CO product

The processing quality flag (PQF) is an important element of the dynamic algorithm output, which indicates the status, errors, data filtering and additional warnings to evaluate the retrieval quality. The PQFs are summarized in Tab. 16.

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Date : 17 May 2019	Page 62/74

Table 16: Processing Quality Flags (PQFs) of SICOR. Four types of flags are provided: status (S), error (E), filter (F) and warning (W)

Short name	Type	Description
successful_retrieval	S	No failures, output contains value. Warnings still possible.
ler_range_filter	F	Lambert-equivalent reflectivity is too low.
sza_range_filter	F	Solar zenith angle out of range.
cloud_filter	F	Scene flagged as cloudy by internal cloud filter protocol.
filter_retrieval_failure	E	A retrieval used for filtering failed.
numerical_error	E	General fatal numerical error occurred during inversion.
convergence_error	E	Inversion did not converge.
wavelength_calibration_warning	W	Offset from wavelength fit is larger than threshold from configuration.
boundary_hit_warning	W	Non-fatal boundary hit during iterations.
pixel_level_input_data_missing	W	Dynamic auxiliary input data (e.g.. METimage) is missing for this ground pixel. A fallback option is used.
altitude_consistency_warning	W	Large, but not too large, difference between model altitude and scene altitude value.
deconvolution_warning	W	Failed deconvolution irradiance spectrum (specific for detector spatial channel index).
data_range_warning	W	Non-fatal boundary hit during iterations.
extrapolation_warning	W	Lookup table needed to be extrapolated during iterations.

9.3.3 Breakpoint Output

Table 17: Breakpoint output of SICOR.

Parameter	Symbol	Dimension	Physical Unit	Range	Destination
wavelengths SWIR-3	RES- λ_{SW3}	N_{SW3}	nm	2300 – 2400	S5 L2 CO product
measured spectrum SWIR-3	RES- $y_{meas,SW3}$	N_{SW3}	$\text{mol}/(\text{m}^2 \text{ s sr nm})$	$0 - 10^{-6}$	S5 L2 CO product
noise spectrum SWIR-3	RES- $\ell_{meas,SW3}$	N_{SW3}	$\text{mol}/(\text{m}^2 \text{ s sr nm})$	$0 - 10^{-8}$	S5 L2 CO product
modeled spectrum SWIR-3	RES- $y_{mod,SW3}$	N_{SW3}	$\text{mol}/(\text{m}^2 \text{ s sr nm})$	$0 - 10^{-6}$	S5 L2 CO product

Table 17 lists additional optional output that is meant for investigation on the retrieval process. In the default run, this output is not written out.

Appendices

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

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Date : 17 May 2019	Page 63/74

$$\frac{dF^\uparrow}{d\tau} = \alpha_{1,n}F^\uparrow - \alpha_{2,n}F^\downarrow - \alpha_{3,n}\frac{S}{\mu_0} \quad (96)$$

$$\frac{dF^\downarrow}{d\tau} = \alpha_{2,n}F^\uparrow - \alpha_{1,n}F^\downarrow - \alpha_{4,n}\frac{S}{\mu_0} \quad (97)$$

$$\frac{dS}{d\tau} = -(1 - \omega_n f_n)\frac{S}{\mu_0} \quad (98)$$

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

$$\alpha_{1,n} = U_\uparrow(1 - \omega_n(1 - \bar{\beta}_n)) \quad (99)$$

$$\alpha_{2,n} = U_\downarrow\bar{\beta}_n\omega_n \quad (100)$$

$$\alpha_{3,n} = (1 - f_n)\omega_n\beta_n(\mu_0) \quad (101)$$

$$\alpha_{4,n} = (1 - f_n)\omega_n(1 - \beta_n(\mu_0)) \quad (102)$$

with the fraction of radiation contained in the forward peak f_n , the single scattering albedo ω_n , the fractional mean backward scattering coefficient of diffuse light $\bar{\beta}_n$, the backward scattering coefficient of primary scattered solar radiation $\beta_n(\mu_0)$, and the diffusivity factors of upward and downward radiation $U_{\downarrow\uparrow}$, respectively.

