



Netherlands Institute for Space Research

Algorithm Theoretical Baseline Document for Sentinel-5 Precursor Methane Retrieval

Otto Hasekamp, Alba Lorente, Haili Hu, Andre Butz, Joost
aan de Brugh, Jochen Landgraf



document number : SRON-S5P-LEV2-RP-001
CI identification : CI-7430-ATBD
issue : 2.4.0
date : 2022-05-30
status : released

Document approval record

	digital signature
prepared:	
checked:	Michael Buchwitz and Thomas Krings Institute of Environmental Physics (IUP) University Bremen, Germany
approved PM:	
approved PI:	

Document change record

issue	date	item	comments
0.00.01	2012-09-25	all	Draft ATBD
0.00.02	2012-11-08		Comments internal review addressed. Note on the dependence of computational performance on compiler included. For GNU compiler significantly more processors are needed (60-90) than earlier estimated for the Intel compiler (page 22).
0.00.02	2012-11-23	doc number	Changed to report (RP) with new doc number
0.05	2013-06-04	all	Several additions from draft to version 1: More extended algorithm description, extended sensitivity study, section on validation.
0.05	2013-06-20	contents + doc number	Addressed comments internal review, changed to new document numbering
0.09	2013-11-27	all	addressed reviewer comments
0.10.00	2014-04-11	all	Removed instrument section. Introduced cirrus filter based on SWIR band. Slightly modified test ensemble to represent more realistic AOT (more demanding than previous ensemble).
0.11.00	2014-07-23	all	added fluorescence, deconvolution of solar spectrum, updated input/output tables, checked fraction of clear pixels with RAL VIIRS processor
0.13.00	2015-08-31	table 5-3 and 6-1	Update for limited release to S5-P Validation Team; Made consistent with actual product
0.14.00	2015-12-11	section 9	Improved quality of Fig 9-17
1.00	2016-02-05	all	V1.0 release; Addressed comments from internal review (all minor)
1.10	2019-02-01	Sec. 5.6	description of bias correction added
	2019-02-01	Sec. 6.3	Tab. 5 updated
	2019-02-01	Sec. 6.4	SZA data filter updated
	2019-02-01	Sec. 6.5	TBC specified
	2019-02-01	Sec. 9	added
2.2.0	2021-02-16	Sec. 5.1.1 and Table 4	Cross-section database update to SEOM-IAS
		Sec. 5.2.2	Inversion procedure update on regularization
		Sec. 5.6. and Figs. 5, 6	Update of bias correction
		Sec. 6.4 and others	Data selection and pre-filtering corrected
2.3.1	2021-10-15	Added Sec. 5.7 and Fig. 7	Information related to ocean measurements
2.4.0	2021-05-30	Added Sec. 5.8	Information related to surface albedo spectral dependence

Contents

Document approval record	2
Document change record	3
List of Tables	6
List of Figures	7
1 Introduction	9
1.1 Identification	9
1.2 Purpose and objectives	9
1.3 Document overview	9
2 Applicable and reference documents	10
2.1 Applicable documents	10
2.2 Standard documents	10
2.3 Reference documents	10
2.4 Electronic references	14
3 Terms, definitions and abbreviated terms	15
3.1 Acronyms and abbreviations	15
4 Introduction to methane retrieval algorithm	16
4.1 Background	16
4.2 The S5P spectral range	16
4.3 Heritage	17
4.4 Requirements	17
5 Algorithm description	18
5.1 Forward model	18
5.1.1 Model Atmosphere and Optical Properties	19
5.1.2 Modeling the top-of-atmosphere radiances	22
5.1.3 Fluorescence	24
5.1.4 Summary of Forward Model	24
5.2 Inverse algorithm	25
5.2.1 Definition of state vector and ancillary parameters	25
5.2.2 Inversion Procedure	26
5.2.3 Regularization of state vector and iteration strategy	28
5.2.4 Convergence criteria	28
5.3 Common aspects with other algorithms	28
5.4 Cloud Filtering	28
5.5 SWIR-NIR Co-Registration	30
5.6 Bias correction	30
5.7 Ocean measurements over sun-glint geometries	32
5.8 Surface reflectance spectral dependence	32
5.9 Algorithm overview	33
5.9.1 Required input	33
5.9.2 Algorithm implementation	35
6 Input-Output file description	38
6.1 High level data product description	38
6.2 Data selection approach	41
6.3 Treatment of Corrupted data	41
6.4 Timeliness	42
7 Error analysis	43
7.1 Simulation of geophysical test cases	43
7.2 Model errors	43
7.2.1 Spectroscopic data	43
7.2.2 Temperature profile	45
7.2.3 Surface Pressure	46
7.2.4 Absorber profiles	46
7.3 Instrument errors	48

7.3.1	Signal to noise ratio	48
7.3.2	Instrument spectral response function	49
7.3.3	Position of the spectral channels	49
7.3.4	Radiometric offset (additive factor)	51
7.3.5	Radiometric gain (multiplicative factor)	52
7.3.6	Combined errors and fitting options	53
7.4	Filtering for clouds / cirrus	55
7.5	Fluorescence	56
7.6	Summary and discussion of error analysis	56
8	Validation	59
9	Examples of TROPOMI CH₄ data	60
10	Conclusions	61
11	Appendix B: Description of test cases	62

List of Tables

1	Spectral ranges from the NIR and SWIR band included in the measurement vector	18
2	A priori values for the different state vector elements.	26
3	Dynamic input data for S5P XCH4 algorithm.	33
4	Static input data for S5P XCH4 algorithm.....	35
5	Contents of the output product.	38
6	Performance of the reference retrieval for the ensemble of 9030 synthetic measurements. .	43
7	Sensitivity of retrieval performance to model errors in methane, water, temperature and pressure profiles. The results for retrievals with a temperature offset and surface pressure fitted are shown in brackets.	47
8	Influence of error in ISRF on the retrieval performance. Multiple simulation runs were performed; with FWHM increased and decreased. The results shown are for the run with poorest performance in terms of RMS error.	50
9	Influence of a constant error in position of spectral channels. Multiple simulation runs were performed; with $\Delta\lambda_{\text{shift}}$ positive and negative. The results shown are for the run with poorest performance. The results for the case a spectral shift is fitted is given in brackets.	51
10	Influence of a wavelength dependent error in position of spectral channel. Multiple simulation runs were performed; with $\Delta\lambda_{\text{squeeze}}$ positive and negative. The results shown are for the run with poorest performance. The results for the case a spectral squeeze is fitted is given in brackets.	51
11	Influence of a radiometric offset. Multiple simulation runs were performed; with positive and negative offset. The results shown are for the run with poorest performance. The results for the case an offset is fitted is given in brackets.	52
12	Influence of an error in radiometric gain. Multiple simulation runs were performed; with multiplicative factor smaller and greater than 1. The results shown are for the run with poorest performance.	53
13	Influence of combined instrument errors on retrieval performance. Retrieval settings are the same as for the reference retrieval.	54

List of Figures

1	Simulated spectrum at S5P spectral resolution for the O ₂ A band (left panel) and the SWIR band (right panel).....	17
2	Aerosol size distribution $n_{aer}(r)$ as a function of particle radius r . The retrieval method relies on a power law (red solid) size distribution. Also shown are more realistic multi-modal lognormal size distributions for a fine mode (black dashed) and a coarse mode (black dotted) dominated aerosol type.....	21
3	Relative difference between a spectrum calculated using the linear- k method and a spectrum obtained using line-by-line calculations. The spectra have been convolved with a Gaussian spectral response function with a Full Width at Half Maximum (FWHM) of 0.4 nm. For the calculations a boundary layer aerosol was used with an optical thickness of 0.3 at 765 nm. Furthermore, we used a solar zenith angle (SZA) of 50° and a viewing zenith angle of 0°...	25
4	Overview of forward model.	25
5	Ratio of reference XCH ₄ to TROPOMI XCH ₄ as a function of the retrieved surface albedo as explained in step 3 in the derivation of the bias correction. Green dots show the average ratio in 0.05 albedo bins and orange line shows the B-spline fit used to derive the bias correction. Data is averaged from 1 Jan 2018 until 31 Dec 2019 in a 0.1° x 0.1° grid.....	31
6	(a) Global TROPOMI XCH ₄ distribution before correction, (b) after correction and (c) their difference ($XCH_4 - XCH_4^{corr}$) for 2019 averaged in a cylindrical equal-area grid with 0.3° x 0.5° resolution at the Equator.	31
7	(a) Daily global TROPOMI XCH ₄ coverage, and (b) global XCH ₄ for 2020 averaged in a cylindrical equal-area grid with 0.3° x 0.5° resolution at the Equator.....	32
8	High level overview of XCH ₄ processing scheme.....	36
9	Overview of processing per ground pixel.	36
10	Retrieval forward model error of XCH ₄ for the (unfiltered) reference retrieval. White colors indicate retrievals did not converge, or synthetic spectra were not calculated (oceans), or SZA > 70 degree. The XCH ₄ error is defined as the difference between the retrieved and the true value and relative values are calculated with respect to the true XCH ₄	44
11	Forward model error of XCH ₄ plotted against retrieved SWIR albedo (left panel) and aerosol filter parameter f (right panel). The vertical lines give the cut-offs for the a posteriori filters. Retrievals with SWIR albedo < 0.02 are already filtered out in the right panel.	44
12	Forward model error of XCH ₄ after applying the a posteriori filter: SWIR albedo > 0.02 and aerosol load not too high ($f < 110$).	45
13	Cumulative probability distribution of the absolute XCH ₄ retrieval error for the reference retrieval that includes scattering of aerosols (green: all converging retrievals, blue: a posteriori filters applied to results).....	46
14	Sensitivity of XCH ₄ retrieval to error in temperature profile. Upper panel: Cumulative probability distribution of the absolute XCH ₄ retrieval error for the reference retrieval (blue), retrieval of ensemble with mean temperature profile per latitude without (green) and with (red) fitting of temperature offset. Lower panels: RMS error of XCH ₄ per latitude bin for same simulation runs as in upper panel. The red areas indicate regions with high SZA that we plan to filter out.	47
15	Influence of model error in prior pressure profile on accuracy (left) and stability (right) of retrievals without (blue) and with (red) fitting of the surface pressure.	48
16	Influence of model errors in prior absorber profiles, CH ₄ (blue) and H ₂ O (red), on retrieval accuracy (left) and stability (right).	48
17	Relative precision of XCH ₄ due to the instrument noise for the a posteriori filtered dataset.	49
18	RMS error of XCH ₄ for the simulation run with $\Delta FWHM_{NIR} = -1\%$ and $\Delta FWHM_{SWIR} = +1\%$ (green line), compared to the reference retrieval (blue line). The red areas indicate regions with high SZA that we plan to filter out.	50
19	RMS error of XCH ₄ retrieval per latitude bin for reference retrieval (blue), offsetted by 0.1% of the continuum but not fitted (green), and offsetted plus fitted (red). The red areas indicate regions with high SZA that we plan to filter out.	52
20	Same as Figure 19 but for an offset of 0.5%.....	53
21	RMS of relative precision of XCH ₄ retrievals per latitude bin for reference retrieval (blue), offsetted (by 0.1% of the continuum) but not fitted (green), and offsetted plus fitted (red). The red areas indicate regions with high SZA that we plan to filter out.	54

22	RMS error of XCH_4 when all instrument errors considered here are applied to the synthetic spectra (green line), compared to reference retrieval (blue line). The retrieval settings are the same for both simulation runs. The red areas indicate regions with high SZA that we plan to filter out.	55
23	Forward model error of XCH_4 after applying the a posteriori filters and the a priori cloud filter based on non scattering H_2O and CH_4 retrievals.	56
24	Cumulative probability distribution of the absolute XCH_4 retrieval error for the reference retrieval that includes scattering of aerosols and cirrus for the baseline filter, baseline and an additional filter based on non scattering CH_4 retrievals, baseline and an additional filter based on non-scattering H_2O retrievals, and the baseline and both CH_4 and H_2O filters. .	57
25	Difference in retrieved columns between weak and strong absorption bands for non-scattering atmosphere for CH_4 (solid) and H_2O (dotted). For the creation of synthetic measurements a cloud with optical thickness 5 was used with cloud top height = 2 km and cloud geometrical thickness = 1 km.	57
26	Aerosol filter f value versus cloud fraction. Cases with $f > 110$ will be filtered out, which roughly corresponds to a cloud fraction of 0.08.	58
27	RMS XCH_4 error as a function of fluorescence emission for retrievals with and without fluorescence included in the fit.	58
28	CH_4 total column mixing ratio bias corrected of TROPOMI averaged from November 28th 2018 to January 16th 2019	60
29	Global distribution of Aerosol Optical Thickness (AOT) adopted for the synthetic ensemble. January (top left), April (top right), July (bottom left), and October (bottom right).	63
30	Global distribution of Cirrus Optical Thickness (COT) as used in the ensemble calculations.	63
31	Surface albedo maps in the SWIR range for the synthetic ensemble.	64

1 Introduction

1.1 Identification

This document describes the algorithm for the retrieval of the column average dry air mixing ratio of methane, X_{CH_4} , from Sentinel-5 Precursor (S5-P) measurements in the Near Infra Red (NIR) and Short Wave Infra Red (SWIR) spectral range. The algorithm name is RemoTeC-S5P and it is one of the deliverables of the ESA project 'Sentinel-5 P level 2 processor development' [AD1].

1.2 Purpose and objectives

The purpose of the document is to describe the theoretical baseline of the algorithm that is used for the operational processing to retrieve the column average dry air mixing ratio of methane X_{CH_4} , from Sentinel-5 Precursor (S5-P) measurements, the input and ancillary data that is needed, and the output that is generated. In addition, information about operational calculation times and the accuracy are provided.

1.3 Document overview

Chapter 4 describes the main characteristics of the TROPOMI instrument relevant for the X_{CH_4} retrieval, and briefly introduces the X_{CH_4} retrieval algorithm. Chapter 5 provides a detailed description of the baseline retrieval algorithm. In Chapter 6 an input-output file description is given. Chapter 7 gives a detailed error analysis, Chapter 8 links to the Mission Performance Centre routine validation reports, and Chapter 9 gives some examples of TROPOMI X_{CH_4} data. Finally, Chapter 10 concludes the document.

2 Applicable and reference documents

2.1 Applicable documents

- [AD1] Sentinel-5P Level 2 Processor Development – Statement of Work.
source: ESA; **ref:** S5P-SWESA-GS-053; **date:** 2012.
- [AD2] GMES Sentinels 4 and 5 mission requirements document.
source: ESA; **ref:** EOP-SMA/1507/JL-dr; **date:** 2011.

2.2 Standard documents

- [SD1] Space Engineering – Software.
source: ESA; **ref:** ECSS-Q-ST-80C; **date:** 2009.
- [SD2] Space Product Assurance – Software Product Assurance.
source: ESA; **ref:** ECSS-E-ST-40C; **date:** 2009.
- [SD3] TROPOMI Instrument and Performance Overview.
source: KNMI; **ref:** S5P-KNMI-L2-0010-RP; **issue:** 0.10.0; **date:** 2014.
- [SD4] Requirements for the Geophysical Validation of Sentinel-5 Precursor Products, draft version.
source: ESA; **ref:** S5P-RS-ESA-SY-164; **date:** 2014.

2.3 Reference documents

- [RD1] Terms, definitions and abbreviations for TROPOMI L01b data processor.
source: KNMI; **ref:** S5P-KNMI-L01B-0004-LI; **date:** 2011.
- [RD2] Terms and symbols in the TROPOMI algorithm team.
source: KNMI; **ref:** SN-TROPOMI-KNMI-049; **date:** 2012.
- [RD3] S. Solomon, D. Qin, M. Manning *et al.* (editors) *IPCC, Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007).
- [RD4] H. Bovensmann, J. P. Burrows, M. Buchwitz *et al.*; SCIAMACHY: Mission Objectives and Measurement Modes. *Journal of Atmospheric Sciences*; **56** (1999), 127.
- [RD5] T. Yokota, H. Oguma, I. Morino *et al.*; Test measurements by a BBM of the nadir-looking SWIR FTS aboard GOSAT to monitor CO₂ column density from space. *In Passive Optical Remote Sensing of the Atmosphere and Clouds IV* (edited by S. C. Tsay, T. Yokota and M.-H. Ahn); volume 5652 of *Proceedings of the SPIE*; (pp. 182–188) (2004); doi:10.1117/12.578497.
- [RD6] A. Kuze, H. Suto, M. Nakajima *et al.*; Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring. *Appl. Opt.*; **48** (2009), 6716; doi:10.1364/AO.48.006716.
- [RD7] J. F. Meirink, H. J. Eskes and A. P. H. Goede; Sensitivity analysis of methane emissions derived from SCIAMACHY observations through inverse modelling. *Atmos. Chem. Phys.*; **6** (2006), 1275.
- [RD8] P. Bergamaschi, C. Frankenberg, J. F. Meirink *et al.*; Satellite cartography of atmospheric methane from SCIAMACHY on board ENVISAT: 2. Evaluation based on inverse model simulations. *J. Geophys. Res.*; **112** (2007), D02304; doi:10.1029/2006JD007268.
- [RD9] 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; doi:10.1029/2010JD014514.
- [RD10] I. Aben, O. Hasekamp and W. Hartmann; Uncertainties in the space-based measurements of CO₂ columns due to scattering in the Earth's atmosphere. *J. Quant. Spectrosc. Radiat. Transfer*; **104** (2007), 450; doi:10.1016/j.jqsrt.2006.09.013.

- [RD11] A.M.S. Gloudemans, H. Schrijver, O.P. Hasekamp *et al.*; Error analysis for CO and CH₄ total column retrieval from SCIAMACHY 2.3 μ m spectra. *Atmos. Chem. Phys.*; **8** (2008), 3999.
- [RD12] 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.
- [RD13] C. Frankenberg, P. Bergamaschi, A. Butz *et al.*; Tropical methane emissions: A revised view from SCIAMACHY onboard ENVISAT. *Geophys. Res. Lett.*; **35** (2008), 15811; doi:10.1029/2008GL034300.
- [RD14] O. Schneising, M. Buchwitz, M. Reuter *et al.*; Long-term analysis of carbon dioxide and methane column-averaged mole fractions retrieved from SCIAMACHY. *Atmos. Chem. Phys.*; **11** (2011), 2863; doi:10.5194/acp-11-2863-2011.
- [RD15] R. Parker, H. Boesch, A. Cogan *et al.*; Methane observations from the Greenhouse Gases Observing SATellite: Comparison to ground-based TCCON data and model calculations. *Geophys. Res. Lett.*; **38** (2011), L15807; doi:10.1029/2011GL047871.
- [RD16] 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; doi:10.1029/2012JD017549.
- [RD17] H. Bösch, G. C. Toon, B. Sen *et al.*; Space-based near-infrared CO₂ measurements: Testing the Orbiting Carbon Observatory retrieval algorithm and validation concept using SCIAMACHY observations over Park Falls, Wisconsin. *J. Geophys. Res.*; **111** (2006), D23302; doi:10.1029/2006JD007080.
- [RD18] B. J. Connor, H. Boesch, G. Toon *et al.*; Orbiting Carbon Observatory: Inverse method and prospective error analysis. *J. Geophys. Res.*; **113** (2008), D05305; doi:10.1029/2006JD008336.
- [RD19] S. Oshchepkov, A. Bril and T. Yokota; PPDF-based method to account for atmospheric light scattering in observations of carbon dioxide from space. *J. Geophys. Res.*; **113** (2008), D23210; doi:10.1029/2008JD010061.
- [RD20] 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; doi:10.1364/AO.48.003322.
- [RD21] A. Bril, S. Oshchepkov and T. Yokota; Retrieval of atmospheric methane from high spectral resolution satellite measurements: a correction for cirrus cloud effects. *Appl. Opt.*; **48** (2009), 2139; doi:10.1364/AO.48.002139.
- [RD22] M. Reuter, M. Buchwitz, O. Schneising *et al.*; A method for improved SCIAMACHY CO₂ retrieval in the presence of optically thin clouds. *Atmos. Meas. Tech.*; **3** (2010), 209.
- [RD23] 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.
- [RD24] H. Boesch, D. Baker, B. Connor *et al.*; Global Characterization of CO₂ Column Retrievals from Shortwave-Infrared Satellite Observations of the Orbiting Carbon Observatory-2 Mission. *Remote Sens. Environ.*; **3** (2011), 270; doi:10.3390/rs3020270.
- [RD25] Y. Yoshida, Y. Ota, N. Eguchi *et al.*; Retrieval algorithm for CO₂ and CH₄ column abundances from short-wavelength infrared spectral observations by the Greenhouse gases observing satellite. *Atmos. Meas. Tech.*; **4** (2011), 717; doi:10.5194/amt-4-717-2011.
- [RD26] C. W. O'Dell, B. Connor, H. Bösch *et al.*; The ACOS CO₂ retrieval algorithm - Part 1: Description and validation against synthetic observations. *Atmos. Meas. Tech.*; **5** (2012), 99; doi:10.5194/amt-5-99-2012.
- [RD27] O. P. Hasekamp and A. Butz; Efficient calculation of intensity and polarization spectra in vertically inhomogeneous scattering and absorbing atmospheres. *J. Geophys. Res.*; **113** (2008), D20309; doi:10.1029/2008JD010379.

- [RD28] S. Guerlet, A. Butz, D. Schepers *et al.*; Impact of aerosol and thin cirrus on retrieving and validating XCO₂ from GOSAT shortwave infrared measurements. *J. Geophys. Res.*; **118** (2013), 4887; doi:10.1002/jgrd.50332.
- [RD29] S. Guerlet, S. Basu, A. Butz *et al.*; Reduced carbon uptake during the 2010 Northern Hemisphere summer from GOSAT. *Geophys. Res. Lett.*; **40** (2013), 2378; doi:10.1002/grl.50402.
- [RD30] J.P. Veefkind, I. Aben, K. McMullan *et al.*; TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sens. Environ.*; **120** (2012), 70; doi:10.1016/j.rse.2011.09.027.
- [RD31] 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; doi:10.1364/AO.46.000243.
- [RD32] 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; doi:10.5194/amt-3-1351-2010.
- [RD33] M. Birk., G. Wagner, J. Loos *et al.*; Methane and water spectroscopic database for TROPOMI Sentinel 5 Precursor in the 2.3 μ m region. volume 19; (p. 4652) (EGU General Assembly, 2017).
- [RD34] SEOM-IAS validation report.
source: DLR; **ref:** IAS-D09-PRJ-066; **date:** 2018.
- [RD35] L.S. Rothman, I.E. Gordon, A. Barbe *et al.*; The HITRAN 2008 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transfer*; **110** (2009), 533.
- [RD36] 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; doi:10.5194/amt-6-879-2013. URL <http://www.atmos-meas-tech.net/6/879/2013/>.
- [RD37] H. Tran, C. Boulet and J.-M. Hartmann; Line mixing and collision-induced absorption by oxygen in the A band: Laboratory measurements, model, and tools for atmospheric spectra computations. *J. Geophys. Res.*; **111** (2006) (D10), D15210; doi:10.1029/2005JD006869.
- [RD38] A. Bucholtz; Rayleigh-Scattering calculations for the terrestrial atmosphere. *Appl. Opt.*; **35** (1995), 2765.
- [RD39] J. E. Hansen and L. D. Travis; Light scattering in planetary atmospheres. *Space Science Reviews*; **16** (1974), 527; doi:10.1007/BF00168069.
- [RD40] M. I. Mishchenko, I. V. Geogdzhayev, B. Cairns *et al.*; Aerosol retrievals over the ocean by use of channels 1 and 2 AVHRR data: sensitivity analysis and preliminary results. *Appl. Opt.*; **38** (1999), 7325; doi:10.1364/AO.38.007325.
- [RD41] P. Stier, J. Feichter, S. Kinne *et al.*; The aerosol-climate model ECHAM5-HAM. *Atmos. Chem. Phys.*; **5** (2005), 1125.
- [RD42] O. Dubovik, A. Sinyuk, T. Lapyonok *et al.*; Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust. *J. Geophys. Res.*; **111** (2006) (D10), D11208; doi:10.1029/2005JD006619.
- [RD43] J. F. de Haan, P. B. Bosma and J. W. Hovenier; The adding method for multiple scattering calculations of polarized light. *Astron. and Astrophys.*; **183** (1987), 371.
- [RD44] Charles Cox and Walter Munk; Measurement of the Roughness of the Sea Surface from Photographs of the Sun's Glitter. *J. Opt. Soc. Am.*; **44** (1954) (11), 838; doi:10.1364/JOSA.44.000838. URL <http://www.osapublishing.org/abstract.cfm?URI=josa-44-11-838>.
- [RD45] 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.

- [RD46] O. P. Hasekamp and J. Landgraf; Linearization of vector radiative transfer with respect to aerosol properties and its use in satellite remote sensing. *J. Geophys. Res.*; **110** (2005), D04203; doi:10.1029/2004JD005260.
- [RD47] B. van Diedenhoven, O. P. Hasekamp and J. Landgraf; Efficient vector radiative transfer calculations in vertically inhomogeneous cloudy atmospheres. *Appl. Opt.*; **45** (2006), 5993; doi:10.1364/AO.45.005993.
- [RD48] M. Duan, Q. Min and J. Li; A fast radiative transfer model for simulating high-resolution absorption bands. *J. Geophys. Res.*; **110** (2005) (D15); doi:10.1029/2004JD005590. D15201; URL <http://dx.doi.org/10.1029/2004JD005590>.
- [RD49] C. Frankenberg, C. O'Dell, L. Guanter *et al.*; Remote sensing of near-infrared chlorophyll fluorescence from space in scattering atmospheres: implications for its retrieval and interferences with atmospheric CO₂ retrievals. *Atmos. Meas. Tech.*; **5** (2012), 2081; doi:10.5194/amt-5-2081-2012.
- [RD50] P.L. Phillips; A technique for the numerical solution of certain integral equations of the first kind. *J. Ass. Comput. Mat.*; **9** (1962), 84.
- [RD51] A.N. Tikhonov; On the solution of incorrectly stated problems and a method of regularization. *Dokl. Acad. Nauk SSSR*; **151** (1963), 501.
- [RD52] P. C. Hansen; *Rank-deficient and discrete ill-posed problems* (SIAM, Philadelphia, 1998).
- [RD53] H. Hu, J. Landgraf, R. Demeters *et al.*; Toward Global Mapping of Methane With TROPOMI: First Results and Intersatellite Comparison to GOSAT. *Geophys. Res. Lett.*; **45** (2018), 3682; doi:10.1002/2018GL077259.
- [RD54] A. Lorente, T. Borsdorff, A. Butz *et al.*; Methane retrieved from TROPOMI: improvement of the data product and validation of the first 2 years of measurements. *Atmospheric Measurement Techniques*; **14** (2021) (1), 665; doi:10.5194/amt-14-665-2021. URL <https://amt.copernicus.org/articles/14/665/2021/>.
- [RD55] C. D. Rodgers and B. J. Connor; Intercomparison of remote sounding instruments. *Journal of Geophysical Research (Atmospheres)*; **108** (2003), 4116; doi:10.1029/2002JD002299.
- [RD56] C.D. Rodgers; *Inverse Methods for Atmospheres: Theory and Practice*; volume 2 (World Scientific, 2000).
- [RD57] S5P-NPP Cloud Processor ATBD, issue 0.13.0.
source: RAL; **ref:** S5P-NPPC-RAL-ATBD-0001; **date:** 2015.
- [RD58] Algorithm Theoretical Baseline Document for Sentinel-5 Precursor: Carbon Monoxide Total Column Retrieval.
source: SRON; **ref:** SRON-S5P-LEV2-RP-002; **date:** 2016.
- [RD59] B. van Diedenhoven, O. P. Hasekamp and I. Aben; Surface pressure retrieval from SCIAMACHY measurements in the O₂ A Band: validation of the measurements and sensitivity on aerosols. *Atmos. Chem. Phys.*; **5** (2005), 2109.
- [RD60] M. Inoue, I. Morino, O. Uchino *et al.*; Bias corrections of GOSAT SWIR XCO₂ and XCH₄ with TCCON data and their evaluation using aircraft measurement data. *Atmospheric Measurement Techniques*; **9** (2016), 3491; doi:10.5194/amt-9-3491-2016.
- [RD61] L. Wu, O. Hasekamp, H. Hu *et al.*; Carbon dioxide retrieval from OCO-2 satellite observations using the RemoTeC algorithm and validation with TCCON measurements. *Atmospheric Measurement Techniques*; **11** (2018), 3111; doi:10.5194/amt-11-3111-2018.
- [RD62] C. W. O'Dell, A. Eldering, P. O. Wennberg *et al.*; Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm. *Atmospheric Measurement Techniques*; **11** (2018) (12), 6539; doi:10.5194/amt-11-6539-2018.
- [RD63] 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:DOI: 10.1016/S0022-4073(98)00007-7.

- [RD64] S5P Mission Performance Centre Methane [L2__CH4__] Readme.
source: ; **ref:** Version 2.2.0, S5P-MPC-SRON-PRF-CH4; **date:** 2021.
- [RD65] A. Galli, A. Butz, R. A. Scheepmaker *et al.*; CH₄, CO, and H₂O spectroscopy for the Sentinel-5 Precursor mission: an assessment with the Total Carbon Column Observing Network measurements. *Atmos. Meas. Tech.*; **5** (2012), 1387; doi:10.5194/amt-5-1387-2012.
- [RD66] A. Butz, A. Galli, O.P. Hasekamp *et al.*; TROPOMI aboard Sentinel-5 Precursor: Prospective performance of CH₄ retrievals for aerosol and cirrus loaded atmospheres. *Remote Sens. Environ.*; **120** (2012), 267.
- [RD67] GMES Sentinel-5 Precursor – S5p System Requirement Document.
source: ESA; **ref:** S5p-RS-ESA-SY-0002; **date:** 2011.
- [RD68] O.P. Hasekamp and J. Landgraf; A linearized vector radiative transfer model for atmospheric trace gas retrieval. *J. Quant. Spectrosc. Radiat. Transfer*; **75** (2002).
- [RD69] M. Hess and M. Wiegner; COP: a data library of optical properties of hexagonal ice crystals. *Appl. Opt.*; **33** (1994) (33), 7740; doi:10.1364/AO.33.007740.
- [RD70] 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; doi:10.1175/1520-0469(1984)041<0846:APOTPS>2.0.CO;2.
- [RD71] D. M. Winker, W. H. Hunt and M. J. McGill; Initial performance assessment of CALIOP. *Geophys. Res. Lett.*; **34** (2007), L19803; doi:10.1029/2007GL030135.
- [RD72] H. Schrijver, A. M. S. Gloudemans, C. Frankenberg *et al.*; Water vapour total columns from SCIAMACHY spectra in the 2.36 μm window. *Atmos. Meas. Tech.*; **2** (2009), 561.
- [RD73] C. Popp, P. Wang, D. Brunner *et al.*; MERIS albedo climatology for FRESCO+ O₂ A-band cloud retrieval. *Atmospheric Measurement Techniques*; **4** (2011), 463; doi:10.5194/amt-4-463-2011.

2.4 Electronic references

There are no electronic references

3 Terms, definitions and abbreviated terms

Terms, definitions and abbreviated terms that are used in the development program for the TROPOMI L0 1b data processor are described in [RD1]. Terms, definitions and abbreviated terms that are used in development program for the TROPOMI L2 data processors are described in [RD2]. Terms, definitions and abbreviated terms that are specific for this document can be found below.

3.1 Acronyms and abbreviations

ECMWF	European Centre for Medium-Range Weather Forecasts
ENVISAT	Environmental Monitoring Satellite
GOSAT	Greenhouse gas Observing SATellite
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
NPP	NPOESS Preparatory Project
NIR	Near Infra Red
OCO	Orbiting Carbon Observatory
SCIAMACHY	SCanning Imaging Absorption SpectroMeter for Atmospheric CHartography
SWIR	Short Wave Infra Red
S5P	Sentinel-5 Precursor
TCCON	Total Carbon Column Observing Network
TROPOMI	Tropospheric Monitoring Instrument
UVN	Ultraviolet, Visible, Near-Infrared
VIIRS	Visible Infrared Imaging Radiometer Suite
XCH4	Column averaged dry air mixing ratio of methane

4 Introduction to methane retrieval algorithm

4.1 Background

Methane (CH_4) is, after carbon dioxide (CO_2), the most important contributor to the anthropogenically enhanced greenhouse effect [RD3]. Monitoring CH_4 abundances in the Earth's atmosphere is the dedicated goal of several current and future satellite missions. Such space borne observations aim at providing CH_4 column concentrations with high sensitivity at the Earth's surface, with good spatiotemporal coverage, and with sufficient accuracy to facilitate inverse modeling of sources and sinks. The Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) on board ENVISAT [RD4], that was operational 2002-2012, and the Greenhouse Gases Observing Satellite (GOSAT) [RD5], [RD6], launched 2009, have the capability to achieve these goals. Their observation strategy relies on measuring spectra of sunlight backscattered by the Earth's surface and atmosphere in the shortwave infrared (SWIR) spectral range. Absorption features of CH_4 allow for retrieval of its atmospheric concentration with high sensitivity to the ground and the lower atmosphere where the major CH_4 sources are located. The benefit of such measurements for estimating source/sink strengths, however, strongly depends on the precision and accuracy achieved. When correlated on the regional or seasonal scale, systematic biases of a few tenths of a percent can jeopardize the usefulness of satellite-measured CH_4 concentrations for source/sink estimates [RD7, RD8, RD9]. Scattering by aerosols and cirrus clouds is the major challenge for retrievals of methane from space-borne observations of backscattered sunlight in the SWIR spectral range. While contamination by optically thick clouds can be filtered out reliably, optically thin scatterers are much harder to detect yet still modify the light path of the observed backscattered sunlight and thus, can lead to underestimation or overestimation of the true methane column if not appropriately accounted for. The net light path effect strongly depends on the amount, the microphysical properties, and the height distribution of the scatterers as well as on the reflectance of the underlying ground surface [RD10, RD11]. Therefore, retrieval strategies rely on inferring the targeted gas concentration either simultaneously with scattering properties of the atmosphere or with a light path proxy. The latter 'proxy' approach has been successfully implemented for methane retrieval from SCIAMACHY measurements around 1600 nm, by using the CO_2 column, also retrieved from SCIAMACHY in the same spectral range, as a lightpath proxy [RD12]. The 'proxy' approach relies on the assumptions that scattering effects cancel in the ratio of the methane column and the CO_2 column and that a prior estimate of the CO_2 column is sufficiently accurate to reliably re-calculate the methane column from the ratio. By definition, the accuracy of the 'proxy' approach is contingent on the uncertainty of CO_2 column used for rescaling and on the cancellation of errors in the $\text{CH}_4 / \text{CO}_2$ ratio. Further applications of the proxy approach for methane retrieval from SCIAMACHY are described by Frankenberg et al., 2018 [RD13] and Schneising et al., 2011 [RD14]. For GOSAT, the proxy approach has been successfully applied by Parker et al., 2011 [RD15] and Schepers et al., 2012 [RD16].

Alternatively, scattering induced lightpath modification can be taken into account by simultaneously inferring the atmospheric CH_4 concentration and physical scattering properties of the atmosphere. Such 'physics-based' methods have been developed for space-based CO_2 and/or CH_4 measurements from SCIAMACHY, GOSAT, and the Orbiting Carbon Observatory (OCO) (e.g. [RD17, RD18, RD19, RD20, RD21, RD22, RD23, RD24, RD25, RD26]). The physics based methods make use of the Oxygen A-band around 760 nm and absorption bands of the target absorber (CH_4 and/or CO_2) in the SWIR spectral range. The advantage of physics based methods for methane retrieval compared to a proxy method is that they do not depend on accurate prior information on the CO_2 column. On the other hand, the physics based algorithms are more complex and may be limited by the information content of the measurement with respect to aerosol properties and/or forward model errors in the description of aerosols. A detailed comparison between the two methods for GOSAT is provided by Schepers et al., 2012 [RD16].

4.2 The S5P spectral range

The spectral range measured by the S5P instrument [SD3] does not allow for a light-path-proxy approach, and thus the effect of aerosols and cirrus should be accounted for using a physics based method as mentioned above. The spectral ranges to be used are shown in Figure 1. The algorithm retrieves 3 aerosol parameters (amount, size, height) simultaneously with the methane column (and other parameters such as surface albedo) in order to account for light path modification by aerosols. The information on aerosol parameters comes from the parts of the spectrum with strong absorption lines (O_2 in the NIR band, CH_4 and H_2O in the SWIR band) of which the depth and shape is modified by aerosol scattering. The basis of the algorithm is to fit a forward model capable of handling (multiple) scattering by molecules and particles in the atmosphere, with the approximate parameterization of atmospheric scattering properties described above to the S5P NIR and

SWIR measurements. An important aspect of the algorithm is computational speed, since the S5P instrument provides significantly more measurements than earlier instruments. For this purpose we developed a highly efficient radiative transfer model that avoids time consuming line-by-line calculations using a k -binning approach [RD27].

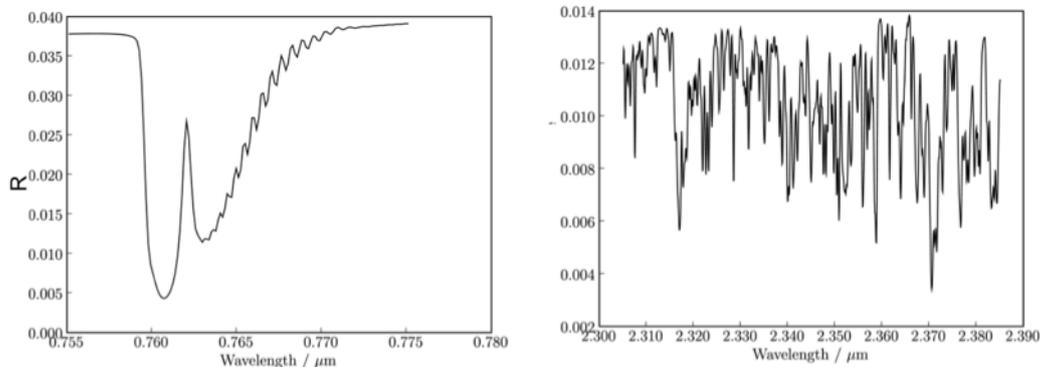


Figure 1: Simulated spectrum at S5P spectral resolution for the O₂ A band (left panel) and the SWIR band (right panel).

4.3 Heritage

The algorithm for retrieval of methane columns from the S5P instrument is based on earlier developments of a CO₂ and CH₄ retrieval algorithm from GOSAT, called RemoTeC [RD20, RD9, RD23, RD16, RD28], making use of measurements at the O₂ A band, and the 1600 nm and 2000 nm CO₂ absorption bands. For CO₂ and CH₄ the algorithm has been thoroughly tested for simulations where the actual scattering properties are unknown [RD20, RD9], and has been successfully applied to real GOSAT measurements [RD23, RD16, RD28, RD29]. For the operational S5P methane column retrieval algorithm we build further on the RemoTeC algorithm for GOSAT. Here, the algorithm had to be adjusted to the spectral range and spectral resolution of the S5P instrument. In the pre-launch phase, testing of the algorithm on real measurements for the S5P spectral range could not be done (SCIAMACHY measurements do not have sufficient quality in the 2.3 μm range), therefore we tested the S5P algorithm for an ensemble of scenarios with realistic combinations of aerosol properties, cirrus properties, surface albedo, and solar zenith angle, similarly to what was done for GOSAT prior to the GOSAT launch [RD9]. Since the S5P spectral range has no heritage for CH₄ retrieval, it is also important to investigate the quality of the relevant spectroscopic data in this spectral range, which is done using ground based Fourier Transform spectroscopy measurements.

4.4 Requirements

The accuracy requirement for the column averaged dry air-mixing ratio of methane (X_{CH_4}) was originally formulated as 2% [AD2]. Veefkind et al. (2012) [RD30] formulated this requirement as 2% accuracy and 0.6% precision (defined as the contribution of purely random instrument noise). Later, the requirement was reformulated as 1% bias and 1% precision [SD4]. From the 1% bias 0.6% is reserved for instrument related errors and 0.8% for forward model errors. It is also important to keep in mind the performance of the Japanese GOSAT satellite, launched 2009, which sets the current benchmark for methane retrievals from space. Performing GOSAT methane retrievals using the same algorithm as the S5P prototype algorithm [RD23, RD16], for methane we achieve a precision of ~0.8% per individual measurement and a relative accuracy (between regions) of ~0.25%.

Given the large number of measurement provided by the S5P instrument, the requirement on processing speed is very demanding. Taking into account that only cloud-free measurements over land are processed, which is ~3% of the total daytime data, the requirement on the processing time for a single S5P CH₄ retrieval is in the order of a few seconds assuming a reasonable amount of processing cores.

Table 1: Spectral ranges from the NIR and SWIR band included in the measurement vector

band	used spectral range
NIR	757-774 nm
SWIR	2305-2385 nm

5 Algorithm description

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

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

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

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \mathbf{e}_y + \mathbf{e}_F . \quad (2)$$

For methane retrieval from the S5P instrument the measurement vector contains the measured radiances in the spectral ranges 757-774 nm of the Near-InfraRed (NIR) channel and 2305-2385 nm in the SWIR channel (see Table 1).

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

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

where \mathbf{x}_n is the state vector for the n -th iteration step and \mathbf{K} is the Jacobian matrix

$$\mathbf{K} = \frac{\partial \mathbf{F}}{\partial \mathbf{x}} . \quad (4)$$

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

5.1 Forward model

The retrieval forward model \mathbf{F} simulates the measurement vector \mathbf{y} for a given model atmosphere defined by the state vector \mathbf{x} and the ancillary parameter vector \mathbf{b} . The simulated radiance for a given spectral pixel i for S5P is given by

$$I_{\text{conv},i} = \int_{\lambda_{\min}}^{\lambda_{\max}} I(\lambda) S_i(\lambda) d\lambda \quad (5)$$

where $I(\lambda)$ is the intensity modeled by the radiative transfer code (see below), and $S_i(\lambda)$ is the Instrument Spectral Response Function (ISRF) for spectral pixel i . In the NIR and SWIR channel $I(\lambda)$ contains many fine spectral structures due to molecular absorption, so it has to be calculated line-by-line at fine spectral sampling (at least 0.1 cm^{-1} in the NIR band and 0.02 cm^{-1} in the SWIR). In discretized form we can write the convolution with the ISRF by a matrix equation

$$\mathbf{I}_{\text{conv}} = \mathbf{S} \mathbf{I}_{\text{fine}} \quad (6)$$

where \mathbf{I}_{conv} is a vector containing the intensity measurements for all spectral pixels under consideration, \mathbf{S} is a matrix containing the ISRF for the different spectral pixels, and \mathbf{I}_{fine} is the modeled intensity spectrum at high spectral resolution. A similar equation can be written for the convolution of the solar spectrum with the ISRF.

In order to model the intensity at high spectral resolution also a high spectral resolution solar spectrum $\mathbf{F}_{0,\text{fine}}$ is needed. We obtain $\mathbf{F}_{0,\text{fine}}$ by performing a deconvolution of the measured solar spectrum $\mathbf{F}_{0,\text{meas}}$:

$$\mathbf{F}_{0,\text{fine}} = \mathbf{S}^T (\mathbf{S}\mathbf{S}^T)^{-1} \mathbf{F}_{0,\text{meas}} \quad (7)$$

The advantage of using a high spectral resolution solar spectrum obtained by deconvolution of the measured solar spectrum is that it is similarly affected by instrument features as the measured intensity spectrum. For further details and overview of the benefits of this method see van Deelen et al., 2007 [RD31].