For a N-layer atmosphere, the general solution of the two-stream model of Zdunkowski et al. [RD47] can be expressed by a linear combination of the internal boundary conditions

$$\begin{pmatrix} S_n \\ F_n^\downarrow \\ F_{n-1}^\uparrow \end{pmatrix} = \begin{pmatrix} a_{1,n} & 0 & 0 \\ a_{2,n} & a_{4,n} & a_{5,n} \\ a_{3,n} & a_{5,n} & a_{4,n} \end{pmatrix} \begin{pmatrix} S_{n-1} \\ F_{n-1}^\downarrow \\ F_n^\uparrow \end{pmatrix}, \quad (103)$$

see also Eq. (30) with coefficients

$$a_{1,n} = \exp\left[-\frac{(1 - \tilde{\omega}_n f_n)\Delta\tau_n}{\mu_0}\right] \quad (104)$$

$$a_{2,n} = -a_{4,n}\gamma_{2,n} - a_{5,n}\gamma_{1,n}a_{1,n} + \gamma_{2,n}a_{1,n} \quad (105)$$

$$a_{3,n} = -a_{5,n}\gamma_{2,n} - a_{4,n}\gamma_{1,n}a_{1,n} + \gamma_{1,n} \quad (106)$$

$$a_{4,n} = E_n \frac{1 - M_n^2}{1 - E_n^2 M_n^2} \quad (107)$$

$$a_{5,n} = M_n \frac{1 - E_n^2}{1 - E_n^2 M_n^2} \quad (108)$$

and with

$$E_n = \exp(-\epsilon_n \Delta\tau_n); \quad M_n = \frac{\alpha_{2,n}}{\alpha_{1,n} + \epsilon_n}; \quad \epsilon_n = \sqrt{\alpha_{1,n}^2 + \alpha_{2,n}^2} \quad (109)$$

$$\gamma_{1,n} = \frac{(1 - \omega_n f_n)\alpha_{3,n} - \mu_0(\alpha_{1,n}\alpha_{3,n} + \alpha_{2,n}\alpha_{4,n})}{(1 - \omega_n f_n)^2 - \epsilon_n^2 \mu_0^2} \quad (110)$$

$$\gamma_{2,n} = \frac{-(1 - \omega_n f_n)\alpha_{4,n} - \mu_0(\alpha_{1,n}\alpha_{4,n} + \alpha_{2,n}\alpha_{3,n})}{(1 - \omega_n f_n)^2 - \epsilon_n^2 \mu_0^2} \quad (111)$$

The resonance at $\epsilon_n \mu_0 = (1 - \omega_n f_n)$ can be avoided by changing μ_0 with a small increment.

	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 64/74
---	--------------------------------	--

The numerical effort to calculate the coefficients $\bar{\beta}$ and $\beta(\mu_0)$ in Eq. (42) and (41) 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. Zdankowski et al. [RD47] proposed the following approximations:

$$\bar{\beta}_n = \frac{3 - p_{1,n}}{8} \quad (112)$$

$$\beta_n(\mu_0) = \frac{1}{2} - \frac{\mu_0}{4} \frac{p_{1,n} - 3f_n}{1 - f_n} \quad (113)$$

$$f_n = \frac{p_{1,n}^2}{3^2} \quad (114)$$

with the diffusivity factors

$$U_{\downarrow} = U_{\uparrow} = U = 2 \quad (115)$$

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 ω_n , the optical depth $\Delta\tau_n$, the asymmetry factor, and the phase function $P_n(\cos \Theta)$ in single scattering geometry.

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

$$\frac{\partial \alpha_{1,n}}{\partial \omega_n} = -U(1 - \bar{\beta}_n) \quad (116)$$

$$\frac{\partial \alpha_{2,n}}{\partial \omega_n} = -U\bar{\beta}_n \quad (117)$$

$$\frac{\partial \alpha_{3,n}}{\partial \omega_n} = (1 - f_n)\beta_n(\mu_0) \quad (118)$$

$$\frac{\partial \alpha_{4,n}}{\partial \omega_n} = (1 - f_n)(1 - \beta_n(\mu_0)) \quad (119)$$

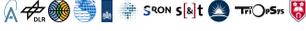
Furthermore,

$$\frac{\partial \epsilon_n}{\partial \omega_n} = \frac{1}{\epsilon_n} (\alpha'_{1,n} \alpha_{1,n} - \alpha'_{2,n} \alpha_{2,n}) \quad (120)$$

$$\frac{\partial E_n}{\partial \omega_n} = \epsilon'_n \tau_n E_n \quad (121)$$

$$\frac{\partial M_n}{\partial \omega_n} = \frac{\alpha'_{2,n} (\alpha_{1,n} + \epsilon_n) - \alpha_{2,n} (\alpha'_{1,n} + \epsilon'_n)}{(\alpha_{1,n} + \epsilon_n)^2} \quad (122)$$

and

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Date : 17 May 2019	Page 65/74

$$\frac{\partial \gamma_{1,n}}{\partial \omega_n} = \frac{-f_n \alpha_{3,n} + (1 - \omega_n f_n) \alpha'_{3,n} - \mu_0 (\alpha'_{1,n} \alpha_{3,n} + \alpha_{1,n} \alpha'_{3,n} + \alpha'_{2,n} \alpha_{4,n} + \alpha_{2,n} \alpha'_{4,n}) ((1 - \omega_n f_n)^2 - \epsilon_n^2 \mu_0^2)}{((1 - \omega_n f_n)^2 - \epsilon_n^2 \mu_0^2)^2} + \frac{[2f_n(1 - \omega_n f_n) + 2\epsilon_n \epsilon'_n \mu_0^2] ((1 - \omega_n f_n) \alpha_{3,n} - \mu_0 (\alpha_{1,n} \alpha_{3,n} + \alpha_{2,n} \alpha_{4,n}))}{((1 - \omega_n f_n)^2 - \epsilon_n^2 \mu_0^2)^2} \quad (123)$$

$$\frac{\partial \gamma_{2,n}}{\partial \omega_n} = \frac{f_n \alpha_{4,n} - (1 - \omega_n f_n) \alpha'_{4,n} - \mu_0 (\alpha'_{1,n} \alpha_{4,n} + \alpha_{1,n} \alpha'_{4,n} + \alpha'_{2,n} \alpha_{3,n} + \alpha_{2,n} \alpha'_{3,n}) ((1 - \omega_n f_n)^2 - \epsilon_n^2 \mu_0^2)}{((1 - \omega_n f_n)^2 - \epsilon_n^2 \mu_0^2)^2} + \frac{[2f_n(1 - \omega_n f_n) + 2\epsilon_n \epsilon'_n \mu_0^2] ((1 - \omega_n f_n) \alpha_{4,n} - \mu_0 (\alpha_{1,n} \alpha_{4,n} + \alpha_{2,n} \alpha_{3,n}))}{((1 - \omega_n f_n)^2 - \epsilon_n^2 \mu_0^2)^2} \quad (124)$$

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 ϵ'_n means the derivative $\frac{\epsilon}{\partial \omega_n}$.

The derivatives with respect to the optical depth $\Delta \tau_n$ are

$$\frac{\partial \alpha_{1,n}}{\partial \Delta \tau_n} = \frac{\partial \alpha_{2,n}}{\partial \Delta \tau_n} = \frac{\partial \alpha_{3,n}}{\partial \Delta \tau_n} = \frac{\partial \alpha_{4,n}}{\partial \Delta \tau_n} = 0 \quad (125)$$

$$\frac{\partial \epsilon_n}{\partial \Delta \tau_n} = \frac{\partial M_n}{\partial \Delta \tau_n} = 0 \quad (126)$$

$$\frac{\partial E_n}{\partial \Delta \tau_n} = -\epsilon_n E_n \quad (127)$$

$$\frac{\partial \gamma_{1,n}}{\partial \Delta \tau_n} = \frac{\partial \gamma_{2,n}}{\partial \Delta \tau_n} = 0. \quad (128)$$

For the derivative with respect to the asymmetry factor $p_{1,n}$, we obtain

$$\frac{\partial \alpha_{1,n}}{\partial p_{1,n}} = \frac{\partial \alpha_{2,n}}{\partial p_{1,n}} = -\frac{U \omega_n}{8} \quad (129)$$

$$\frac{\partial \alpha_{3,n}}{\partial p_{1,n}} = \omega_n \left[\beta'_n(\mu_0)(1 - f_n) - \beta_n(\mu_0) \frac{2p_{1,n}}{9} \right] \quad (130)$$

$$\frac{\partial \alpha_{4,n}}{\partial p_{1,n}} = -\omega_n \left[\beta'_n(\mu_0)(1 - f_n) + (1 - \beta_n(\mu_0)) \frac{2p_{1,n}}{9} \right] \quad (131)$$