5.1.1 Model Atmosphere and Optical Properties

For the S5P $X\text{CH}_4$ algorithm described here, the model atmosphere is defined for $N\text{LAY} = 36$ homogeneous vertical layers that are equidistant in pressure, the largest pressure level being defined by the surface pressure. The absorbing trace gases of interest are O_2 (in the NIR band) and CH_4 , H_2O , and CO in the SWIR band. The layer sub-columns of these gases are for the first iteration step of each retrieval calculated from the input profiles of CH_4 , CO (TM5) and H_2O (ECMWF) and the temperature and pressure profiles (ECMWF). They are obtained on the grid of the model atmosphere by linear interpolation. Here, first the surface pressure p_{surf} is obtained by interpolating the input pressure profile as function of height to the surface height (input) for the corresponding ground pixel. Next the pressure values at the layer boundaries are calculated, with the pressure p_k at the lower boundary of layer k (counting from top to bottom) is given by:

$$p_{\text{lev},k} = p_{\text{min}} + \Delta p \cdot k \quad (8)$$

$$\Delta p = (p_{\text{surf}} - p_{\text{min}})/N\text{LAY} \quad (9)$$

where p_{min} is the pressure value of the upper boundary of the input (ECMWF) atmosphere. The different atmospheric profiles are constructed on this pressure grid. For example, the methane sub-column DV_CH_{4k} for the layer bounded by pressure levels $p_{\text{lev},k-1}$ and $p_{\text{lev},k}$ is given by:

$$DV_CH_{4k} = X\text{CH}_{4k} DV_AIR_k \quad (10)$$

where $X\text{CH}_{4k}$ is the methane dry air mixing ratio linearly interpolated from the input pressure grid to the pressure at the 'middle' of layer k defined by $(p_k + p_{k+1})/2$. DV_AIR_k is the sub-column of air in layer k , given by

$$DV_AIR_k = \frac{(p_{\text{lev},k+1} - p_{\text{lev},k}) R}{M g_k (1 + \frac{X\text{H}_2\text{O}_k}{1.60855})} \quad (11)$$

where R is Avogadro's number, M is the molecular mass of air, g_k is the gravity constant in altitude layer k , and 1.60855 is the mass of air relative to the mass of water [RD32]. The sub columns of CO and H_2O are calculated in the same manner as for CH_4 , and the O_2 sub-column is obtained by multiplying the air sub-column by the O_2 mixing ratio ($=0.2095$).

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

$$\tau_{\text{abs},\text{CH}_4}(\lambda_j) = \frac{1}{N} \sum_{i=1}^N \sigma(p_i, T_i, \lambda_j) DV_CH_4 \quad (12)$$

where N is the number of sub-layers in which the model atmosphere layers are divided (set at $N = 2$) and σ is the absorption cross-section of CH_4 at wavelength λ_j , pressure p_i and temperature T_i at the center of model sub-layer i . The absorption optical thickness for the other trace gases is calculated in the same way.

Pre-calculated absorption cross-sections for CH_4 , CO , H_2O , and O_2 are stored as lookup-table as a function of pressure, temperature, and wavenumber. These cross-section lookup-tables are calculated from the latest spectroscopic databases. In preparation for the Sentinel 5 Precursor mission DLR established an improved spectroscopic database, the so-called Scientific Exploitation of Operational Missions – Improved Atmospheric Spectroscopy Databases (SEOM-IAS) for the interpretation of TROPOMI observations (<https://www.wdc.dlr.de/seom-ias/>) [RD33, RD34]. In the release of the processor version 2.2.0 we have updated the cross-section lookup-tables from HITRAN 2008 [RD35] with updates for water vapor from Scheepmaker et al. (2013) ([RD36]) to SEOM-IAS. In the NIR spectral range the absorption cross sections of O_2

are calculated according to Tran et al. (2006) [RD37] taking into account line mixing and collision induced absorption. The cross-section for pressure p_i , temperature T_i and wavelength λ_j are obtained by linear interpolation from the tabulated values. The reason that each of the $NLAY$ layers of the model atmosphere is further divided into N sub-layers, is to properly account for the strong dependence of temperature and pressure of the absorption cross-sections. So, the absorption cross-sections are needed for $NLAY \times N = 36 \times 2 = 72$ vertical layers equidistant in pressure.

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

$$\tau_{ray,k} = \sigma_{ray}(\lambda_j) DV_AIR_k \quad (13)$$

where σ_{ray} is the Rayleigh scattering cross section given by [RD38]

$$\sigma_{ray}(\lambda) = A \lambda^{-(4+X)} \quad (14)$$

$$X = B\lambda + \frac{C}{\lambda} - D \quad (15)$$

with $A = 4.02E-28$, $B = 0.389$, $C = 0.04926$, and $D = 0.3228$. The Rayleigh scattering phase function is given by (e.g. [RD39])

$$P(\Theta) = \frac{3}{4}(1 + \cos^2 \Theta) \frac{1 - \delta}{1 + \delta/2} \quad (16)$$

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

$$\cos \Theta = -u_0 u_v + \sqrt{(1 - u_v)^2} \sqrt{(1 - u_0)^2} \cos(\phi_0 - \phi_v) \quad (17)$$

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

In addition to trace gases, also aerosols are present in the model atmosphere. In our algorithm, aerosols are described by the following parameters (following [RD20, RD9]):

1. Number of particles in each layer of the model atmosphere. This is provided by the total amount of particles, N_{aer} , and a normalized altitude distribution described by a Gaussian function of center height z_{aer} and width w_0 . Hence, in model layer k with central height z_k , the number of particles is given by

$$h(z_k) = N_{aer} \exp \left[-\frac{4 \ln(z_k - z_{aer})^2}{4w_0^2} \right] \quad (18)$$

2. A size distribution of $n_{aer}(r)$, described by a power law size function, characterized by power law exponent and upper and lower cut-off (e.g. [RD40]):

$$n(r) = \begin{cases} A & \text{for } r \leq r_1 \\ A \left(\frac{r}{r_1}\right)^{-\alpha} & \text{for } r_1 < r \leq r_2 \\ 0 & \text{for } r > r_2 \end{cases} \quad (19)$$

The cut-offs are $r_1 = 0.1 \mu\text{m}$, $r_2 = 10 \mu\text{m}$ and the constant A is determined from normalization of the size distribution. Figure 2 illustrates $n_{aer}(r)$ and compares it to a more realistic multimodal lognormal size distribution [RD41]. Through its parameter α , the aerosol size distribution controls the spectral dependence of aerosol optical properties among the considered retrieval windows.

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

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

$$\tau_{scat,aer}(z_k) = \sigma_{scat,aer} h(z_k) \quad (20)$$

$$\tau_{abs,aer}(z_k) = \sigma_{abs,aer} h(z_k) \quad (21)$$

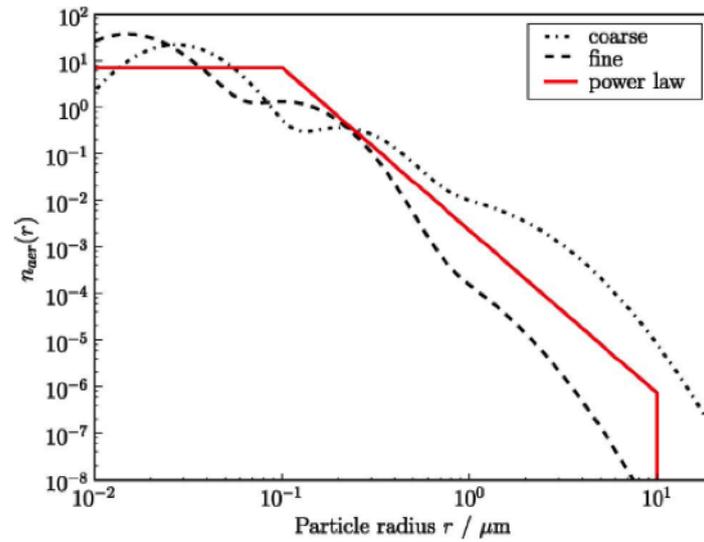


Figure 2: Aerosol size distribution $n_{aer}(r)$ as a function of particle radius r . The retrieval method relies on a power law (red solid) size distribution. Also shown are more realistic multi-modal lognormal size distributions for a fine mode (black dashed) and a coarse mode (black dotted) dominated aerosol type.

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

$$\sigma_{scat,aer} = \sum_{i=1}^M K_{scat,i}(m) r_i n_{aer}(r_i) v(r_i) \quad (22)$$

$$K_{scat,i}(m) = \int_{\Delta \ln r_i} \frac{\sigma_{scat}(r, m)}{v(r)} d \ln r \quad (23)$$

where v denotes particle volume and $K_{scat,i}$, representative for particle radius r_i , is pre-calculated for M size bins according to Dubovik et al., 2006 [RD42] and stored in a lookup table as function of aerosol size parameter $x = 2\pi r/\lambda$, real refractive index, and imaginary refractive index. The values for the actual aerosol refractive index are obtained by linear interpolation from the tabulated values. The lookup table contains values for spheres (Mie calculations) as well as for spheroids with a pre-defined axis ratio distribution [RD42], but in our algorithm baseline we only consider spherical particles. Similar expressions hold for the absorption cross-section and the aerosol scattering phase function.

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

$$\tau_{abs} = \tau_{abs,mol} + \tau_{abs,aer} \quad (24)$$

$$\tau_{sca} = \tau_{sca,mol} + \tau_{sca,aer} \quad (25)$$

$$P(\Theta) = \frac{\tau_{sca,mol} P_{mol}(\Theta) + \tau_{sca,aer} P_{aer}(\Theta)}{\tau_{sca}} \quad (26)$$

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

$$\alpha^l = \frac{2l+1}{2} \int_{-1}^1 P_{0,0}^l(\cos \Theta) P(\cos \Theta) d \cos \Theta \quad (27)$$

where $P_{0,0}^l$ is element (1,1) of the Generalized Spherical Function matrix (e.g. [RD43]).

To summarize, the forward model needs the following inputs:

- Surface pressure to define the equidistant pressure grid

- Sub-columns of CH₄, CO, H₂O, O₂, and air for the vertical layers of the model atmosphere.
- Pressure and temperature at the middle of the model sub-layers for absorption cross-sections.
- The aerosol column N_{aer} .
- The aerosol size parameter α (power of the power law size distribution function).
- The central height z_{aer} and width w_0 of the Gaussian aerosol altitude distribution.
- Solar Zenith Angle (SZA).
- Viewing Zenith Angle (VZA).
- Relative Azimuth Angle (RAA).
- The aerosol complex refractive index $m = m_r + i m_i$
- A high spectral resolution solar reference spectrum.
- Lookup tables with absorption cross-sections of CH₄, CO, H₂O, and O₂ as function of pressure, temperature, and wavenumber.
- Lookup tables with pre-calculated aerosol properties as function of aerosol size parameter, real refractive index, and imaginary refractive index (cross sections and phase functions integrated over each size parameter bin).
- The surface albedo and its spectral dependence (up to 3 coefficients).
- The fluorescence emission at 755 nm and its first order spectral dependence (see below).

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

5.1.2 Modeling the top-of-atmosphere radiances

Based on the optical properties (τ_{abs} , τ_{sca} , $P(\Theta)$) defined for each wavelength and layer of the model atmosphere we can compute the top-of-atmosphere radiance as measured by the S5P instrument. The first step is to separate the radiation field in a singly scattered component I_{ss} and a multiply scattered component I_{ms} , respectively:

$$I = I_{\text{ss}} + I_{\text{ms}} \quad (28)$$

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

$$I_{\text{ss}} = F_0 \sum_{k=1}^{\text{NLAY}} \omega_k P_k(\Theta) \left[1 - e^{-\tau_{\text{tot},k} \left(\frac{1}{u_0} + \frac{1}{u_v} \right)} \right] \frac{u_0}{4\pi(u_0 + u_v)} e^{-\sum_{i=1}^k \tau_{\text{tot},i} \left(\frac{1}{u_0} + \frac{1}{u_v} \right)} + e^{-\frac{1}{u_0} \sum_{i=1}^{\text{NLAY}} \tau_{\text{tot},i} \left(\frac{1}{u_0} + \frac{1}{u_v} \right)} R_{\text{surf}} \quad (29)$$

where F_0 is the incoming total flux, $\tau_{\text{tot}} = \tau_{\text{abs}} + \tau_{\text{sca}}$, $\omega = \tau_{\text{sca}}/\tau_{\text{tot}}$, u_0 is the cosine of the solar zenith angle, u_v is the cosine of the viewing zenith angle, and R_{surf} is the surface reflection for the specific solar and viewing geometry under consideration. Ocean surface reflection is modeled using the wind-speed-dependent Cox-and-Munk reflection model [RD44].

Besides I_{ss} itself, also the derivatives with respect to $\tau_{\text{sca},k}$, $\tau_{\text{abs},k}$, ω_k and P_k are needed:

$$\frac{\partial I_{\text{ss}}}{\partial \tau_{\text{tot},k}} = F_0 \omega_k P_k(\Theta) e^{-\tau_{\text{tot},k} \left(\frac{1}{u_0} + \frac{1}{u_v} \right)} \frac{u_0}{4\pi(u_0 + u_v)} e^{-\sum_{i=1}^k \tau_{\text{tot},i} \left(\frac{1}{u_0} + \frac{1}{u_v} \right)} - \sum_{k=1}^{\text{K}} \omega_k P_k(\Theta) \left[1 - e^{-\tau_{\text{tot},k} \left(\frac{1}{u_0} + \frac{1}{u_v} \right)} \right] \frac{u_0}{4\pi(u_0 + u_v)} e^{-\sum_{i=1}^k \tau_{\text{tot},i} \left(\frac{1}{u_0} + \frac{1}{u_v} \right)} \quad (30)$$

$$\frac{\partial I_{ss}}{\partial P_k} = F_0 \frac{u_0}{4\pi(u_0 + u_v)} e^{-\sum_{i=1}^k \tau_{\text{tot},i}(\frac{1}{u_0} + \frac{1}{u_v})} \left[1 - e^{-\tau_{\text{tot},k}(\frac{1}{u_0} + \frac{1}{u_v})} \right] \omega_k \quad (31)$$

$$\frac{\partial I_{ss}}{\partial \omega_k} = F_0 \frac{u_0}{4\pi(u_0 + u_v)} e^{-\sum_{i=1}^k \tau_{\text{tot},i}(\frac{1}{u_0} + \frac{1}{u_v})} \left[1 - e^{-\tau_{\text{tot},k}(\frac{1}{u_0} + \frac{1}{u_v})} \right] P_k \quad (32)$$

$$\frac{\partial I_{ss}}{\partial \tau_{\text{sca},k}} = \frac{1}{\omega_k} \frac{\partial I_{ss}}{\partial \omega_k} - \frac{\partial I_{ss}}{\partial \tau_{\text{tot},k}} \quad (33)$$

$$\frac{\partial I_{ss}}{\partial \tau_{\text{abs},k}} = \frac{\partial I_{ss}}{\partial \tau_{\text{tot},k}} - \frac{\omega_k}{\tau_{\text{tot},k}} \frac{\partial I_{ss}}{\partial \omega_k} \quad (34)$$

The computation of the multiply scattered radiation for a given wavelength involves the solution of the plane-parallel radiative transfer equation. We solve this equation using the Gauss-Seidel iterative method. The solution is described in detail by Landgraf et al., 2001 [RD45] and will not be repeated here. The calculation of the derivatives with respect to optical properties per layer k of the model atmosphere is performed using the forward adjoint perturbation theory and is described in detail by Hasekamp and Landgraf, 2005 [RD46].

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

$$\frac{\partial I_{ss}}{\partial x} = \sum_{k=1}^{\text{NLAY}} \left[\frac{\partial I_{ss}}{\partial \tau_{\text{sca},k}} \frac{\partial \tau_{\text{sca},k}}{\partial x} + \frac{\partial I_{ss}}{\partial \tau_{\text{abs},k}} \frac{\partial \tau_{\text{abs},k}}{\partial x} + \frac{\partial I_{ss}}{\partial P_k} \frac{\partial P_k}{\partial x} \right] \quad (35)$$

$$\frac{\partial I_{ms}}{\partial x} = \sum_{k=1}^{\text{NLAY}} \left[\frac{\partial I_{ms}}{\partial \tau_{\text{sca},k}} \frac{\partial \tau_{\text{sca},k}}{\partial x} + \frac{\partial I_{ms}}{\partial \tau_{\text{abs},k}} \frac{\partial \tau_{\text{abs},k}}{\partial x} + \sum_{l=0}^M \frac{\partial I_{ss}}{\partial \alpha_k^l} \frac{\partial \alpha_k^l}{\partial x} \right]. \quad (36)$$

In order to avoid time consuming multiple scattering calculations on the high spectral resolution line-by-line grid we aim at reducing the number of spectral calculations, following the approach of Hasekamp and Butz, 2008 [RD27]. For this purpose, we consider the intensity I_{ms} as a function of total absorption optical thickness τ_{abs} and its normalized vertical distribution \mathbf{n} (similar to k -distribution and spectral mapping methods),

$$I_{ms}(\lambda) = I_{ms}(\tau_{\text{abs}}(\lambda), \mathbf{n}(\lambda)), \quad (37)$$

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

$$\tau_{\text{abs},k}(\lambda) = n_k(\lambda) \tau_{\text{abs}}(\lambda), \quad (38)$$

where $\tau_{\text{abs},k}$ is the absorption optical thickness of layer k of the model atmosphere.

For a vertically homogeneous atmosphere, the advantage of the description as function of absorption optical thickness is obvious. Namely, the intensity depends smoothly on absorption optical thickness, which means that only calculations at a limited number of values of τ_{abs} are needed from which the Stokes parameters at other values of τ_{abs} can be obtained by interpolation (see e.g. [RD47]). Finally, the Stokes parameters can be mapped back into wavelength space. To apply this procedure for a non-homogeneous atmosphere, one has to assume that the vertical distribution $\mathbf{n}(z)$ of τ_{abs} can be approximated by a vertical distribution that is independent of wavelength in the spectral interval under consideration. This is the underlying assumption of the correlated k method. For the simulation of moderate- or high resolution spectra in the near- and shortwave infrared spectral ranges, this assumption may cause errors of several percent in reflectance spectra for realistic inhomogeneous terrestrial atmospheres (see e.g. [RD48]).

Clearly, an efficient radiative transfer model is needed that accounts for the vertical distribution of absorption optical thickness at different wavelengths. For this purpose we use the transformation into absorption optical

thickness space, and perform calculations for a limited number N values of $\tau_{\text{abs},k}$ of the absorption optical thickness and corresponding vertical distributions \mathbf{n}_k , with $k = 1, \dots, N$. From the reference calculations $I_{ms}(\tau_{\text{abs},k}, \mathbf{n}_k)$ we want to obtain the multiply scattered intensity vector $I_{ms}(\lambda_j)$ at any wavelength λ_j in the spectral range under consideration with absorption optical thickness $\tau_{\text{abs}}(\lambda_j)$ and its vertical distribution $\mathbf{n}(\lambda_j)$.

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

$$I_{ms}(\tau_{\text{abs},k}, \mathbf{n}(\lambda_j)) \approx I_{ms}(\tau_{\text{abs},k}, \mathbf{n}^k) + \frac{\partial I_{ms}}{\partial \mathbf{n}} [\mathbf{n}(\lambda_j) - \mathbf{n}^k], \quad (39)$$

where the derivatives with respect to \mathbf{n} follow from the forward adjoint perturbation theory [RD46]. Applying this equation yields the intensity vector of the multiply scattered radiation at the grid points $\tau_{\text{abs},k}$, corrected for the actual vertical distribution. In order to obtain the intensity I_{ms} at $\tau_{\text{abs}}(\lambda_j)$, we fit a second order polynomial to the logarithm of the (absorption profile corrected) intensities at the grid points, using the grid points closest to $\tau_{\text{abs}}(\lambda_j)$ and the two neighbouring points. In order to correct for variation of scattering properties and surface albedo within the spectral range under consideration, also a linear correction is used.

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

$$\tau_{\text{grid},k} = e^{\log(\tau_{\text{min}}) + \frac{k-1}{N} [\log(\tau_{\text{max}}) - \log(\tau_{\text{min}})]}, \quad (40)$$

where τ_{min} and τ_{max} are respectively the minimum and maximum absorption optical thickness in the spectral window under consideration. If τ_{max} is larger than 15 its value is set to 15, because for larger values of the absorption optical thickness the radiation field is dominated by single scattering (being calculated exactly) and hence interpolation errors are of minor importance. The rationale of choosing a logarithmic scale is to obtain more grid points at small values of absorption optical thickness, where multiple scattering effects are most important. If a certain spectral range is influenced by considerable absorption by two or more species we use 2 grids: one for the target absorber and one for the total of other absorbers. For the radiative transfer calculations in the methane retrieval algorithm we use 10 grid points in the NIR band and $5 \times 4 = 20$ grid points (5 for CH₄ and 4 for H₂O and CO combined) in the SWIR band. Figure 3 illustrates the accuracy of the linear- k method for the S5P NIR bands. For more information on the linear- k method we refer to [RD27].

5.1.3 Fluorescence

Fluorescence is included in the forward model using the approximation proposed by Frankenberg et al., 2012 [RD49]:

$$I_{\text{TOA}} = I_{\text{TOA}}^{nf} + F_s^{\text{surf}}(\lambda) e^{-\tau(\lambda)/\mu_v} \quad (41)$$

where the fluorescent emission $F_s^{\text{surf}}(\lambda)$ is approximated by

$$F_s^{\text{surf}}(\lambda) = F_{s,755\text{nm}}^{\text{surf}} (1 - s(\lambda - 755)) \quad (42)$$

where the fluorescent emission at 755 nm, $F_{s,755\text{nm}}^{\text{surf}}$, and the wavelength dependence coefficient s are input parameters to the forward model.

5.1.4 Summary of Forward Model

The different steps of the forward model calculation are summarized in Figure 4.

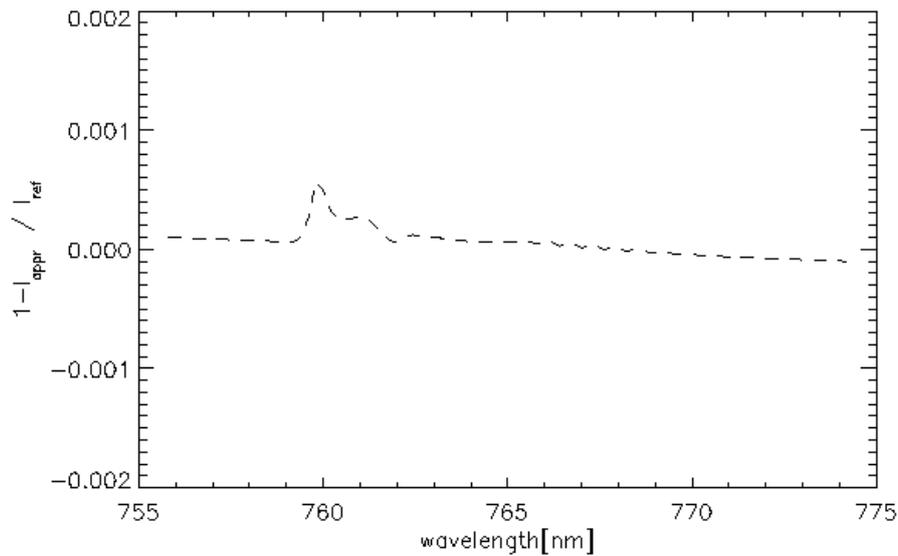


Figure 3: Relative difference between a spectrum calculated using the linear- k method and a spectrum obtained using line-by-line calculations. The spectra have been convolved with a Gaussian spectral response function with a Full Width at Half Maximum (FWHM) of 0.4 nm. For the calculations a boundary layer aerosol was used with an optical thickness of 0.3 at 765 nm. Furthermore, we used a solar zenith angle (SZA) of 50° and a viewing zenith angle of 0° .