with

$$\beta'_n(\mu_0) = \frac{\partial \beta_n(\mu_0)}{\partial p_{1,n}} = -\frac{\mu_0}{4} \left[\frac{1 - \frac{2p_{1,n}}{3}}{1 - f_n} + \frac{p_{1,n} - 3f_n}{(1 - f_n)^2} \frac{2p_{1,n}}{9} \right] \quad (132)$$

and

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Date : 17 May 2019	Page 66/74

$$\frac{\partial \epsilon_n}{\partial p_{1,n}} = -\frac{U\omega_n}{8\epsilon_n}(\alpha_{1,n} - \alpha_{2,n}) \quad (133)$$

$$\frac{\partial E_n}{\partial p_{1,n}} = \epsilon'_n E_n \quad (134)$$

$$\frac{\partial M_n}{\partial p_{1,n}} = \frac{\alpha'_{2,n}(\alpha_{1,n} + \epsilon_n) + \alpha_{2,n}(\alpha'_{1,n} + \epsilon'_n)}{(\alpha_{1,n} + \epsilon_n)^2} \quad (135)$$

$$\frac{\partial \gamma_1}{\partial p_{1,n}} = \frac{(1 - \omega_n f_n)\alpha_{3,n} - \omega_n \alpha_{3,n} \frac{2}{9} p_{1,n} - \mu_0(\alpha_{1,n}\alpha'_{3,n} + \alpha'_{1,n}\alpha_{3,n} + \alpha_{2,n}\alpha'_{4,n} + \alpha'_{2,n}\alpha_{4,n})}{((1 - \omega_n f_n)^2 - \epsilon_n^2 \mu_0^2)^2} \quad (136)$$

$$\frac{\partial \gamma_2}{\partial p_{1,n}} = \frac{(1 - \omega_n f_n)\alpha_{4,n} - \omega_n \alpha_{4,n} \frac{2}{9} p_{1,n} - \mu_0(\alpha_{1,n}\alpha'_{4,n} + \alpha'_{1,n}\alpha_{4,n} + \alpha_{2,n}\alpha'_{3,n} + \alpha'_{2,n}\alpha_{3,n})}{((1 - \omega_n f_n)^2 - \epsilon_n^2 \mu_0^2)^2} \quad (137)$$

With Eq. (116)–(137), 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. (104)–(108):

$$\frac{\partial a_{1,n}}{\partial \omega_n} = \frac{f_n \Delta \tau_n}{\mu_0} a_{1,n} \quad (138)$$

$$\frac{\partial a_{2,n}}{\partial \omega_n} = -a'_{4,n}\gamma_{2,n} - a_{4,n}\gamma'_{2,n} - a'_{5,n}\gamma_{1,n}a_{1,n} - a_{5,n}\gamma'_{1,n}a_{1,n} - a_{5,n}\gamma_{1,n}a'_{1,n} + \gamma_{2,n}a'_{1,n} + \gamma'_{2,n}a_{1,n} \quad (139)$$

$$\frac{\partial a_{3,n}}{\partial \omega_n} = -a'_{5,n}\gamma_{2,n} - a_{5,n}\gamma'_{2,n} - a'_{4,n}\gamma_{1,n}a_{1,n} - a_{4,n}\gamma'_{1,n}a_{1,n} - a_{4,n}\gamma_{1,n}a'_{1,n} + \gamma'_{1,n} \quad (140)$$

$$\frac{\partial a_{4,n}}{\partial \omega_n} = \frac{[E'_n(1 - M_n^2) - 2E_n M'_n M_n](1 - E_n^2 M_n^2) + 2E_n(1 - M_n^2)E_n M_n[E'_n M_n + E_n M'_n]}{(1 - E_n^2 M_n^2)^2} \quad (141)$$

$$\frac{\partial a_{5,n}}{\partial \omega_n} = \frac{[M'_n(1 - E_n^2) - 2M_n E'_n E_n](1 - E_n^2 M_n^2) + 2M_n(1 - E_n^2)E_n M_n[E'_n M_n + E_n M'_n]}{(1 - E_n^2 M_n^2)^2} \quad (142)$$

$$\frac{\partial a_{1,n}}{\partial \Delta \tau_n} = \frac{1 - \omega_n f_n}{\mu_0} a_{1,n} \quad (143)$$

$$\frac{\partial a_{2,n}}{\partial \Delta \tau_n} = a'_{4,n}\gamma_{2,n} - a'_{5,n}\gamma_{1,n}a_{1,n} - a_{5,n}\gamma_{1,n}a'_{1,n} + \gamma_{2,n}a'_{1,n} \quad (144)$$