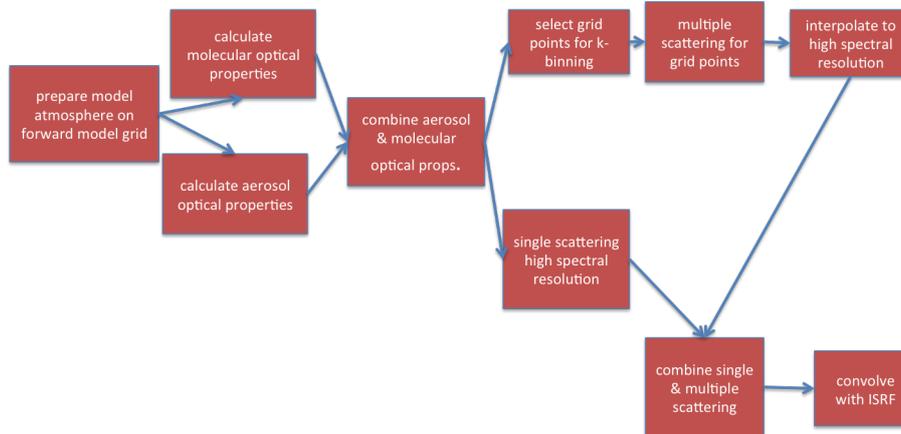


Figure 4: Overview of forward model.

5.2 Inverse algorithm

5.2.1 Definition of state vector and ancillary parameters

The **state vector** \mathbf{x} contains the following elements:

- CH₄ sub-columns in 12 vertical layers (layer interfaces coincide with NLAY layers of forward model grid).
- CO total column.
- H₂O total column.
- Aerosol parameter N_{aer} (number column).
- Aerosol parameter α (size parameter).

- Aerosol parameter z_{aer} (central height of Gaussian height distribution).
- Lambertian surface albedo in the NIR band.
- Third order spectral dependence of surface albedo in the NIR band.
- Lambertian surface albedo in the SWIR band.
- Third order spectral dependence of surface albedo in the SWIR band.
- Spectral shift of Earth radiances in the NIR band (higher orders optional).
- Spectral shift of Earth radiances in the SWIR band (higher orders optional).
- offset in the input temperature profile (optional).
- surface pressure (optional).

The **ancillary parameter vector \mathbf{b}** contains the following parameters:

- H₂O sub-columns in 36 vertical layers of the forward model grid.
- CO sub-columns in 36 vertical layers of the forward model grid.
- Temperature vertical profile at 72 layers of the cross-section vertical grid.
- Pressure vertical profile at 72 layers of the cross-section vertical grid.
- The aerosol complex refractive index, fixed at $m = 1.4 - 0.01i$ for the NIR and $m = 1.47 - 0.008i$ for the SWIR (no information present in measurement).
- The width w_o of the Gaussian aerosol height distribution, fixed at 2000 m (no information present in measurement).

State vector element	A priori value
CH ₄ sub-columns	TM5
CO total column	TM5
H ₂ O total column	ECMWF
N_{aer}	AOT=0.1 at 760 nm
α	fixed at 3.5
z_{aer}	fixed at 5000 m
surface albedo (NIR + SWIR)	no prior value needed (first guess at maximum of measured reflectance)
spectral shifts	no prior needed (first guess = 0)
temperature offset	no prior needed (first guess = 0)
surface pressure	ECMWF + GMTED2010 DEM
fluorescence emission $F_{s,755\text{nm}}^{\text{surf}}$	from FRESCO
spectral dependence s of fluorescence	fixed at $s = 0$

Table 2: A priori values for the different state vector elements.

5.2.2 Inversion Procedure

The inverse method optimizes the state vector \mathbf{x} with respect to the measurements \mathbf{y} after applying the forward model \mathbf{F} to \mathbf{x} . The inverse method is based by default on a Phillips-Tikhonov regularization scheme [RD50, RD51, RD46]. Regularization is required because the inverse problem is ill-posed, i.e., the measurements \mathbf{y} typically contain insufficient information to retrieve all state vector elements independently. The inverse

algorithm finds \mathbf{x} by minimizing the cost function that is the sum of the least-squares cost function and a side constraint weighted by the regularization parameter γ according to

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \left\{ \|\mathbf{S}_y^{-\frac{1}{2}}(\mathbf{F}(\mathbf{x}) - \mathbf{y})\|^2 + \gamma \|\mathbf{W}(\mathbf{x} - \mathbf{x}_a)\|^2 \right\}, \quad (43)$$

where \mathbf{S}_y is the diagonal measurement error covariance matrix, which contains the noise estimate. \mathbf{x}_a is an a priori state vector (see Table 2), and \mathbf{W} is a diagonal weighting matrix that renders the side constraint dimensionless and ensures that only the CH₄ parameters and the scattering parameters contribute to its norm: $W_{jj} = 1/x_{a,j}$ for the CH₄ column number densities and the three aerosol parameters, and $W_{jj} = 10^{-7}/x_{a,j}$ for all other state vector elements. The latter are thus retrieved in a least-squares sense. For determining γ , the L-curve criterion was applied [RD52] in the first version of the operational algorithm, as described by Hu et al. (2018) [RD53]. In the processor update for version 2.0.0., after two years of measurements, it is possible to select a constant regularization optimized for real observations. This includes a dedicated regularization parameter for the target absorber CH₄ and one for each of the scattering parameters (effective aerosol distribution height and size parameter, and effective aerosol column) [RD54].

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

$$\mathbf{x}_{n+1} = \min_{\mathbf{x}} \left\{ \|\mathbf{K}'(\mathbf{x}' - \mathbf{x}'_n) - \mathbf{y}'\|^2 + \gamma \|\mathbf{x}' - \mathbf{x}'_a\|^2 \right\} \quad (44)$$

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

The solution reads

$$\mathbf{x}_{n+1} = \mathbf{G}'\mathbf{y}' + \mathbf{A}'\mathbf{x}'_n + (\mathbf{I} - \mathbf{A}')\mathbf{x}'_{\text{apr}} \quad (45)$$

with \mathbf{A}' the averaging kernel matrix and \mathbf{G}' the contribution function matrix given by $\mathbf{A}' = \mathbf{G}'\mathbf{K}'$ and $\mathbf{G}' = (\mathbf{K}'^T\mathbf{K}' + \gamma\mathbf{I})^{-1}\mathbf{K}'^T$. If the retrieval converges after a given number of steps N (typically 7-8), the final state vector $\mathbf{x}_{\text{retr}} = \mathbf{x}_N$ is related to the true state vector and to the prior via

$$\mathbf{x}_{\text{retr}} = \mathbf{A}\mathbf{x}_{\text{true}} + (\mathbf{I} - \mathbf{A})\mathbf{x}_a + \mathbf{G}\mathbf{e}_y + \mathbf{G}\mathbf{e}_f. \quad (46)$$

The covariance matrix \mathbf{S}_x describing the retrieval noise $\mathbf{G}\mathbf{e}_y$ is given by

$$\mathbf{S}_x = \mathbf{G}\mathbf{S}_y\mathbf{G}^T. \quad (47)$$

For the S5P instrument, the target retrieval quantity is the column averaged dry air methane mixing ratio, $X\text{CH}_4$. This quantity is obtained from the methane entries of the retrieved state vector through

$$X\text{CH}_4 = \frac{1}{V_{\text{air,dry}}} \mathbf{h}^T \mathbf{x}_{\text{retr}} \quad (48)$$

where \mathbf{h} is the total column operator for methane (summing up the partial columns in the state vector) and $V_{\text{air,dry}}$ is the dry air column calculated from the surface pressure and water vapor profile, both obtained from a meteorological model (required as input). The retrieval noise $\Delta X\text{CH}_4$ on $X\text{CH}_4$ is given by

$$\Delta X\text{CH}_4 = \frac{\sum_{i,j=1}^{12} S_{x,i,j}}{V_{\text{air,dry}}}. \quad (49)$$

This is the error estimate that is given in the output together with $X\text{CH}_4$.

For validation and application purposes, it is important to realize that the retrieved $X\text{CH}_4$ is in fact a representation of $\mathbf{a}\mathbf{x}_{\text{true}}/V_{\text{air,dry}}$ where the quantity

$$\mathbf{a} = \mathbf{h}^T \mathbf{A} \quad (50)$$

is referred to as the column averaging kernel [RD55].

5.2.3 Regularization of state vector and iteration strategy

If the true solution is far from the solution of the current iteration \mathbf{x}_n and the forward model \mathbf{F} far from being linear around (\mathbf{x}_n) , the linear approximation may fail and yield a solution \mathbf{x}_{n+1} that is further away from the true solution than \mathbf{x}_n . To avoid diverging retrievals, we use a Gauss-Newton scheme (see e.g. [RD56]) with reduced step-size by introducing a filter factor that limits the update per iteration step of the state vector. The updated state vector is then given by:

$$\mathbf{x}_{n+1} = \Lambda \mathbf{G}' \mathbf{y}' + \mathbf{A}' \mathbf{x}'_n + (\mathbf{I} - \Lambda \mathbf{A}') \mathbf{x}'_{\text{apr}} \quad (51)$$

with the filter factor Λ given by

$$\Lambda = \frac{1}{1 + \xi} \quad \text{with } \xi > 0. \quad (52)$$

If ξ is large the update of the state vector is small. If $\xi = 0$, the equation for the updated state vector is equivalent to the pure Phillips-Tikhonov equation. The iteration is started with a large ξ , typically on the order of 10. It is then reduced or increased in the following iteration steps according to an empirically found scheme similar to Levenberg-Marquardt strategies [RD56]. We accept solution \mathbf{x}'_{n+1} and decrease ξ by a factor of 2.5 if the least squares norm of iteration $n + 1$ is smaller than 1.1 times the least squares norm of iteration n . Otherwise, we discard the solution of iteration $n + 1$, increase ξ by a factor of 2.5 and solve again for \mathbf{x}'_{n+1} . If ξ is smaller than a threshold value of 0.05, it is set to zero and the iteration is continued without a reduction in step size, assuming that the current state vector is sufficiently close to the true solution to finally converge.

5.2.4 Convergence criteria

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

1. The update of the $X\text{CH}_4$ has become smaller than its theoretical uncertainty.
2. The merit-function has not increased for the current iteration step, and the step-size factor ξ has reached 0.
3. The state vector entries have never reached unrealistic values during the retrieval (negative methane densities for instance).

5.3 Common aspects with other algorithms

The most important common aspect with the CO algorithm is that non-scattering retrievals performed by the CO algorithm is input to the CH_4 algorithm for cloud screening purposes (see below). Furthermore, a common aspect with the retrieval algorithm for CO retrieval is the use of a common lookup table for absorption cross sections for CH_4 , CO, and H_2O . Another common aspect with the CO algorithm is the use of a common dataset with prior information on methane (and CO) from an atmospheric transport model. Finally, the here developed highly efficient linear- k binning method for radiative transfer calculations in absorption bands is useful for aerosol retrievals from the O_2 A band.

5.4 Cloud Filtering

As mentioned above, S5P methane retrievals are only performed for cloud free ground pixels. Here, the baseline is to use the VIIRS cloud mask re-gridded to the S5P spatial pixels for the NIR and SWIR band (work performed by RAL under ESA contract [RD57]). VIIRS provides 4 different 'confidence' levels for cloudiness ranging from 'confidently clear' to 'confidently cloudy'. The RAL VIIRS product provides a cloud mask for the S5P pixel itself, referred to as Inner Field Of View (IFOV), and for areas larger than the IFOV by a certain factor, referred to as the Outer Field Of Views (OFOVS).

For GOSAT, where cloud information is provided by the Cloud and Aerosol Imager (CAI), we found that an appropriate criterion is that 99% of an area equal 4 times that of a ground pixel should be confidently clear. We also use the VIIRS cirrus reflectance to detect scenes with high cirrus loading which are challenging for the retrieval algorithm. These scenes can be filtered a posteriori. In addition to VIIRS data we use a cloud filter based on S5P measurements itself. This cloud filter is used as a backup approach in case VIIRS data are not

available. For this purpose we use CH_4 and H_2O columns retrieved from weak and strong absorption bands in the SWIR channel under the assumption of a non-scattering atmosphere. For the CH_4 bands, we choose 2310-2315 nm (weak) and 2363-2373 nm (strong). The difference in columns retrieved from the weak and strong bands increases by increasing cloud optical thickness and/or cloud fraction. So, we filter out scenes for which this difference exceeds a certain threshold. In the error analysis section (Sect. 7.4) we determine appropriate values for these thresholds. The non-scattering retrieval results for the weak and strong CH_4 and H_2O absorption bands is expected as input for the CH_4 algorithm. They are produced by a pre-processing step of the S5P CO retrieval algorithm SICOR (see CO ATBD [RD58]). A similar cloud screening method was developed by the ACOS/OCO-2 team for cloud screening for GOSAT and OCO-2 CO_2 retrievals.

In addition we use the FRESCO (L2) apparent surface pressure retrieved under the assumption of no scattering by aerosols and clouds. The FRESCO algorithm is provided by KNMI and runs in the L1-2 processor. We compare the retrieved "apparent" surface pressure with the "true" surface pressure provided by ECMWF. The difference between the values is a measure for modification of the light path due to aerosols and clouds. If the difference exceeds a certain threshold the scene is flagged cloudy. The most important limitation of this cloud filter is that for some values of the surface albedo light path effects in the O_2 A band cancel out [RD59] while clouds are present. Another limitation is due to the ground pixel size of the S5P measurements, because it is required that a larger area than just the ground pixel under consideration is cloud free. Using a cloud filter from S5P itself, this means that also the full adjacent ground pixels should be cloud free. When using VIIRS, with its 1 km pixel size, one is much more flexible in defining the area that is required to be cloud free. Due to the limitations in cloud flagging from S5P itself, it is strongly preferred to use VIIRS for cloud and cirrus filtering. Doing cloud filtering using S5P itself should only be considered as a backup approach, but is fully implemented in the operational processor.

To summarize, the following checks for cloudiness and/or homogeneity are performed within **the baseline approach using VIIRS**:

- Fraction of 'confidently and probably clear' VIIRS pixels within S5P SWIR ground pixel > threshold T_1
- Fraction of 'confidently and probably clear' VIIRS pixels within S5P SWIR Outer Field Of View > threshold T_2 .
- Fraction of 'confidently and probably clear' VIIRS pixels within S5P NIR ground pixel > threshold T_3 .
- Fraction of 'confidently and probably clear' VIIRS pixels within S5P NIR Outer Field Of View > threshold T_4 .
- Average VIIRS cirrus reflectance within SWIR ground pixel < threshold T_5 .
- Average VIIRS cirrus reflectance within NIR ground pixel < threshold T_6 .
- The difference in VIIRS average cirrus reflectance between the SWIR ground pixel and the SWIR OFOV < threshold T_7 .

The additional cloud screening, and fall-back approach in case VIIRS data are not available consists of the following checks:

- The difference between the water column $[\text{H}_2\text{O}]_{\text{weak}}$ retrieved from the weak water band and the water column $[\text{H}_2\text{O}]_{\text{strong}}$ retrieved from the strong water band should be < threshold BT_1 .
- The difference between $[\text{CH}_4]_{\text{weak}}$ and $[\text{CH}_4]_{\text{strong}}$ < threshold BT_2 .
- The standard deviation of $[\text{H}_2\text{O}]_{\text{weak}} / [\text{H}_2\text{O}]_{\text{strong}}$ within the SWIR pixel and the 8 neighboring pixels < Threshold BT_3 .
- The standard deviation of $[\text{CH}_4]_{\text{weak}} / [\text{CH}_4]_{\text{strong}}$ within the SWIR pixel and the 8 neighboring pixels < Threshold BT_4 .
- The difference between the FRESCO apparent pressure and the ECMWF surface pressure < threshold BT_5 .
- The standard deviation of the FRESCO apparent surface pressure in the NIR pixel and the 8 surrounding pixels < threshold BT_6 .

5.5 SWIR-NIR Co-Registration

Due to the fact that the measurements in the NIR and SWIR spectral bands are performed by different instrument modules with different characteristics, the measurements in the two different bands are not performed at the same spatial grids. The SWIR pixel size is approximately $7 \times 7 \text{ km}^2$, while the NIR pixel size is $3.5 \times 7 \text{ km}^2$ (i.e. factor 2 smaller in the across track direction). On 6 August 2019, the spatial resolution was increased to 5.5 km in the along track direction. The **initial baseline** to deal with spatial co-registration is the most simple and robust one, namely to select the 2 NIR spatial pixels that have largest overlap with the SWIR $7 \times 7 \text{ km}^2$ spatial pixel. If the scattering properties are different for the SWIR and NIR ground pixel (constructed by 2 sub-pixels) this leads to errors in the retrieved $X\text{CH}_4$. Therefore, we use the VIIRS cirrus reflectance co-added to the SWIR and NIR ground pixels to identify situations where the cirrus contamination for the NIR and SWIR ground pixels deviate too much. The corresponding retrieval results are flagged as erroneous. In case VIIRS data are not available we use variability of the FRESCO (L2) retrieved apparent surface pressure to identify inhomogeneous scenes (looking at the NIR ground pixel and its neighbour pixels).

5.6 Bias correction

We perform a bias correction of the TROPOMI $X\text{CH}_4$ data based on the retrieved surface albedo to further improve the accuracy and the fitness for purpose of the TROPOMI CH_4 product. This type of correction is also performed for GOSAT CO_2 and CH_4 and OCO-2 CO_2 retrievals (e.g. [RD28, RD60, RD61]) to correct dependencies of the products on different parameters such as goodness of fit, surface albedo or aerosol parameters.

The comparison of TROPOMI and TCCON $X\text{CH}_4$ measurements shows a dependence of the bias (i.e. difference between TROPOMI and TCCON) on surface albedo retrieved in the SWIR spectral range. For low albedo values, TROPOMI $X\text{CH}_4$ strongly underestimates TCCON measurements, while for relatively high albedo values TROPOMI overestimates TCCON measurements. The comparison of TROPOMI $X\text{CH}_4$ with $X\text{CH}_4$ retrieved from measurements of GOSAT shows the same dependence of the bias with the retrieved surface albedo. For scenes with low albedo values, generally the retrieval's sensitivity is low due to the large measurement noise, therefore errors from unaccounted light path modification due to scattering processes can be more significant than for scenes with a relatively higher albedo.

To account for the albedo dependence, we apply an a posteriori bias correction to the retrieved $X\text{CH}_4$. In the operational algorithm version 01.xx few months after TROPOMI was operational, we applied a correction based on the comparison of TROPOMI $X\text{CH}_4$ with GOSAT retrievals. After more than two years of measurements, we have sufficient data to derive a new correction using only TROPOMI $X\text{CH}_4$ measurements. We use a similar approach to the "small area approximation" applied to OCO-2 [RD62], assuming a uniform $X\text{CH}_4$ distribution as a function of albedo in several regions. This approach makes the correction completely independent of any reference data (e.g. GOSAT, TCCON) that could introduce additional biases when applying the correction and does not allow for an independent verification of the correction.

The new correction is derived as follows [RD54]:

1. We select areas at several latitudes and longitudes throughout the globe, small enough so we can assume that $X\text{CH}_4$ does not vary, but large enough to cover scenes with a wide range of albedo values.
2. For each region we estimate a $X\text{CH}_4$ reference value for a surface albedo around 0.2 and then we calculate the ratio of the retrieved $X\text{CH}_4$ to the reference value to obtain the albedo dependence. The specific value for surface albedo is selected because $X\text{CH}_4$ retrieval errors are lower in the SWIR for that albedo range: errors because of unaccounted light path modifications due to scattering and surface albedo are minimal around a surface albedo of 0.2 (e.g. [RD28, RD10]).
3. We combine the albedo dependence for all the areas, we fit the curve using B-spline interpolation and least squares fitting.

Figure 5 shows the distribution of the reference to TROPOMI $X\text{CH}_4$ ratio for all the areas and the result of the B-spline fit [RD54]. We observe two distinct features: (1) the strong underestimation for low albedo values for which the B-spline fit corrects more strongly than the regular polynomial fit that was previously derived (2) an overestimation for high albedo values, not captured by TCCON due to the limited albedo range values but reported in the TROPOMI and GOSAT comparisons.

The correction applied to the retrieved $X\text{CH}_4$ can be expressed as:

$$X\text{CH}_{4i}^{\text{corr}} = X\text{CH}_{4i} \cdot f(A_{si}). \quad (53)$$

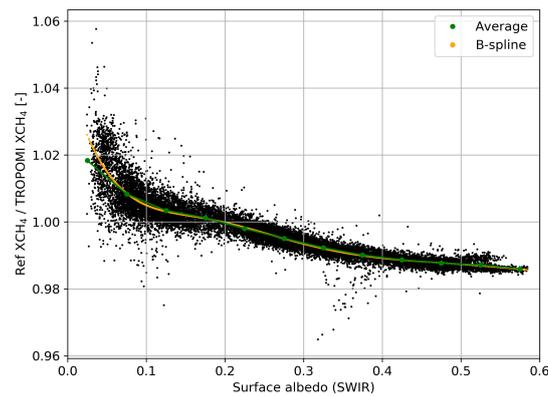


Figure 5: Ratio of reference XCH_4 to TROPOMI XCH_4 as a function of the retrieved surface albedo as explained in step 3 in the derivation of the bias correction. Green dots show the average ratio in 0.05 albedo bins and orange line shows the B-spline fit used to derive the bias correction. Data is averaged from 1 Jan 2018 until 31 Dec 2019 in a $0.1^\circ \times 0.1^\circ$ grid

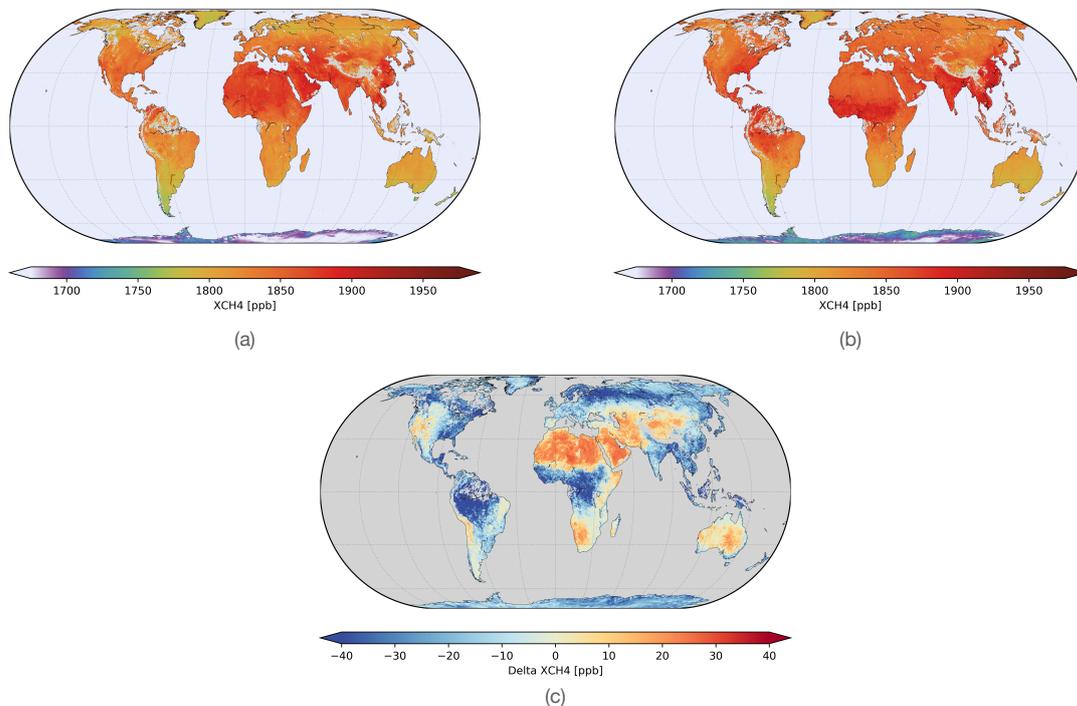


Figure 6: (a) Global TROPOMI XCH_4 distribution before correction, (b) after correction and (c) their difference ($XCH_4 - XCH_4^{corr}$) for 2019 averaged in a cylindrical equal-area grid with $0.3^\circ \times 0.5^\circ$ resolution at the Equator.

The correction function f depends on the retrieved surface albedo A_s in the SWIR spectral range at each pixel i .

Figure 6 shows the global distribution of XCH_4 before and after applying the correction [RD54]. Distinctive features that correspond with low and high surface albedo areas are visible in the difference map. After correction, for example, the XCH_4 underestimation for low albedo values (e.g. over high latitudes over Canada and Russia) is corrected. Similarly, the XCH_4 overestimation for high albedo values over desert areas like Sahara is accounted for in the correction. The change in XCH_4 induced by the bias correction is in the range of 2%, in agreement with the errors observed in the TCCON comparison.

As the correction is derived using only TROPOMI XCH_4 data, the comparison with TCCON and GOSAT is an independent verification of the approach. The validation with TCCON shows a reduction of 5.9 ppb

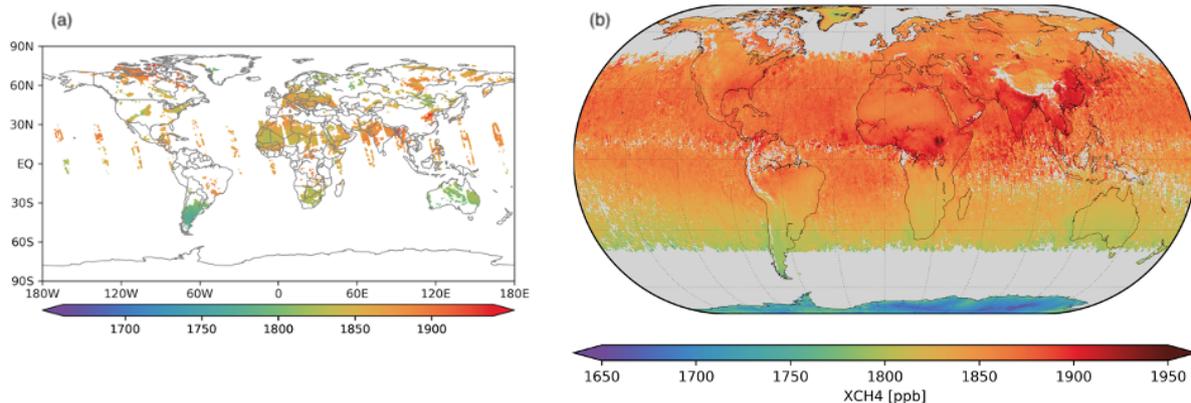


Figure 7: (a) Daily global TROPOMI XCH₄ coverage, and (b) global XCH₄ for 2020 averaged in a cylindrical equal-area grid with 0.3° x 0.5° resolution at the Equator.

(50%) in the station-to-station variability and of 13.6 ppb in the bias due to the albedo correction [RD54]. The dependence of the bias on surface albedo is removed and the dependence on other parameters remains negligible. The comparison with GOSAT measurements shows that bias dependence on albedo is removed after applying the correction, which reduces by 4 ppb the scatter of the differences in XCH₄ measured by the two satellites.

5.7 Ocean measurements over sun-glint geometries

Since processor version 2.3.1, ocean measurements over sun-glint geometries have been included in the processing. The sun-glint angle is defined as follows:

$$\cos(\text{SGA}) = ((\cos(\theta_0 + \theta_v) + \cos(\theta_0 - \theta_v) + (\cos(\theta_0 + \theta_v) - \cos(\theta_0 - \theta_v)) \cdot \cos(\phi_0 - \phi_v)) \cdot 0.5 \quad (54)$$

where θ_0 and θ_v are the solar and viewing zenith angle, respectively (absolute values) and ϕ_0 and ϕ_v are the solar and viewing azimuth angles.

When SGA is lower than 20°, the pixel is considered to fall under the sun glint area and the retrieval is performed for that pixel. An example of a typical daily cloud-free coverage of TROPOMI XCH₄ measurements is shown in Fig. 7(a).

Ocean surface reflection is modeled using the wind-speed-dependent Cox-and-Munk reflection model [RD44], which might result in negative retrieved surface reflectance values over ocean. The surface reflection model over ocean is different to that over land, so the physical meaning of the retrieved surface albedo over land and over ocean is different.

In order to have an homogeneous XCH₄ distribution over land and ocean, we apply a correction to the ocean data based on the XCH₄ distribution. Ocean measurements do not show any clear dependence on surface reflectivity, differently to what was found for measurements over land, so we apply a constant global correction factor to ocean measurements. To calculate this correction factor, we estimate the median XCH₄ over land (after correction) and over ocean and estimate the ratio between the two. We estimate the ratio separately for the Northern and Southern hemisphere in order to account for the fact that most of the XCH₄ sources are located in the Northern hemisphere. The estimated correction factor for ocean is 1.00385; the correction applied to the retrieved XCH₄ over ocean is then:

$$\text{XCH}_{4oc}^{\text{corr}} = \text{XCH}_{4oc} \cdot 1.00385. \quad (55)$$

5.8 Surface reflectance spectral dependence

Since processor version 2.4.0, the surface reflectance spectral dependence is taken into account using a 3rd order polynomial. This is particularly important to take into account surface spectral features of specific surfaces for which a first order spectral dependence might not be representative.

5.9 Algorithm overview

5.9.1 Required input

Required dynamic input data are listed in Table 3. The methane algorithm expects all dynamic inputs per ground pixel and hence they need to be pre-processed accordingly. The required static input data are listed in Table 4.

Table 3: Dynamic input data for S5P XCH4 algorithm.

Data	Symbol	Units	Source	Pre-Process needs	If not available
S5P level 1b Earth radiance SWIR band	I	mol/s/m ² /nm/sr	S5P level 0-1b product	per ground pixel	no retrieval
S5P level 1b Earth radiance NIR band	I	mol/s/m ² /nm/sr	S5P level 0-1b product	per ground pixel, spatially co-located with SWIR ground pixel	no retrieval
S5P level 1b Solar irradiance SWIR band	$F_{0,meas}$	mol/s/m ² /nm	S5P level 0-1b product		use previous measurement
S5P level 1b solar irradiance NIR band	$F_{0,meas}$	mol/s/m ² /nm	S5P level 0-1b product		use previous measurement
latitude	lat	degree	S5P level 0-1b product		no retrieval
longitude	lon	degree	S5P level 0-1b product		no retrieval
solar zenith angle	θ_0	degree	S5P level 0-1b product		no retrieval
viewing zenith angle - SWIR band	θ_v	degree	S5P level 0-1b product		no retrieval
relative azimuth angle –SWIR band	ϕ	degree	S5P level 0-1b product		no retrieval
viewing zenith angle - NIR band	θ_v	degree	S5P level 0-1b product		no retrieval
relative azimuth angle - NIR band	ϕ	degree	S5P level 0-1b product		no retrieval
cloud fraction for S5P SWIR and NIR ground pixel			VIIRS / RAL algorithm	per ground pixel, spatially co-located with SWIR and NIR ground pixel.	use FRESCO apparent pressure and backup cloud screening
cirrus reflectance for S5P SWIR and NIR ground pixel		[-]	VIIRS / RAL algorithm	per ground pixel, spatially co-located with SWIR and NIR ground pixel.	use FRESCO apparent pressure and backup cloud screening
apparent pressure for ground pixel and surrounding		Pa	FRESCO (L2)	find corresponding and neighbouring ground pixel	filter based on retrieved scattering parameters

Continued on next page

Table 3 – Continued from previous page

Data	Symbol	Units	Source	Pre-Process needs	If not available
non scattering retrieval results for weak and strong CH ₄ and water bands for ground pixel and surroundings	[CH ₄] _{weak} [CH ₄] _{strong} [H ₂ O] _{weak} [H ₂ O] _{strong}	mol cm ⁻²	CO algorithm SICOR (L2)	Find corresponding and neighbouring ground pixel	filter based on retrieved scattering parameters
fluorescence emission at 755 nm	$F_{s,755nm}^{surf}$	mol s ⁻¹ m ⁻² nm ⁻¹ sr ⁻¹	FRESCO algorithm		Assume no fluorescence signal.
atmospheric temperature profile	T	K	ECMWF	interpolated in space and time to SWIR ground pixel / time	no retrieval
atmospheric pressure profile hybrid coefficients	P		ECMWF	calculate pressure profile Interpolated in space and time to SWIR ground pixel / time	no retrieval
vertical profile of specific humidity			ECMWF	interpolated in space and time to SWIR ground pixel / time	use closest in space
surface pressure	P_{surf}	Pa	ECMWF	interpolated in space and time to SWIR measurement	no retrieval
altitude corresponding to ECMWF surface pressure	$z_{surf,ECMWF}$	m	ECMWF	Interpolated in space and time to SWIR measurement	no retrieval
CH ₄ a priori vertical profile		mol mol ⁻¹	TM5 / SRON	interpolated in space and time to SWIR ground pixel / time	use last available input (in time)
surface wind speed in 2 directions		m/s	ECMWF	interpolated in space and time to SWIR measurement, calculate total wind speed	use last available input (in time)
CO a priori vertical profile		mol mol ⁻¹	TM5 / SRON	interpolated in space and time to SWIR ground pixel / time	use last available input (in time)
mean pixel altitude (NIR and SWIR)		m	GMTED2010	interpolated in space to SWIR ground pixel	no retrieval
standard deviation of altitude within one ground pixel (NIR and SWIR)		m	GMTED2010	Interpolated in space to SWIR ground pixel	no retrieval

Continued on next page

Table 3 – Continued from previous page

Data	Symbol	Units	Source	Pre-Process needs	If not available
land / water mask			GMTED2010 (GTOPO30 not accurate enough)	for SWIR and NIR ground pixel	flag

Table 4: Static input data for S5P XCH4 algorithm.

Data	Symbol	Units	Source	Pre-Process needs
absorption cross sections O ₂	σ_{O_2}	cm ²	[RD37]	calculate cross-sections on fixed wavenumber, pressure, temperature grid (SRON)
absorption cross sections CH ₄	σ_{CH_4}	cm ²	[RD33] [RD34]	same as O ₂
absorption cross sections H ₂ O	σ_{H_2O}	cm ²	[RD33] [RD34]	same as O ₂
model high resolution solar spectrum	F_0	mol/s/m ² /nm	TBD	TBD
absorption cross sections CO	σ_{CO}	cm ²	[RD33] [RD34]	same as O ₂
LUT with optical aerosol properties			[RD42]	
LUT with cirrus optical properties			[RD63]	
Instrument Spectral Response Function (ISRF) NIR & SWIR	S		Calibration Key Data	
retrieval input settings			SRON	

5.9.2 Algorithm implementation

Figure 8 gives a high level overview of the processing for the CH₄ product. The static input (GMTED Digital Elevation MAP, retrieval input settings, cross section lookup tables, and lookup tables with aerosol/cirrus optical properties) needs to be read in only once when the processing starts. Here, the cross section lookup tables are provided in NetCDF format. The dynamical input consists of the NIR and SWIR level 1b data (e.g. per orbit), the VIIRS cloud data (co-located with the S5P level 1b measurements), the S5P FRESCO level-2 cloud data (apparent pressure), and ECMWF data, TM5 CH₄ and CO data (e.g. per day). From the dynamical input data and the GMTED2010 DEM data per ground pixel are being prepared in a pre-processing step. This pre-processing step consists of interpolating the ECMWF data and GMTED2010 DEM data to the coordinates of each S5P ground pixel (assumed to be performed by general KNMI software as this is the same for all S5P products) and co-registration of level 1b data between NIR and SWIR ground pixels. It is assumed that retrievals for individual ground pixels are distributed over different processing threads, where the retrieval per ground pixel is being performed by Fortran90 modules.

Figure 9 shows an overview of the retrieval per ground pixel (see Figure 8), making use of the pre-processed input data for each ground pixel. The first step is to perform the a priori filtering based on cloud data provided

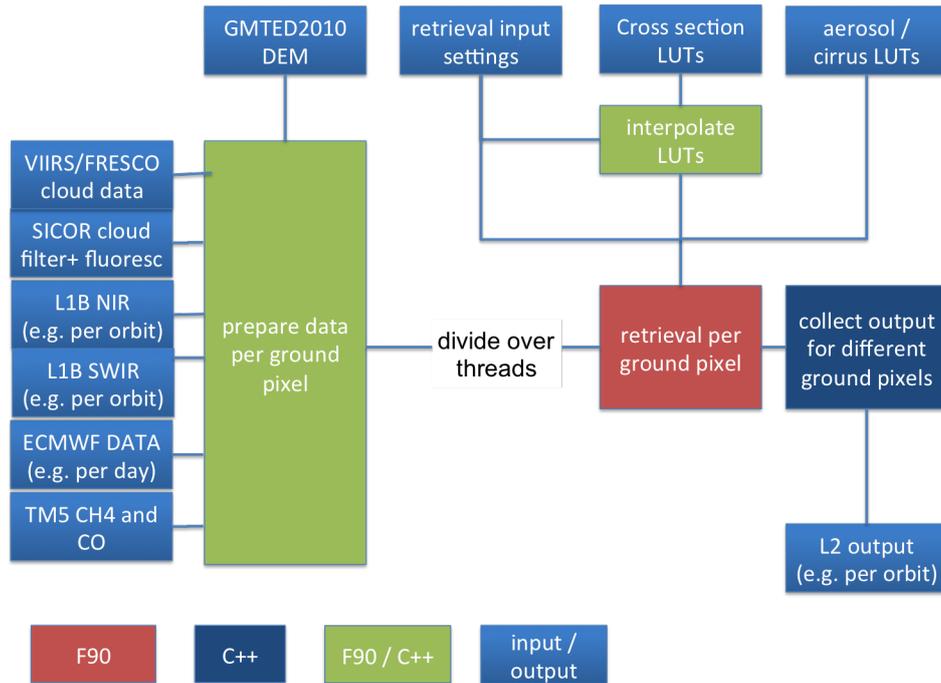


Figure 8: High level overview of XCH₄ processing scheme.

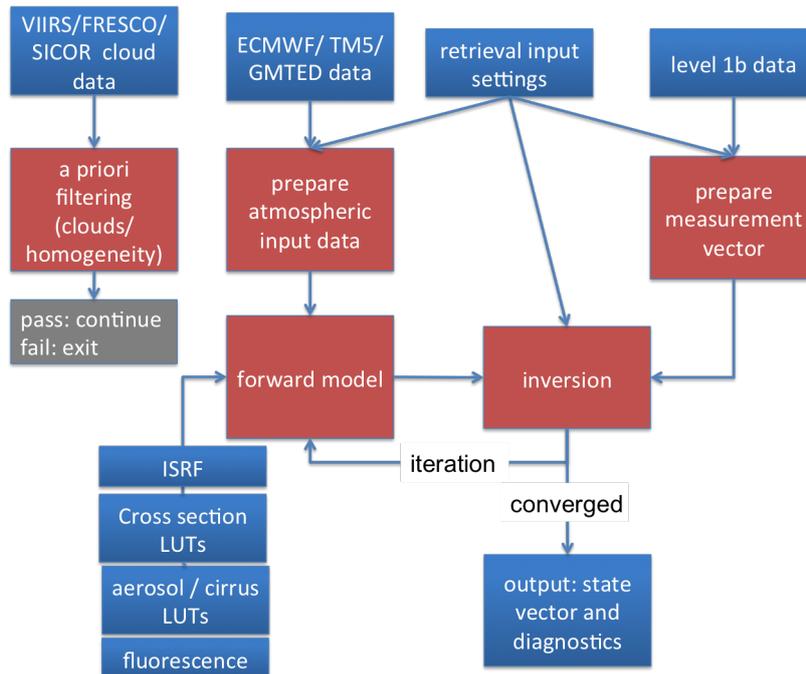


Figure 9: Overview of processing per ground pixel.

by VIIRS. If the ground pixel under consideration contains too much cloud contamination and if the variation in cirrus reflectance between the NIR and SWIR spatial ground pixel is too large the retrieval for that ground pixel is not performed. In case VIIRS cloud data is not available homogeneity is being checked by means of the FRESCO apparent pressure for the ground pixel under consideration and its neighbouring pixels. Furthermore, a cloud filtering is being performed based on the retrieved O₂ column for a model atmosphere without aerosol and cloud scattering. When this O₂ column deviates too much from the value predicted by ECMWF the pixel

is considered cloudy.

The next step is to prepare the atmospheric input data to be used in the retrieval, i.e. interpolation to the vertical grid of the retrieval forward model and filling the values of the initial state vector for the retrieval (CH_4 profile, CO and H_2O column, aerosol parameters, surface albedo). Also, the measurement vector is being prepared from the level 1b input data. (selection of measurements within specified spectral range). Then the iterative scheme starts. A forward model calculation is performed for the first guess state vector and the inversion is performed based on the forward calculation, the measurement vector, and the prior information. This procedure is repeated till convergence has been reached.

6 Input-Output file description

6.1 High level data product description

The output of the XCH_4 algorithm is specified in Table 5.

Table 5: Contents of the output product.

Name/ Data	Symbol	Units	Description	Data type	Dimension
time		s	time of measurements	float	1
latitude, latitude_bounds	lat	degree	SWIR pixel latitude (center & corners)	float	1,4
longitude, longitude_bounds	lon	degree	SWIR pixel longitude (center & corners)	float	1,4
solar_zenith_angle	θ_0	degree	solar zenith angle at pixel center	float	1
viewing_zenith_angle	θ_v	degree	viewing zenith angle at pixel center	float	1
methane_mixing_ratio	XCH_4	ppb	column averaged dry air mixing ratio of CH_4	float	1
methane_mixing_ratio_bias_corrected	XCH_4	ppb	bias corrected XCH_4	float	1
methane_mixing_ratio_precision	ΔXCH_4	ppb	precision of XCH_4	float	1
qa_value			QA value for CH_4	int	1
number_of_iterations			number of iterations for the CH_4 retrieval	int	1
column_averaging_kernel	a		column averaging kernel ¹	float	12
altitude_levels	z	m	height levels of retrieval grid	float	13
surface_pressure	p_{surf}	hPa	surface pressure	float	1
pressure_interval	dp	Pa	pressure difference between levels in the retrieval	float	1
dry_air_subcolumns	DVAIR	$mol\ m^{-2}$	dry air layer sub columns	float	12
methane_profile_apriori		$mol\ m^{-2}$	a priori CH_4 profile in layer sub-columns	float	12
carbonmonoxide_total_column	VCO	$mol\ m^{-2}$	CO total column (not official product)	float	1
carbonmonoxide_total_column_precision	ΔVCO	$mol\ m^{-2}$	error on VCO	float	1
water_total_column	VH2O	$mol\ m^{-2}$	H_2O total column	float	1
water_tota_column_precision	ΔVH_2O	$mol\ m^{-2}$	error on VH2O	float	1
aerosol_size	α		aerosol size parameter	float	1
aerosol_size_precision	$\Delta\alpha$		error on aerosol size parameter	float	1
aerosol_number_column	N_{aer}	m^{-2}	aerosol number column	float	1
aerosol_number_column_precision	ΔN_{aer}	m^{-2}	error on aerosol_amount	float	1

¹ works on subcolumns

Continued on next page

Table 5 – Continued from previous page

Name/ Data	Symbol	Units	description	Data type	dimension
aerosol_mid_altitude	z_{aer}	m	central height of aerosol altitude distribution	float	1
aerosol_mid_altitude_precision	Δz_{aer}	m	error on aerosol height z_{aer}	float	1
surface_albedo_SWIR	A_{SWIR}		surface albedo in SWIR band	float	1
surface_albedo_SWIR_precision	ΔA_{SWIR}		precision of SWIR albedo	float	1
surface_albedo_NIR	A_{NIR}		surface albedo in NIR band	float	1
surface_albedo_NIR_precision	ΔA_{NIR}		precision of the NIR albedo	float	1
aerosol_optical_thickness_SWIR	τ	τ_{aer}	AOT in SWIR band	float	1
aerosol_optical_thickness_NIR	τ	τ_{aer}	AOT in NIR band	float	1
chi_square	χ^2		chi squared of fit in both bands	float	1
chi_square_swir			chi squared of fit in SWIR band	float	1
chi_square_nir			chi squared of fit in NIR band	float	1
degrees_of_freedom	DFS		Degrees of Freedom for Signal	float	1
degrees_of_freedom_methane			DFS for CH ₄ profile	float	1
degrees_of_freedom_aerosol			DFS for aerosol parameters	float	1
cloud_fraction_VIIRS_NIR_IFOV			Cloud fraction from VIIRS data in the NIR channel for the instantaneous field of view (band 6)	float	1
cloud_fraction_VIIRS_NIR_OFOVa			Cloud fraction from VIIRS data in the SWIR channel for the 10% upscaled field of view	float	1
cloud_fraction_VIIRS_NIR_OFOVb			Cloud fraction from VIIRS data in the SWIR channel for the 50% upscaled field of view	float	1
cloud_fraction_VIIRS_NIR_OFOVc			Cloud fraction from VIIRS data in the SWIR channel for the 100% upscaled field of view	float	1
cloud_fraction_VIIRS_SWIR_IFOV			Cloud fraction from VIIRS data in the NIR channel for the instantaneous field of view (band 6)	float	1