$$\frac{\partial a_{3,n}}{\partial \Delta \tau_n} = -a'_{5,n}\gamma_{2,n} - a'_{4,n}\gamma_{1,n}a_{1,n} - a_{4,n}\gamma_{1,n}a'_{1,n} \quad (145)$$

$$\frac{\partial a_{4,n}}{\partial \Delta \tau_n} = \frac{(1 - M_n^2)E'_n(1 + E_n^2 M_n^2)}{(1 - E_n^2 M_n^2)^2} \quad (146)$$

$$\frac{\partial a_{5,n}}{\partial \Delta \tau_n} = \frac{2E_n E'_n M_n(1 - M_n^2)}{(1 - E_n^2 M_n^2)^2} \quad (147)$$

and

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Date : 17 May 2019	Page 67/74

$$\frac{\partial a_{1,n}}{\partial p_{1,n}} = -\frac{2\omega_n \Delta \tau_n p_{1,n}}{9\mu_0} a_{1,n} \quad (148)$$

$$\frac{\partial a_{2,n}}{\partial p_{1,n}} = -a'_{4,n} \gamma_{2,n} - a'_{4,n} \gamma'_{2,n} - a'_{5,n} \gamma_{1,n} a_{1,n} - a_{5,n} \gamma'_{1,n} a_{1,n} - a_{5,n} \gamma_{1,n} a'_{1,n} + \gamma_{2,n} a'_{1,n} + \gamma'_{2,n} a_{1,n} \quad (149)$$

$$\frac{\partial a_{3,n}}{\partial p_{1,n}} = -a'_{5,n} \gamma_{2,n} - a_{5,n} \gamma'_{2,n} - a'_{4,n} \gamma_{1,n} a_{1,n} - a_{4,n} \gamma'_{1,n} a_{1,n} - a_{4,n} \gamma_{1,n} a'_{1,n} + \gamma'_{1,n} \quad (150)$$

$$\frac{\partial a_{4,n}}{\partial p_{1,n}} = \frac{[E'_n(1 - M_n^2) - 2E_n M'_n M_n](1 - E_n^2 M_n^2) + 2E_n(1 - M_n^2)E_n M_n[E'_n M_n + E_n M'_n]}{(1 - E_n^2 M_n^2)^2} \quad (151)$$

$$\frac{\partial a_{5,n}}{\partial p_{1,n}} = \frac{[M'_n(1 - E_n^2) - 2M_n E'_n E_n](1 - E_n^2 M_n^2) + 2M_n(1 - E_n^2)E_n M_n[E'_n M_n + E_n M'_n]}{(1 - E_n^2 M_n^2)^2} \quad (152)$$

Finally to calculate the Jacobian in Eq. (63), we need the corresponding derivatives of the response vector

$$\mathbf{R} = (u_0, v_0, w_0, \dots, u_N, v_N, w_N) \quad (153)$$

with respect to the atmospheric parameters $\Delta \tau_n$, ω_n , β_n and the phase function P_n . Using Eq. (54), we obtain:

$$\frac{\partial u_n}{\partial \omega_m} = \delta_{n+1,m} \frac{1}{4\pi\mu_0\mu_v} \tilde{\mu} P_{n+1} t_{n+1}(\tilde{\mu}) \quad (154)$$

$$\frac{\partial v_n}{\partial \omega_m} = \frac{U^\downarrow}{4\pi\mu_v} [\delta_{nm} \beta_n t_n(\mu_v) + \delta_{n+1,m} \beta_{n+1} t_{n+1}(\mu_v)] \quad (155)$$

$$\frac{\partial w_n}{\partial \omega_m} = \frac{U^\uparrow}{4\pi\mu_v} [\delta_{nm} (1 - \beta_n) t_n(\mu_v) + \delta_{n+1,m} (1 - \beta_{n+1}) t_{n+1}(\mu_v)] \quad (156)$$

and

$$\frac{\partial u_n}{\partial P_m} = \delta_{n+1,m} \frac{\omega_{n+1}}{4\pi\mu_0\mu_v} \tilde{\mu} t_{n+1}(\tilde{\mu}) \quad (157)$$