Continued on next page

Table 5 – Continued from previous page

Name/ Data	Symbol	Units	description	Data type	dimension
cloud_fraction_VIIRS_SWIR_OFOVa			Cloud fraction from VIIRS data in the SWIR channel for the 10% upscaled field of view	float	1
cloud_fraction_VIIRS_SWIR_OFOVb			Cloud fraction from VIIRS data in the SWIR channel for the 50% upscaled field of view	float	1
cloud_fraction_VIIRS_SWIR_OFOVc			Cloud fraction from VIIRS data in the SWIR channel for the 100% upscaled field of view	float	1
reflectance_cirrus_VIIRS_SWIR			Cirrus reflectance from VIIRS for the SWIR ground pixel	float	1
reflectance_cirrus_VIIRS_NIR			Cirrus reflectance from VIIRS for the NIR ground pixel	float	1
methane_weak_twoband_total_column	$[\text{CH}_4]_{\text{weak}}$	mol m^{-2}	SICOR retrieved CH_4 column from weak absorption band	float	1
methane_strong_twoband_total_column	$[\text{CH}_4]_{\text{strong}}$	mol m^{-2}	SICOR retrieved CH_4 column from strong absorption band	float	1
water_weak_twoband_total_column	$[\text{H}_2\text{O}]_{\text{weak}}$	mol m^{-2}	SICOR retrieved H_2O column from weak absorption band	float	1
water_strong_twoband_total_column	$[\text{H}_2\text{O}]_{\text{strong}}$	mol m^{-2}	SICOR retrieved H_2O column from strong absorption band	float	1
methane_ratio_weak_strong_standard_deviation			Standard deviation of ratio of CH_4 columns from weak and strong band for 9 pixels (ground pixel plus 8 neighbours)	float	1
water_ratio_weak_strong_standard_deviation			Standard deviation of ratio of H_2O columns from weak and strong band for 9 pixels (ground pixel plus 8 neighbours)	float	1
apparent_scene_pressure		hPa	apparent surface pressure of FRESCO for NIR band	float	1
apparent_scene_pressure_standard_deviation		hPa	standard deviation of PSURF_FRESCO for 9 pixels (center and 8 neighbors)	float	1
surface_altitude		m	surface elevation of S5P SWIR pixel	float	1

Continued on next page

Table 5 – Continued from previous page

Name/ Data	Symbol	Units	description	Data type	dimension
surface_altitude_precision		m	standard deviation of surface elevation within SWIR pixel	float	1
fluorescence	$F_{s,755nm}^{surf}$	mol/s/cm ² /nm	retrieved fluorescence signal at 755 nm	float	1
fluorescence_apriori		mol/s/cm ² /nm	apriori fluorescence signal at 755 nm	float	1
wavelength_calibration_offset_swir		nm	fitted correction to wavelength calibration	float	1
wavelength_calibration_offset_nir		nm	fitted correction to wavelength calibration	float	1

The output product consists of the retrieved methane column and a row vector referred to as the column averaging kernel **A**. The column averaging kernel describes how the retrieved column relates to the true profile and should be used in validation exercises (when possible) or use of the product in source/sink inverse modeling. The output product also contains altitude levels of the layer interfaces to which the column averaging kernel corresponds. Additional output for Level-2 data products: viewing geometry, precision of retrieved methane, residuals of the fit, quality flags (cloudiness, terrain roughness etc.) and retrieved albedo and aerosol properties. The latter properties are required for a posteriori filtering and for estimation of total retrieval error.

6.2 Data selection approach

First, only select those data that meet the following criteria (a priori filtering in Figure 9). The pre-filtering determines for each measured pixel in the S5P swath if:

- Fraction of 'confidently and probably clear' VIIRS pixels > 98% to assure cloud free pixels.
- Spectrum intensity above threshold level (prefilters for ocean and snow). There is a lower threshold for SWIR signal-to-noise ratio (SNR) of 50.
- Above land.
- Cloud / cirrus coverage and variation: perform cloud checks as described in section 5.4.
- SZA < 75 degree. For larger angles, the plane-parallel treatment of scattering processes introduces too large errors, and the signal-to-noise-ratio is generally insufficient for a retrieval.
- Instrument zenith angle below 60 degree, i.e. preference is given to those pixels in the swath that are closer to instrument nadir.
- Mountainous areas are excluded; for that the pre-filtering filters out all pixels with a surface roughness (i.e. standard deviation of altitude) larger than 100 m (derived from GMTED2010 data).

After the processing, there a post-filter applied to assure the highest quality. The specifics of the post filter are specified in the Product Readme File [RD64].

6.3 Treatment of Corrupted data

A pixel mask (good or bad pixels) is being determined per orbit as part of the in-flight calibration. The methane retrieval for a certain ground pixel starts only if the number of good pixels is larger than a certain threshold (70%). As an additional safety check for corrupted / unphysical data we quit the iterative retrieval in the first iteration step if the χ^2 difference between the measurement and forward model calculation (based on the first guess parameters) exceeds a certain threshold.

6.4 Timeliness

The methane retrieval does not take place in the near-real-time data stream. All of the necessary auxiliary information either exists already today or is delivered by the S5P housekeeping data. The only critical issue is the availability of VIIRS information on cloud coverage as soon as possible after a S5P overpass.

7 Error analysis

The error analysis presented in this section was performed pre-launch. First, the performance of the reference retrieval for a global ensemble of synthetic measurements is presented in Sect. 7.1. Thereafter, the impact of model and instrument errors on the retrieved XCH_4 is discussed in Sect. 7.2 and 7.3, respectively. It should be noted that the effect of chlorophyll-fluorescence on the measured radiances in the NIR band is not taken into account in this. In analogy with the study by Frankenberg et al. (2012) [RD49], who studied the effect of fluorescence on XCO_2 retrievals from GOSAT and OCO-2, we expect this error to be smaller than the other error sources discussed here.

7.1 Simulation of geophysical test cases

In order to quantify the retrieval error on the XCH_4 induced by the presence of cirrus and aerosol, an ensemble of 9030 simulated measurements has been generated that covers the range of scenes that are likely encountered by the TROPOMI instrument. The details how these synthetic spectra were created are explained in the Appendix. Here, we present the performance of the algorithm when retrieving them.

The term “retrieval error” in this section describes the forward model error $\mathbf{G}e_F$ (caused by approximate treatment of aerosols) divided by the true XCH_4 . The error due to instrument noise ($\mathbf{G}e_F$) is subtracted from the retrieved state vector. The retrieval noise is evaluated separately in Sect. 7.3.1.

The world map in Figure 10 shows the forward model error of XCH_4 for the reference retrieval with three aerosol retrieval parameters. The relative error is largest for scenarios with a high aerosol load over bright surfaces (red pixels at the Sahara in summer time), high cirrus load over dark surfaces (e.g. tropical Africa), and/or with a large SZA (Northern Eurasia in January).

From the simulated retrievals we found that the information content of the average spectrum is sufficient to retrieve the three aerosol parameters in our algorithm: The total number of aerosols (N_{aer}), the height of the aerosol layer (h) and the exponent of the power-law size-distribution (α). Here, the information on size comes from the spectral dependence between the NIR and SWIR band, and information on amount and height comes from the large range in absorption optical thickness in both bands. The same selection was found to be optimal for GOSAT data.

The retrieval errors in Figure 10 are in particular large if there is a layer at large altitude with large AOT and with large particles (i.e. small value of α). Therefore, we filtered the converged retrievals a posteriori by rejecting all retrievals for which $f = AOT_{SWIR} \times \frac{h}{\alpha} < 110$. (This relation was found empirically.) Additionally, we also requested that the retrieved surface albedo in the SWIR band should be > 0.02 .

Figure 11 shows the XCH_4 error as a function of the retrieved SWIR albedo (left panel) and aerosol filter parameter f (right panel). It can be seen that for cases with low SWIR albedo and high f , the error scatter increases. About 50% of all retrievals converged and passed the a posteriori filters. Figure 12 shows the a posteriori filtered world map.

Figure 13 and Table 6 summarizes the error characteristics of the reference retrieval with three aerosol parameters, without (Fig. 10) and with (Figure 12) the two a posteriori filters for f and SWIR albedo.

Table 6: Performance of the reference retrieval for the ensemble of 9030 synthetic measurements.

	unfiltered	filtered
mean CPU time per retrieval	7.6 s	
success rate	95% (converged)	61% (converged and filtered)
RMS forward model error	1.1%	0.45%
RMS retrieval noise (due to instrument noise)	0.77%	0.37%
fraction of retrievals with $ XCH_4 \text{ error} < 1\%$	86%	96%

7.2 Model errors

7.2.1 Spectroscopic data

Galli et al. 2012 [RD65] investigated if the absorption features of the three relevant molecules CH_4 , CO , and H_2O in the SWIR range were adequately known by inverting ground-based Fourier transform spectrometers (FTS) measurements obtained by the TCCON group. The TCCON FTS observe direct sun light, therefore all

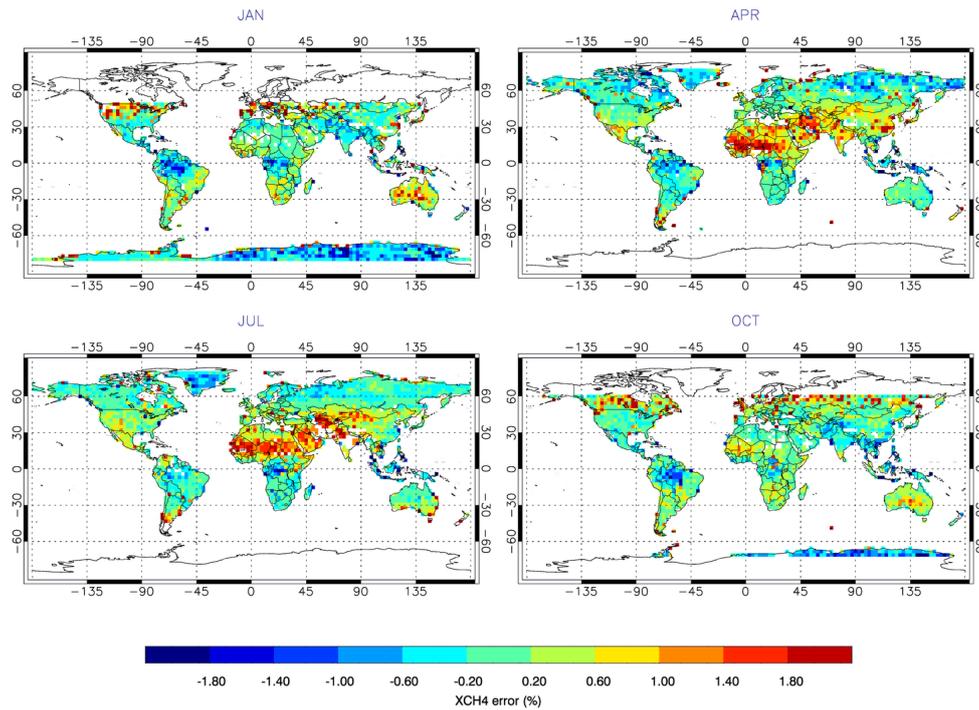


Figure 10: Retrieval forward model error of XCH_4 for the (unfiltered) reference retrieval. White colors indicate retrievals did not converge, or synthetic spectra were not calculated (oceans), or SZA > 70 degree. The XCH_4 error is defined as the difference between the retrieved and the true value and relative values are calculated with respect to the true XCH_4 .

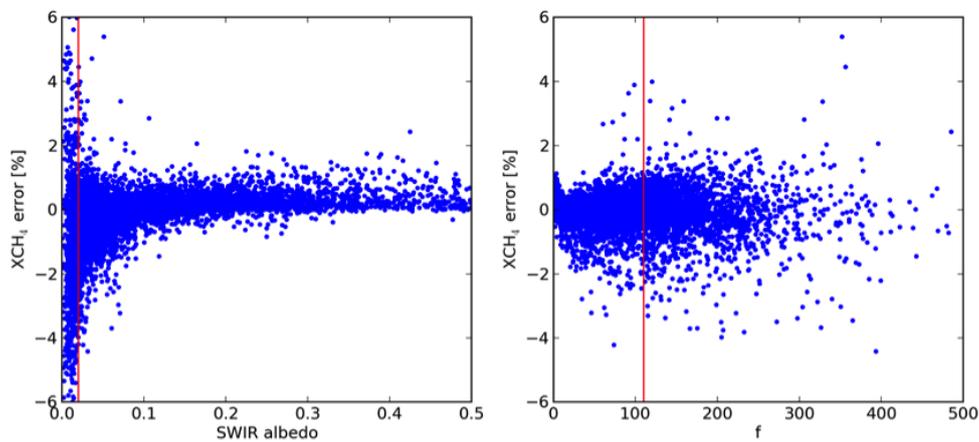


Figure 11: Forward model error of XCH_4 plotted against retrieved SWIR albedo (left panel) and aerosol filter parameter f (right panel). The vertical lines give the cut-offs for the a posteriori filters. Retrievals with SWIR albedo < 0.02 are already filtered out in the right panel.

scattering effects can be neglected, and only molecular absorption needs to be included in the retrieval forward model. On the other hand, the TCCON FTS measure a broad spectral range in the NIR and SWIR at a 30 times higher spectral resolution than the TROPOMI SWIR channel. This makes the TCCON spectra ideally suited to test spectroscopic line lists.

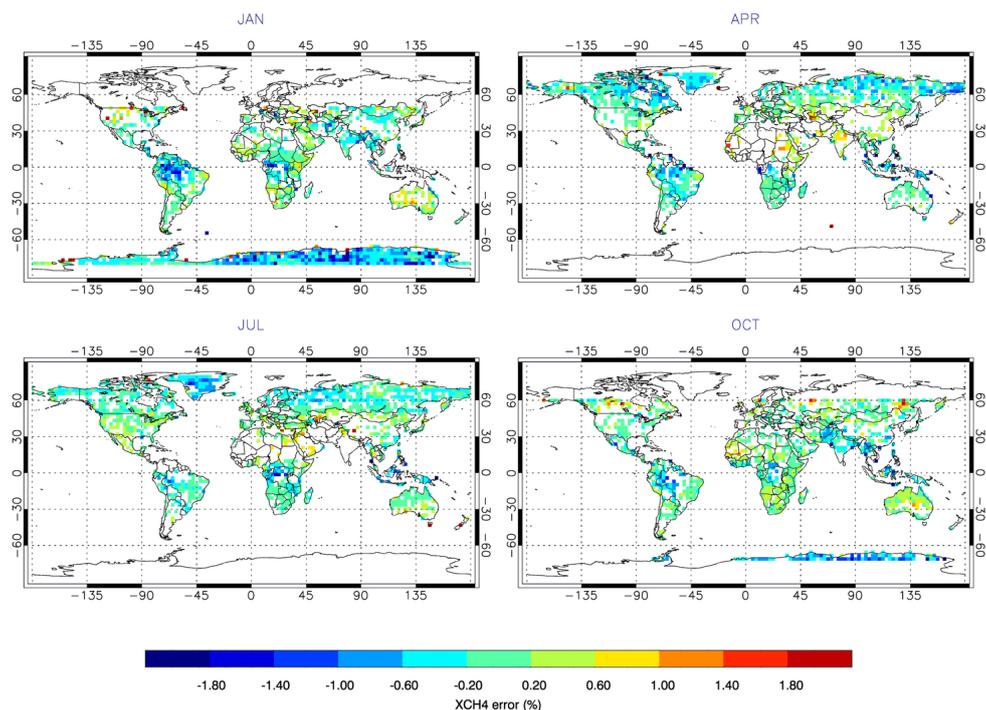


Figure 12: Forward model error of XCH_4 after applying the a posteriori filter: SWIR albedo > 0.02 and aerosol load not too high ($f < 110$).

Combining the inverse algorithm presented by Butz et al. 2012 [RD66] with a simple forward model without scattering, Galli et al. 2012 [RD65] retrieved CH_4 , H_2O , and CO from various spectral ranges at different spectral resolutions. Most relevant to this document is the comparison between the CH_4 columns retrieved from the TROPOMI range at the TROPOMI resolution and the CH_4 results obtained from the $1.6\ \mu m$ region at the higher TCCON resolution. The CH_4 columns were reproduced to an accuracy of 0.3% apart from a constant bias of 1%. No strong correlation with H_2O abundance was found. We caution, however, that the intrinsic retrieval accuracy of CH_4 from TCCON spectra is a few tenths of a percent. This means only spectroscopic deficiencies could be identified that result in column errors with the same order of magnitude as the intended accuracy for the S5P mission. When surface pressure was retrieved as well, Galli et al., 2012 [RD65] found an incorrect pressure-dependence at the TROPOMI SWIR range, both for CH_4 and for H_2O . This implies that not only the H_2O but also the CH_4 line list (in particular the pressure broadening and pressure shift parameters) is less accurate at the TROPOMI range than around $1.6\ \mu m$.

7.2.2 Temperature profile

An erroneous temperature profile affects the retrieval performance through the temperature dependence of the cross-sections. To investigate this, the mean temperature profile per latitude is used in the retrieval forward model instead of the true temperature profile. This is an overestimation of the temperature error expected from ECMWF and should hence be considered as a worst case scenario. The algorithm allows the option to fit a temperature offset and the performance is evaluated with this option switched on and off.

In Figure 14 the error characteristics are compared to the reference retrieval with a posteriori filters. The reduction in successful cases is mainly due to non-convergence. Fitting a temperature offset can to some extent compensate for an error in the temperature profile.

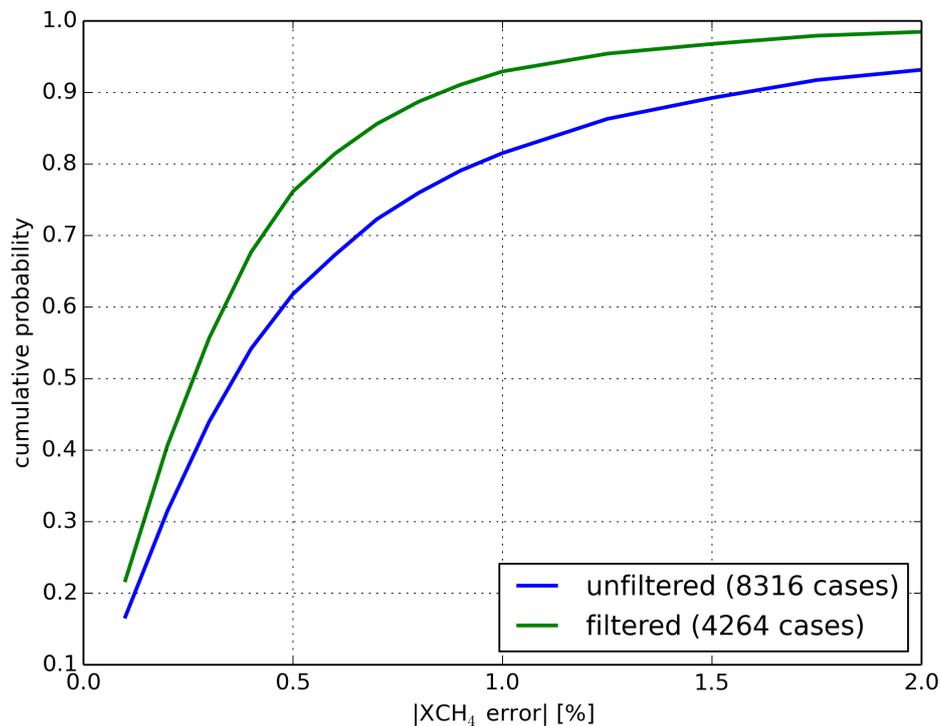


Figure 13: Cumulative probability distribution of the absolute XCH_4 retrieval error for the reference retrieval that includes scattering of aerosols (green: all converging retrievals, blue: a posteriori filters applied to results).

7.2.3 Surface Pressure

An erroneous pressure profile affects the retrieval of XCH_4 in two ways: first of all, through pressure dependence of the cross-sections and, secondly, through the retrieved air column that is used to convert the CH_4 total column to the dry air mixing ratio, XCH_4 . The latter introduces a retrieval error of the same magnitude as the pressure error. The algorithm allows the option to fit the surface pressure. To investigate the performance of this optional fit, the prior pressure profile is perturbed with a scaling factor up to $\pm 1\%$. Figure 15 and Table 7 show the results with and without fitting the surface pressure. In case the pressure is fitted the forward model error is the same as for the reference case. However, it should be noted that fitting the surface pressure from real measurements may be difficult. For example, in our retrievals from GOSAT, with its much higher spectral resolution, retrieval of surface pressure from real measurements is not possible. From the simulations where we do not fit the surface pressure we conclude that for surface pressure errors up to 0.2% (roughly 2 hPa) the effect on the retrieved XCH_4 is $< 0.2\%$. It is expected that the ECMWF pressure together with the GMTED2010 Digital Elevation map allow to achieve this accuracy.

7.2.4 Absorber profiles

CH₄ profile

A robust retrieval algorithm should be able to solve the inverse problem regardless of its first guess information. To test this, a normalized mean CH_4 profile per latitude is taken as first guess. We normalize the mean profile such that the prior total column is a factor times the true column. In this way, the effect of an error on the profile can be distinguished from an error on the column. The CH_4 column error is varied up to $\pm 2\%$. Figure 16 shows reassuringly that the first guess CH_4 profile has a negligible influence on the retrievals.

H₂O profile

The error on the prior H_2O profile is established in the same way as for CH_4 , i.e. by taking a normalized mean profile per latitude. Note that for H_2O , there is an additional (minor) influence on the retrieval of XCH_4 through the dry air column. The H_2O column error is varied up to $\pm 10\%$. Figure 16 shows that the prior H_2O

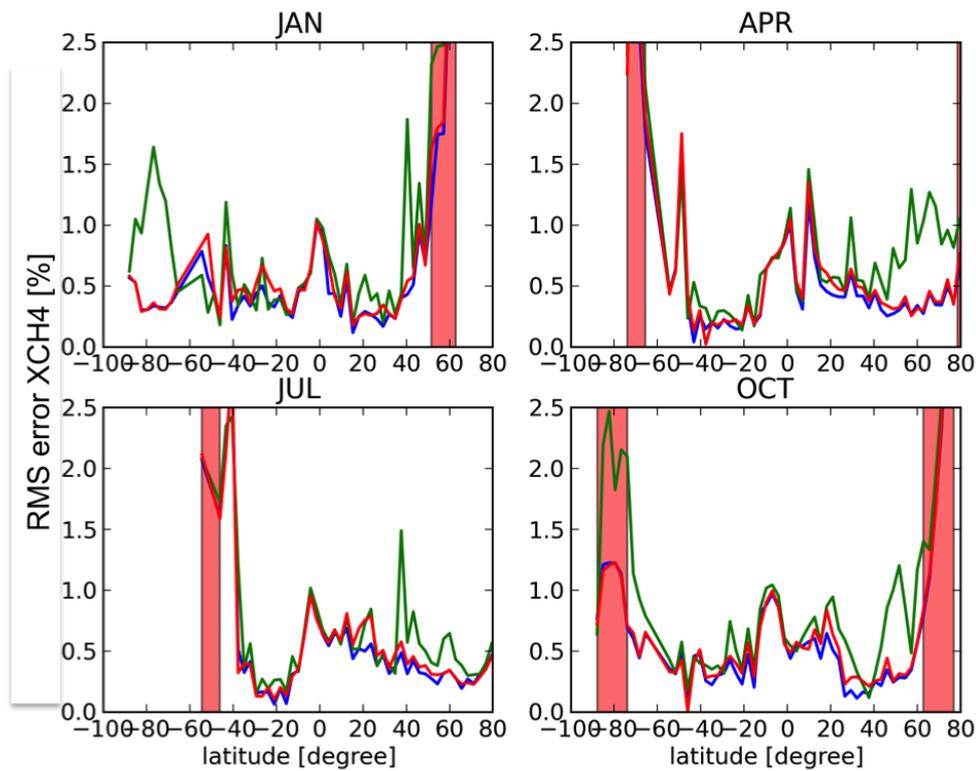


Figure 14: Sensitivity of XCH_4 retrieval to error in temperature profile. Upper panel: Cumulative probability distribution of the absolute XCH_4 retrieval error for the reference retrieval (blue), retrieval of ensemble with mean temperature profile per latitude without (green) and with (red) fitting of temperature offset. Lower panels: RMS error of XCH_4 per latitude bin for same simulation runs as in upper panel. The red areas indicate regions with high SZA that we plan to filter out.

Table 7: Sensitivity of retrieval performance to model errors in methane, water, temperature and pressure profiles. The results for retrievals with a temperature offset and surface pressure fitted are shown in brackets.

	converged	converged and filtered	RMS forward model error (filtered data)	CPU time per retrieval
reference	95%	61%	0.45%	7.6 s
mean CH_4 profile, $\Delta CH_{4col}=2\%$	95%	61%	0.46%	7.6 s
mean H_2O profile, $\Delta H_{2Ocol}=10\%$	88%	54%	0.48%	9.4 s
mean T profile (T_{offset} fitted)	72% (91%)	40% (57%)	0.77% (0.50%)	11.1 s (7.9 s)
$\delta P=0.5\%$ ($P_{surface}$ fitted)	83% (96%)	51% (61%)	= 0.76% (0.46%)	7.6 s (8.2 s)

profile has some influence on the RMS error (+0.03%) and convergence (-7%). It is expected that the ECMWF information on H_2O is more accurate than what we used in our sensitivity study so the errors shown here should be considered as worst case.

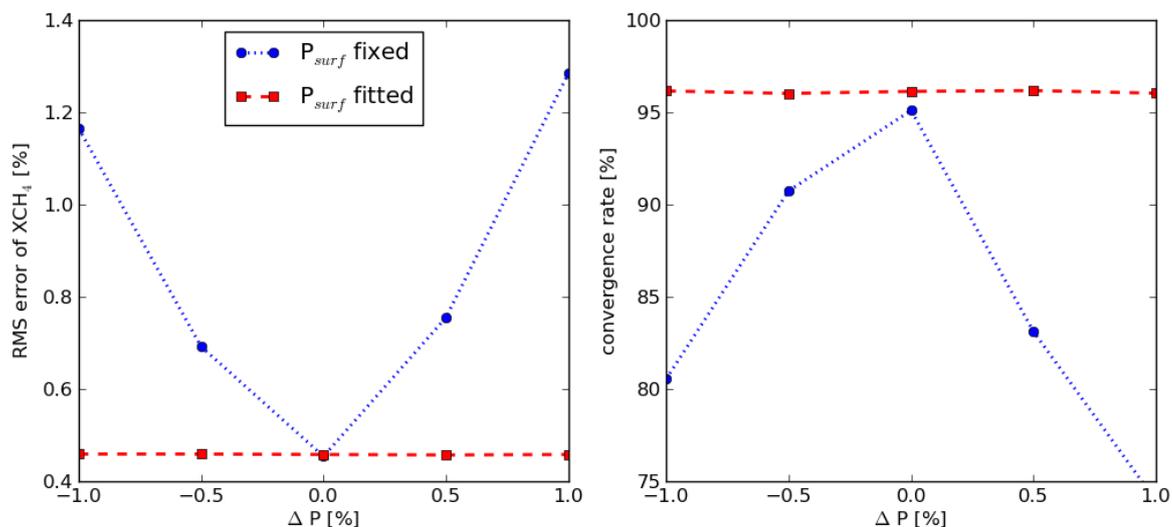


Figure 15: Influence of model error in prior pressure profile on accuracy (left) and stability (right) of retrievals without (blue) and with (red) fitting of the surface pressure.

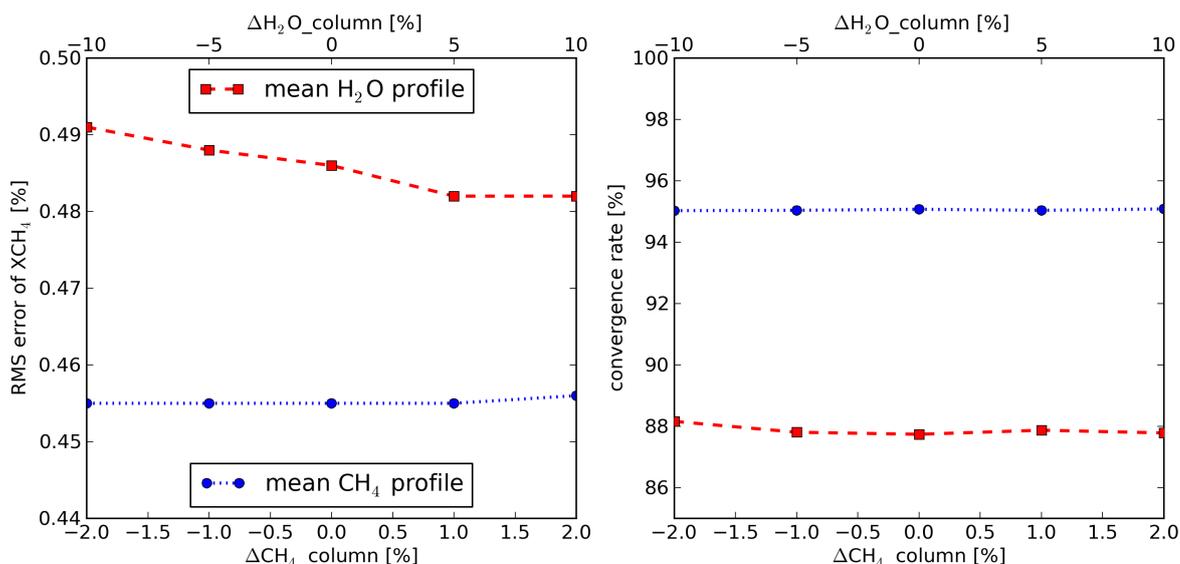


Figure 16: Influence of model errors in prior absorber profiles, CH_4 (blue) and H_2O (red), on retrieval accuracy (left) and stability (right).

7.3 Instrument errors

In this section, the sensitivity of the retrieval performance to relevant instrument and calibration errors is investigated. Where possible, the errors are set by the spectral and radiometric requirements as described in the System Requirement Document (SRD) [RD67].

7.3.1 Signal to noise ratio

The simulated spectra include noise based on the TROPOMI requirements as described in the SRD. For the NIR and the SWIR band signal-to-noise ratios in the continuum of 500 and 100, respectively, have been adopted. The precision is given by the retrieval noise $\sigma_{X_{CH_4}}$. The world map in Figure 17 shows the precision relative to the retrieved X_{CH_4} . Typically the precision is much better than the accuracy. The signal to noise

ratio becomes only a limiting factor for scenarios with snow-covered ground and large SZA.

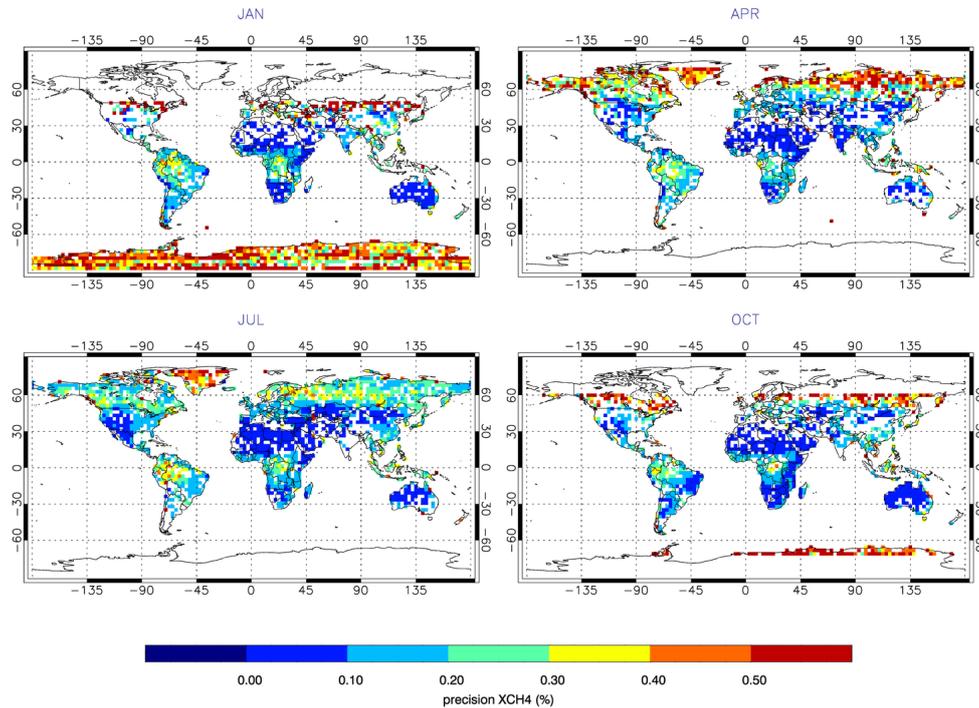


Figure 17: Relative precision of XCH_4 due to the instrument noise for the posteriori filtered dataset.

7.3.2 Instrument spectral response function

For the creation of the synthetic measurement spectra, the underlying line-by-line spectra have been convolved with a Gaussian spectral response function with $FWHM_{NIR} = 0.5 \text{ nm}$ and $FWHM_{SWIR} = 0.25 \text{ nm}$. According to the SRD, the ISRF shall be known within 1% of its maximum (where the ISRF is greater than 1% of its maximum). Approximately, this is achieved by varying the FWHM by 1%. Table 8 and Figure 18 gives the results for simulated retrievals with an assumed error in the FWHM. The retrieval appears to be more sensitive to the accuracy of the ISRF in the SWIR band. From Figure 18 it follows that for some cases (especially latitude bands 20-40 degrees north) the error can be quite significant, increasing the forward model error by up to 0.5% compared to the reference case. It is important to note that the error due to incorrect knowledge of the ISRF is a factor 2 larger than the budget reserved for it in the SRD [AD1]. Therefore we suggest that accurate calibration of the ISRF should have high priority.

7.3.3 Position of the spectral channels

According to the SRD, the positions of the centre of spectral channels are known within 2 pm. For the ensemble of simulated spectra, a spectral shift of 2 pm has negligible effect on the error characteristics. The spectral stability is 1/20 of the spectral sampling distance, i.e. 10 pm for the NIR band and 5 pm for the SWIR band. Both a spectral shift (wavelength independent) and a spectral squeeze (1st order wavelength dependence) are investigated here, where the maximum wavelength shift is taken to be 10 pm and 5 pm, respectively, for the NIR and SWIR band.

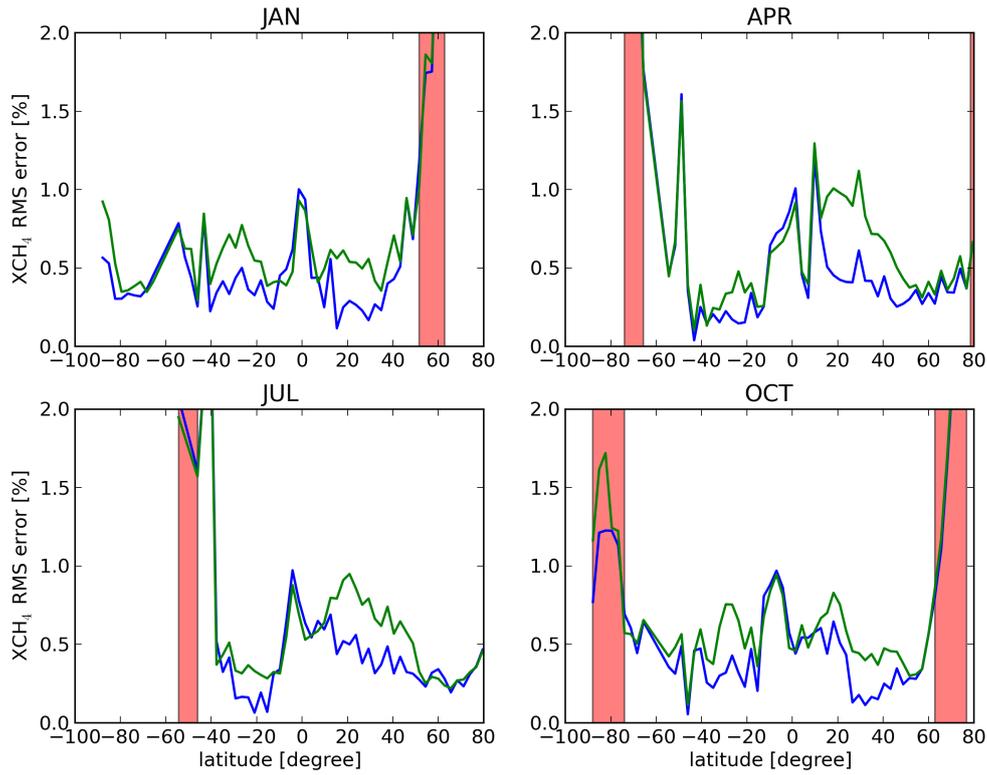


Figure 18: RMS error of XCH_4 for the simulation run with $\Delta FWHM_{NIR} = -1\%$ and $\Delta FWHM_{SWIR} = +1\%$ (green line), compared to the reference retrieval (blue line). The red areas indicate regions with high SZA that we plan to filter out.

Table 8: Influence of error in ISRF on the retrieval performance. Multiple simulation runs were performed; with FWHM increased and decreased. The results shown are for the run with poorest performance in terms of RMS error.

	converged	converged and filtered	RMS forward model error (filtered data)	CPU time per retrieval
$\Delta FWHM_{NIR} = -1\%$ $\Delta FWHM_{SWIR} = 0\%$	95%	61%	0.47%	7.6 s
$\Delta FWHM_{NIR} = 0\%$ $\Delta FWHM_{SWIR} = +1\%$	91%	60%	0.57%	7.6 s
$\Delta FWHM_{NIR} = -1\%$ $\Delta FWHM_{SWIR} = +1\%$	91%	60%	0.58%	7.6 s

Spectral shift

The reference retrieval fits a spectral shift. To test this fitting option, the synthetic spectra were shifted with a constant wavelength shift:

$$\lambda'_k = \lambda_k + \Delta\lambda_{\text{shift}} \quad (56)$$

where λ_k is the real wavelength, λ'_k is the measured wavelength at pixel k . In Table 9 results are given for the case that a spectral shift is not fitted and is fitted (between brackets). When fitted, the performance is as good as for the reference retrieval, i.e. simulations without an error in the spectral position. This indicates that the

Table 9: Influence of a constant error in position of spectral channels. Multiple simulation runs were performed; with $\Delta\lambda_{\text{shift}}$ positive and negative. The results shown are for the run with poorest performance. The results for the case a spectral shift is fitted is given in brackets.

	converged	converged and filtered	RMS forward model error (filtered data)	CPU time per retrieval
$\Delta\lambda_{\text{shift,NIR}} = -10 \text{ pm}$ $\Delta\lambda_{\text{shift,SWIR}} = 0 \text{ pm}$	78%	42%	0.53%	6.6 s
$\Delta\lambda_{\text{shift,NIR}} = 0 \text{ pm}$ $\Delta\lambda_{\text{shift,SWIR}} = +5 \text{ pm}$	95%	60%	0.47%	4.1 s
$\Delta\lambda_{\text{shift,NIR}} = -10 \text{ pm}$ $\Delta\lambda_{\text{shift,SWIR}} = +5 \text{ pm}$	77% (95%)	41% (61%)	0.55% (0.45)	5.8 s (7.6 s)

Table 10: Influence of a wavelength dependent error in position of spectral channel. Multiple simulation runs were performed; with $\Delta\lambda_{\text{squeeze}}$ positive and negative. The results shown are for the run with poorest performance. The results for the case a spectral squeeze is fitted is given in brackets.

	converged	converged and filtered	RMS forward model error (filtered data)	CPU time per retrieval
$\Delta\lambda_{\text{squeeze,NIR}} = -10 \text{ pm}$ $\Delta\lambda_{\text{squeeze,NIR}} = 0 \text{ pm}$	95%	60%	0.47%	7.6 s
$\Delta\lambda_{\text{squeeze,NIR}} = 0 \text{ pm}$ $\Delta\lambda_{\text{squeeze,NIR}} = -5 \text{ pm}$	95%	60%	0.49%	7.9 s
$\Delta\lambda_{\text{squeeze,NIR}} = -10 \text{ pm}$ $\Delta\lambda_{\text{squeeze,NIR}} = -5 \text{ pm}$	94% (95%)	60% (61%)	0.50% (0.46%)	7.7 s (9.4 s)

spectral shift fitting is robust.

Spectral squeeze

Optionally a wavelength dependent shift can be fitted. To test this fitting option, the synthetic wavelength grid was "squeezed":

$$\lambda'_k = \lambda_k + \Delta\lambda_{\text{squeeze}} \frac{\lambda_k - \lambda_{\text{mid}}}{\lambda_{\text{end}} - \lambda_{\text{mid}}} \quad (57)$$

where λ_{end} and λ_{mid} are the wavelengths at the end and middle of the band, respectively.

Table 10 shows the performance of retrievals with this assumed error on the measured wavelength grid. It follows that a spectral squeeze has a small but significant impact on the retrieved XCH_4 .

7.3.4 Radiometric offset (additive factor)

The effect of an unknown systematic offset in the Earth radiance is investigated. The offsets in the NIR and SWIR bands are independently varied with $\pm 0.1\%$ of the continuum. The algorithm (technically) allows fitting an additive constant to the spectrum. This option is tested here. The results in Table 11 and Figures 19 and 20 show that the retrieval algorithm cannot correct for an error in the radiometric offset, the performance is actually degraded when attempting to fit an offset. Because it is expected that the effect of noise will be larger when fitting an offset, the relative precision is shown in Figure 21. The noise contribution typically increases with 0.2% when fitting a radiometric offset.

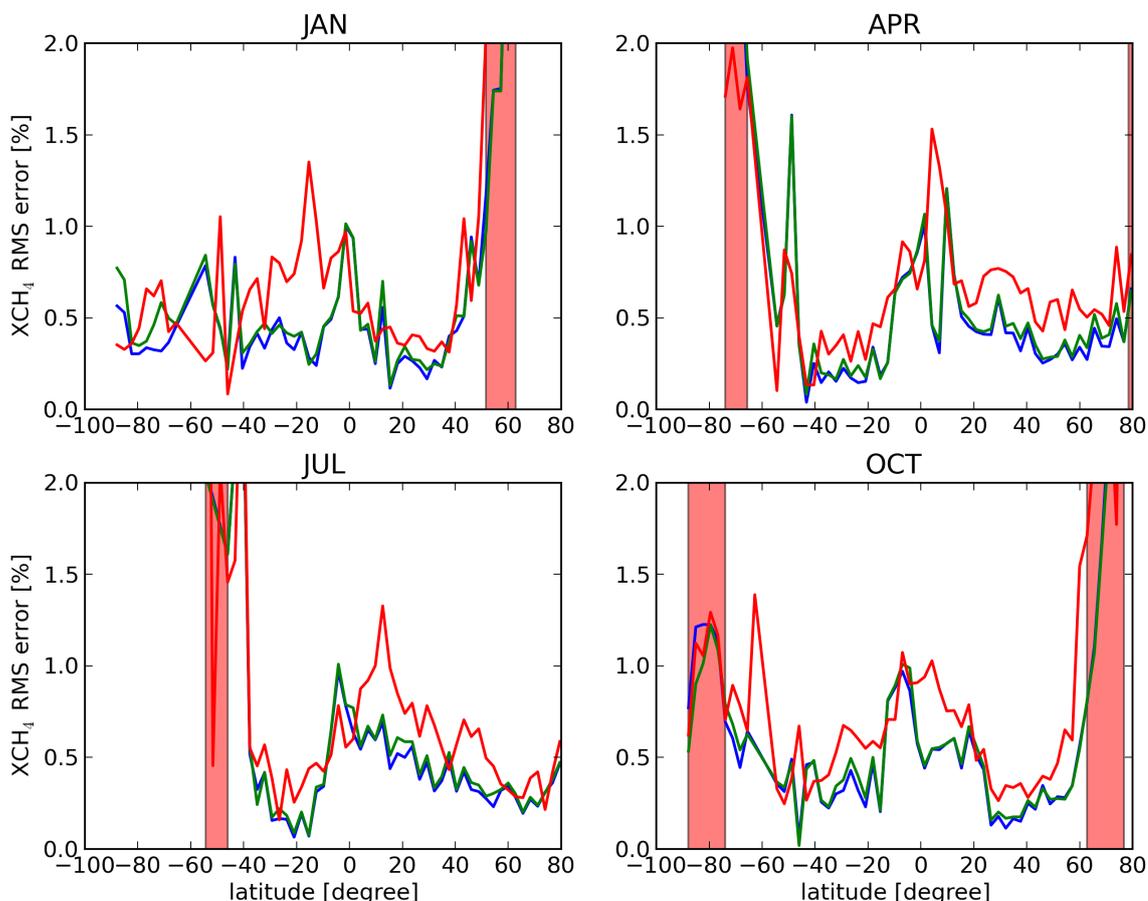


Figure 19: RMS error of XCH₄ retrieval per latitude bin for reference retrieval (blue), offsetted by 0.1% of the continuum but not fitted (green), and offsetted plus fitted (red). The red areas indicate regions with high SZA that we plan to filter out.