$$\frac{\partial v_n}{\partial P_m} = 0 \quad (158)$$

$$\frac{\partial w_n}{\partial P_m} = 0 \quad (159)$$

and

$$\frac{\partial u_n}{\partial \beta_m} = 0 \quad (160)$$

$$\frac{\partial v_n}{\partial \beta_m} = \frac{U^\downarrow}{4\pi\mu_v} [\delta_{nm} \omega_n t_n(\mu_v) + \delta_{n+1,m} \omega_{n+1} t_{n+1}(\mu_v)] \quad (161)$$

$$\frac{\partial w_n}{\partial \beta_m} = -\frac{U^\uparrow}{4\pi\mu_v} [\delta_{nm} \omega_n t_n(\mu_v) + \delta_{n+1,m} \omega_{n+1} t_{n+1}(\mu_v)] \quad (162)$$

and

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 68/74
--	--------------------------------	--

$$\frac{\partial u_n}{\partial \Delta \tau_m} = \frac{\omega_{n+1}}{4\pi\mu_0\mu_v} \tilde{\mu} P_{n+1} t'_{n+1,m}(\tilde{\mu}) \quad (163)$$

$$\frac{\partial v_n}{\partial \Delta \tau_m} = \frac{U^\downarrow}{4\pi\mu_v} \left[\omega_n \beta_n t'_{n,m}(\mu_v) + \omega_{n+1} \beta_{n+1} t'_{n+1,m}(\mu_v) \right] \quad (164)$$

$$\frac{\partial w_n}{\partial \Delta \tau_m} = \frac{U^\uparrow}{4\pi\mu_v} \left[\omega_n (1 - \beta_n) t'_{n,m}(\mu_v) + \omega_{n+1} (1 - \beta_{n+1}) t'_{n+1,m}(\mu_v) \right] \quad (165)$$

with the derivative

$$t'_{n,m}(\mu) = \frac{\partial t_n(\mu)}{\partial \Delta \tau_m} = \begin{cases} -\frac{1}{\mu_v} t_n(\mu) & \text{for } m < n \\ \frac{1}{\mu} e^{-\tau_{n-1}/\mu_v} e^{-\Delta \tau_n/\mu} & \text{for } m = n = 1, \dots, N \\ 0 & \text{for } m > n \end{cases} \quad (166)$$

of the auxiliary function t_n in Eq. (49), and after adding the contribution by surface reflection

$$\frac{\partial w_N}{\partial \Delta \tau_m} = \frac{U^\uparrow}{4\pi\mu_v} \omega_N (1 - \beta_N) t'_{N,m}(\mu_v) - \frac{1}{\pi\mu_v} e^{-\tau_N/\mu_v}, \quad (167)$$

because the contribution of surface reflection is included in the response to diffuse upward radiation.

	S5L2PP	Reference : SRON-ESA-S5L2PP-ATBD-002
	Carbon Monoxide ATBD	Version : 3.1 Page Date : 17 May 2019 69/74

B Algorithm settings

This appendix describes the user settings for SICOR-S5, which are provided in an ASCII format to be read in by the S5 framework and passed to SICOR as a character string via the framework interface. SICOR parses and interprets the content. The core of the SICOR algorithm is programmed for several missions and so parts of the user settings are predefined for the S5 mission. This section describes only setting parameters, which are relevant in the context of this ATBD. Changes of other predefined setting parameters are not recommended and may result in a performance which is not compliant with the described SICOR-S5 algorithm.

B.1 Filter settings

As part of the S5 CO protocol up to two retrievals are performed. At two designated points in the protocol, there is a filter step that, if triggered, stops the execution. All filter thresholds are part of the user settings.

Table 18: Filter settings used in SICOR-S5.

Tag	Description
radiance_threshold	Minimum intensity in SWIR-3 (T_{A1}).
sza_threshold	Maximum solar zenith angle (T_{A2}).
prefit_bias_threshold	Maximum difference between retrieved CH_4 from non-scattering retrieval and a-priori CH_4 (T_{B1}).

B.2 Atmospheric gridding

Both trace gas retrievals share the same atmospheric grid. Therefore, settings on the atmospheric grid are listed in a specific field of the settings. These comprise of the following.

Table 19: Atmospheric settings used in SICOR.

Tag	Description
nlay	Number of atmospheric layers.
dz	Geometric thickness per layer (m).
effective_gravity	Gravitational constant representative for the entire column (m s^{-2}).

B.3 Retrieval settings

Both retrievals (non-scattering and physics-based) have their own retrieval-specific settings. The retrieval settings include specific settings for each state parameter. The state parameter is explained in the ATBD and is different for the two different retrievals. Besides that, the retrieval settings comprise of the following.

Table 20: Retrieval settings used in SICOR.

Tag	Description
internal_sampling	Internal sampling of the band as wavenumber interval cm^{-1} .
external_start	Retrieval window starting wavelength (nm).
external_end	Retrieval window ending wavelength (nm).
internal_start	Set this one the ISRF range lower than external_start (nm).
internal_end	Set this one the ISRF range higher than external_end (nm).

Continued on next page

	S5L2PP	Reference : SRON-ESA-S5L2PP-ATBD-002
	Carbon Monoxide ATBD	Version : 3.1 Page
		Date : 17 May 2019 70/74

Table 20 – Retrieval settings - Continued from previous page

Tag	Description
minimum_pixels	Minimum number of living measurements to perform the retrieval.
minimum_pixels_nowarning	Minimum number of living measurements in order not to generate a warning.
cross_triangle_convolution	Flag for triangle convolution of cross section LUT.
sun_triangle_convolution	Flag for triangle convolution of solar spectrum.
measure_reflectance	Flag for reflectance retrieval or radiance retrieval.
execute_deconvolution	Flag for deconvolution solar spectrum (only applicable for radiance retrieval).
execute_spectral_shift	Flag for fitting spectral shift in solar spectrum.
spectral_shift_range	Range in which spectral shifts in solar spectrum will be adequately fitted (nm).
spectral_shift_sampling	Sampling for fitting spectral shifts in solar spectrum (nm).
protect_negative_absorption	Option how to protect the RTM for negative absorption optical thickness.
tikhonov_regularization	General strength of Tikhonov regularization.
stepcontrol_reducer_start	Starting value of Levenberg-Marquardt parameter.
stepcontrol_reducerfactor_success	Reduction factor of Levenberg-Marquardt parameter for a successful step.
stepcontrol_reducerfactor_fail	Multiplication factor of Levenberg-Marquardt parameter for a failed step.
stepcontrol_reduceraddition_fail	Additive value of Levenberg-Marquardt parameter for a failed step.
stepcontrol_max_rejections	Maximum number of allowed rejected steps.
unphysical_freeze_iterations	Number of iterations a fit parameter is left out of the inversion when it tends to unphysical values.
iterations	Maximum number of iterations.
minimum_iterations	Minimum number of iterations.
maximum_chi_square_change	Maximum χ^2 change for convergence.
maximum_chi_square	Maximum absolute χ^2 for convergence.
maximum_stepcontrol_reducer	Maximum Levenberg-Marquardt parameter for convergence.
maximum_scaled_desired_step	Maximum change in state vector in scaled space for convergence.

The state vector includes parameters, which describe a wavelength-dependent feature by a polynomial expression of limited order, using one state parameter for each order. For each state parameter, the following has to be added.

Table 21: State parameter settings used in SICOR.

Tag	Description
start	A-priori value, if not derived from the input.
fit	Flag for fitting the parameter.
regularization_weight	Regularization weight factor.
low	Lower limit for what is considered physical (not for polynomials).
high	Upper limit for what is considered physical (not for polynomials).
order	Polynomial order (only for polynomials).

Note that to comply the ATBD, the methane column is fixed to a prior value. For the physics-based retrieval, particular scattering parameters are added to the state parameter, where other scattering parameters are fixed to prior values.

 	<p style="text-align: center;">S5L2PP Carbon Monoxide ATBD</p>	<p>Reference : SRON-ESA-S5L2PP-ATBD-002 Version : 3.1 Page Date : 17 May 2019 71/74</p>
--	--	---

Table 22: Scattering settings used in SICOR.

Tag	Description
angstrom	Ångström exponent.
ssa	Single-scattering albedo.
asymmetry	Asymmetry parameter.

C Usage of L1b-flagging at framework level

The S5 Level-2 prototype processor uses the flag information in the L1b product as follows:

1. For processing-critical flags that apply to a ground pixel, an error flag for the ground pixel will be set in the L2 output product, and processing for the ground pixel is skipped.
2. For flags that apply to a ground pixel that potentially impact processing in a non-critical way, a warning flag for the ground pixel is set in the L2 output product, and processing for the ground pixel is performed.
3. For processing-critical flags that apply to a spectral pixel, the processing will assume a missing value for the spectral pixel and apply an algorithm specific evaluation of the missing values for each ground pixel
4. Flags that do not impact processing are ignored.

Table 23 lists details on the L1b flag usage (note that the .

Flag name	Possible values	Meaning	Action	L2 flag-ging
ground_pixel_quality	See below	Quality assessment information for each ground pixel		
	no_error	No error	Use pixel	
	solar_eclipse	Potentially affected by solar eclipse	Use pixel	Warning
	sun_glint_possible	Potentially affected by sun glint	Use pixel	Warning
	ascending	Pixel measured in ascending node (dlatitude/dt > 0)	Use pixel	
	night	Pixel measured on night side (SZA > 90)	Use pixel	
	geo_boundary_crossing	Pixel crosses a geo-boundary	Use pixel	
measurement_quality	geolocation_error	Error in geolocation assignment	Do not use pixel	Error
	See below	Overall quality information for a single measurement i.e. one value for each scanline		
	no_error	No error	Use pixel	
	proc_skipped	Processing skipped	Do not use pixel	Error
	no_residual	Undefined meaning (leftover from TROPOMI)	Use pixel	
	saa	measurement in South Atlantic Anomaly	Use pixel	Warning
	spacecraft_manoeuvre	Measurement during spacecraft manoeuvre	Use pixel	Warning
	sub_grp	Undefined meaning (leftover from TROPOMI)	Use pixel	
	irr_out_range	Irradiance out of range	Do not use pixel	Error
	sub_group	Undefined meaning (leftover from TROPOMI)	Use pixel	
spectral_channel_quality	See below	Quality assessment information for each spectral pixel	*	*
	no_error	No error	Use pixel	
	missing	Missing pixel	*	*

detector_column_ - qualification	bad_pixel	Bad pixel	*	*
	processing_error	Processing error	*	*
	saturated	Saturated pixel	*	*
	transient	Transient pixel	*	*
	rts	Pixel affected by RTS	*	*
	underflow	Underflow error	*	*
	See below	Qualification flag for a full column indicating detector column type or state		
	no_qualification	(Normal column)	Use pixel	
	skipped	Skipped column	Do not use pixel (i.e. ground pixel that includes this column)	Error
	uvn_odd	Odd Column number	Use pixel	
uvn_prepost	Pre/postscan column	Do not use pixel (i.e. ground pixel that includes this column)	Error	
uvn_overscan	Overscan column	Do not use pixel (i.e. ground pixel that includes this column)	Error	
detector_row_ qualification	swir_adc0	SWIR ADC0 used (TBC)	Use pixel	
	swir_adc1	SWIR ADC1 used (TBC)	Use pixel	
	swir_adc2	SWIR ADC2 used (TBC)	Use pixel	
	swir_adc3	SWIR ADC3 used (TBC)	Use pixel	
	See below	Qualification flag indicating row type or state.	**	**
	no_qualification	(Normal row)	Use pixel	
	uvn_ror	Read-out register row	**	**
	uvn_dump	Dump row	**	**
	uvn_covered	Covered row	**	**
	uvn_overscan	Overscan row	**	**
uvn_higain	Hi-gain row	TBD	TBD	
swir_reference	SWIR reference row	**	**	
gen_transition	Transition row	**	**	
gen_non_illuminated	Non illuminated row	**	**	
spectral_calibration_ - quality	Several	Spectral calibration quality assessment information for each ground pixel.	Do not use spectral calibration in case any flag set, instead use nominal calibration	Warning
quality_level	0-1	Overall quality assessment information for each spectral pixel	*Use pixel in case quality exceeds configurable threshold	*

 	S5L2PP Carbon Monoxide ATBD	Reference : SRON-ESA-S5L2PP-ATBD-002
		Version : 3.1 Date : 17 May 2019

overall_quality_flag	Several values	Overall quality flag for product	Use pixels in product	Copy quality flag to output
-----------------------------	----------------	----------------------------------	-----------------------	-----------------------------

Table 23: Usage of L1b flags by algorithm

*: Setting of this flag implies a “missing” spectral pixel that is excluded from processing. The following rules apply for processing of a ground pixel in relation to one or more missing spectral pixels: **: Flagged columns are treated as a set of missing spectral pixels for each ground pixel in the corresponding scanline. Warning and error criteria for the occurrence of such missing spectral pixels have been defined in the previous foot note.

SICOR-S5 does not apply any filtering on the amount of consecutive missing spectral pixel but uses a posteriori filter criteria for overall data quality control.