Table 11: Influence of a radiometric offset. Multiple simulation runs were performed; with positive and negative offset. The results shown are for the run with poorest performance. The results for the case an offset is fitted is given in brackets.

	converged	converged and filtered	RMS forward model error (filtered data)	CPU time per retrieval
offset _{NIR} = 0.1% offset _{SWIR} = 0%	91%	58%	0.47%	7.6 s
offset _{NIR} = 0% offset _{SWIR} = -0.1%	94%	63%	0.47%	7.7 s
offset _{NIR} = 0.1% offset _{SWIR} = -0.1%	90% (82%)	60% (56%)	0.49% (0.65%)	7.7 s (7.7 s)

7.3.5 Radiometric gain (multiplicative factor)

The absolute radiometric accuracy of the measurement of the Earth spectral radiance shall be better than 2%, see SRD. To investigate the effect of such an error, the synthetic spectra were multiplied with a scaling factor *A*. Table 12 shows that there is negligible effect of an error of 2% in radiometric gain. This error is largely

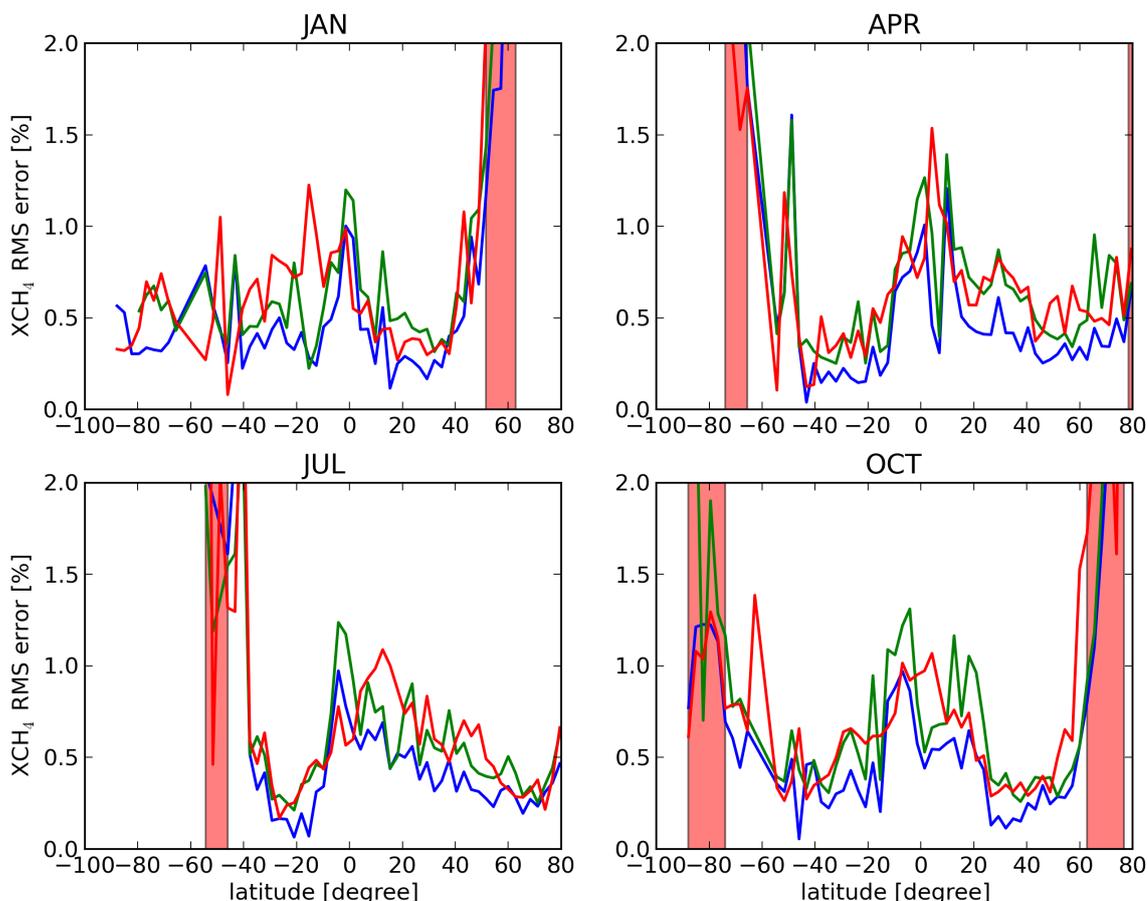


Figure 20: Same as Figure 19 but for an offset of 0.5%.

Table 12: Influence of an error in radiometric gain. Multiple simulation runs were performed; with multiplicative factor smaller and greater than 1. The results shown are for the run with poorest performance.

	converged	converged and filtered	RMS forward model error (filtered data)	CPU time per retrieval
$A_{NIR} = 1.02$ $A_{SWIR} = 1.00$	95%	61%	0.46%	7.6 s
$A_{NIR} = 1.00$ $A_{SWIR} = 1.02$	95%	60%	0.46%	7.6 s
$A_{NIR} = 1.02$ $A_{SWIR} = 1.02$	95%	61%	0.47%	7.6 s

compensated by the retrieved surface albedo.

7.3.6 Combined errors and fitting options

Although the performance of optional fitting parameters have been tested here, one should remain cautious of fitting too many parameters. The information content in the observed spectra limits the allowed number and choice of fitting parameters. For example, simulated retrievals where both a radiometric offset and a spectral squeeze are fitted only have a convergence rate of less than 30%.

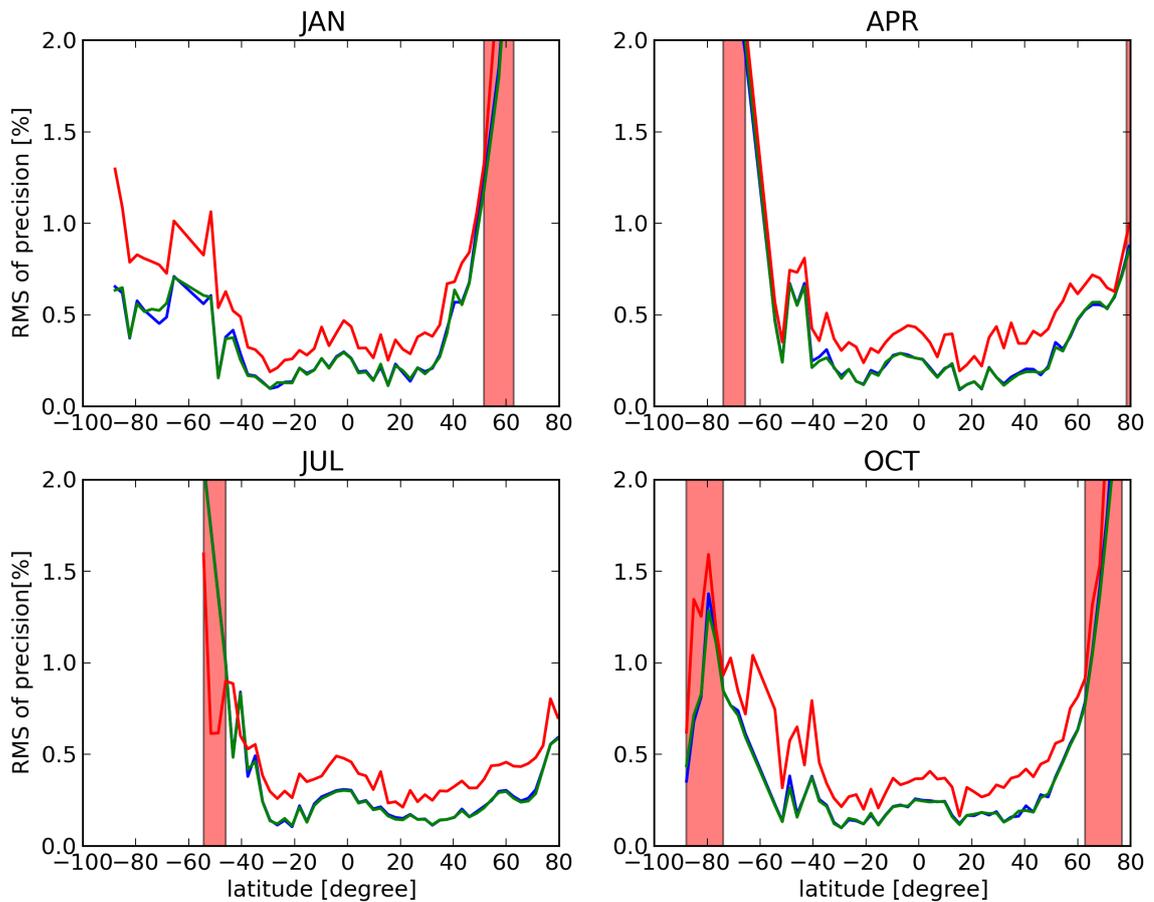


Figure 21: RMS of relative precision of XCH_4 retrievals per latitude bin for reference retrieval (blue), offsetted (by 0.1% of the continuum) but not fitted (green), and offsetted plus fitted (red). The red areas indicate regions with high SZA that we plan to filter out.

Table 13: Influence of combined instrument errors on retrieval performance. Retrieval settings are the same as for the reference retrieval.

	converged	converged and filtered	RMS error without noise (filtered data)	RMS of retrieval noise (filtered data)	CPU time per retrieval
reference	95%	61%	0.45%	0.37%	7.6 s
instrument errors	83%	56%	0.64%	0.36%	7.8 s

Table 13 and Figure 22 show the results for the baseline inversion as described in Sec. 1 but with assumed instrument errors in both NIR and SWIR band for the synthetic spectra: 1% error in ISRF, radiometric offset error of 0.1% of the continuum, error in radiometric gain of 2%, spectral shift of 2 pm, spectral squeeze of 2 pm (at the ends of spectral bands). This gives an estimate of the total effect of the instrument errors considered here. The XCH_4 error due to an error in the ISRF appears to be the dominating error.

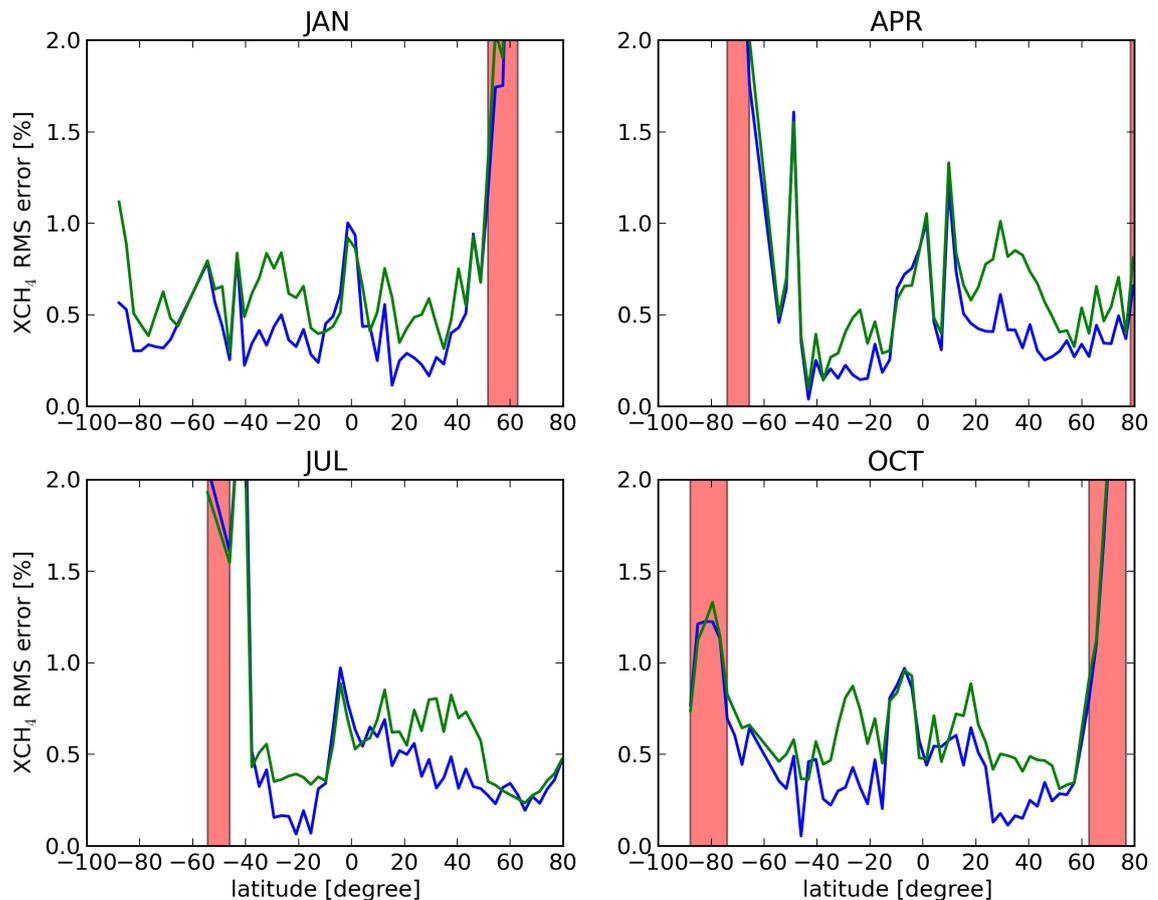


Figure 22: RMS error of XCH_4 when all instrument errors considered here are applied to the synthetic spectra (green line), compared to reference retrieval (blue line). The retrieval settings are the same for both simulation runs. The red areas indicate regions with high SZA that we plan to filter out.

7.4 Filtering for clouds / cirrus

Overall, the S5P methane algorithm works satisfactory. However, the limited number of cases where the error is larger than 1% corresponds mostly to tropical regions which are very important for methane source/sink inversions. These scenes are difficult because of the combination of low surface albedo and high cirrus load. As described in section 5 we perform an (additional) cirrus filtering based on the difference in H_2O columns $[H_2O]_{weak}$ and $[H_2O]_{strong}$, and CH_4 columns $[CH_4]_{weak}$ and $[CH_4]_{strong}$, retrieved from weak and strong absorption bands, respectively. Here, the weak absorption band is chosen as 2310-2315 nm for both H_2O and CH_4 , and the strong absorption bands are 2375-2380 nm for H_2O and 2363-2373 nm for CH_4 . We found that a threshold value of 22% for H_2O ($(weak-strong)/strong \times 100$) and 6% for CH_4 is appropriate to filter out the cases with large ($>1\%$) underestimation of the methane column caused by cirrus over a dark surface. The resulting methane errors are shown in Figure 23 (world maps) and Figure 24 (cumulative distribution). From Figure 23 it can be seen that the cases with large underestimation of methane are effectively filtered out. From Figure 24 we see that the filter based on H_2O retrievals works better as stand-alone filter (in addition to aerosol filter) than the filter based on CH_4 , but both filters improve considerably on the aerosol filter only. Using all filters together still improves the performance slightly compared to the use of the H_2O filter.

Figure 25 shows the difference in retrieved columns between weak and strong absorption bands for non-scattering atmosphere for CH_4 (solid) and H_2O (dotted) as a function of cloud fraction for a thick ($COT=5$) water cloud with cloud top at 2 km. It can be seen that, in contrast to cirrus clouds, water clouds cannot be filtered out to the level required as a maximum cloud fraction of 1-2% is allowed to keep the methane error

below 0.25%-0.5%. With the thresholds defined above (22% for H₂O and 6% for CH₄) only cloud fractions > 0.15 will be filtered out. For these clouds the aerosol filter (defined in section 7.1) is more suited, and it will filter out cases with cloud fraction > 8% (see Figure 26). However, such cloud fractions still lead to unacceptable errors in the retrieved XCH₄ and hence this emphasizes the need for the VIIRS cloud imager, which is particularly strong in detecting small fractions (within a S5P pixel) of thick water clouds.

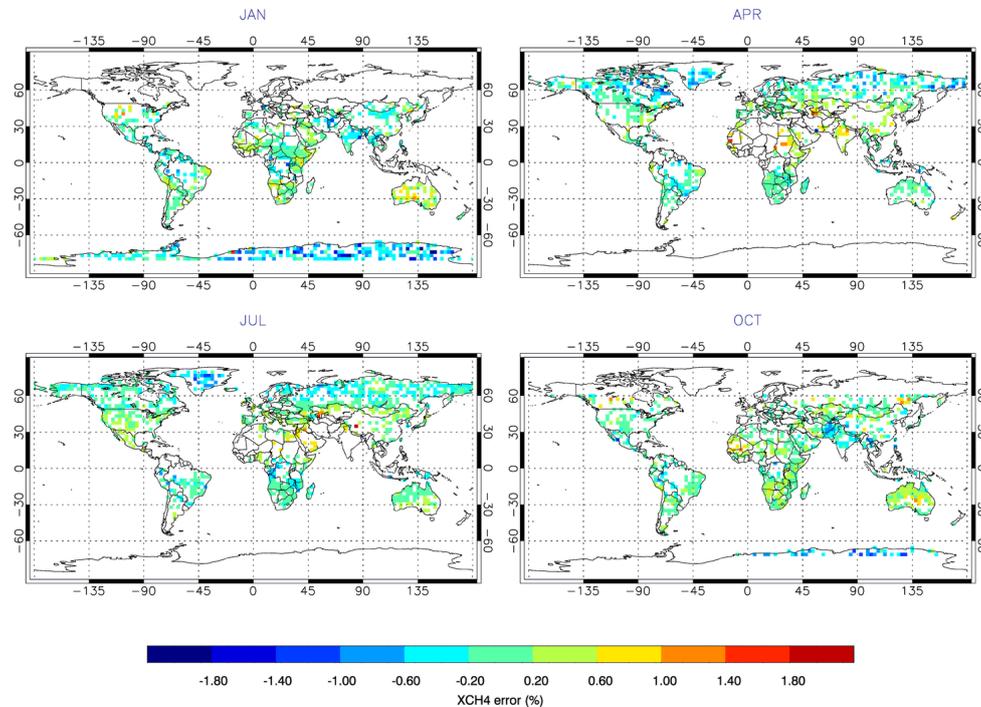


Figure 23: Forward model error of XCH₄ after applying the a posteriori filters and the a priori cloud filter based on non scattering H₂O and CH₄ retrievals.

7.5 Fluorescence

To test our approach for accounting for plant fluorescence we included fluorescence in the synthetic measurements for those scenes of the ensemble, which have a vegetated surface. To create the synthetic measurements we used a more detailed fluorescence model, where the spectral dependence by two co-added Gaussians (see [RD49, Eq. 1]), than the simplified model used in the retrieval. Figure 27 shows the RMS retrieval error on XCH₄ as a function of fluorescence emission for retrievals where fluorescence is not included and for retrievals with fluorescence included. If fluorescence is not included, the RMS retrieval error increases significantly from 0,7% to 0,9% with increasing fluorescence signal. On the other hand, if fluorescence is included in the retrieval (i.e. fitting $F_{s,755nm}^{surf}$ and s in Eq. 42 the RMS error is similar for different values of the fluorescence signal (there is a slight decrease in RMS error, probably as a result of compensating aerosol / fluorescence errors). So, the inclusion of fluorescence in the retrieval making use of Eq. 42 is sufficiently accurate for our purposes.

7.6 Summary and discussion of error analysis

Overall, the developed algorithm does well in correcting and filtering for the effect of aerosols and cirrus clouds on the retrieved XCH₄. After filtering for cirrus contamination for the ensemble of synthetic measurements 85% of the retrievals yield an error on XCH₄ that is less than 0.5%. Apart from strict cirrus filtering, we have performed a bias correction of the XCH₄ data (see Sect. 5.6), similar to what we did for GOSAT [RD29].

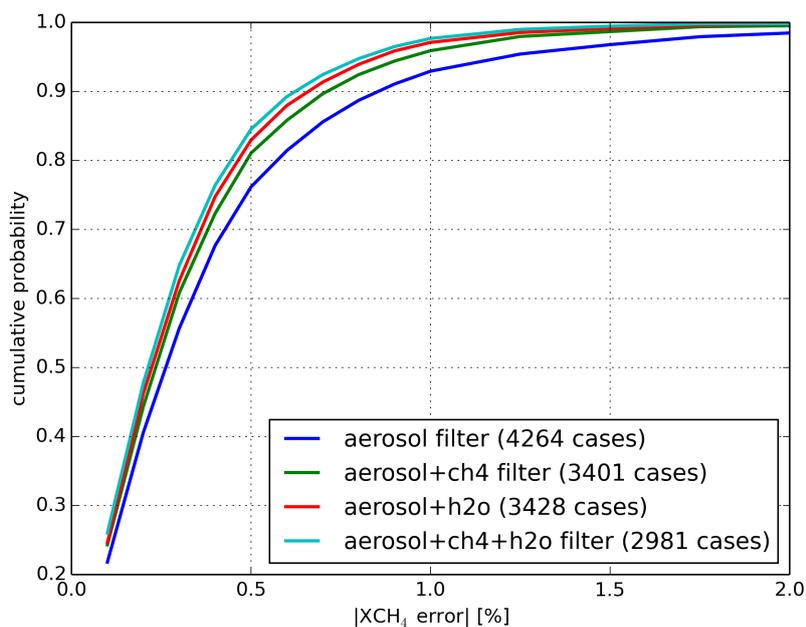


Figure 24: Cumulative probability distribution of the absolute XCH_4 retrieval error for the reference retrieval that includes scattering of aerosols and cirrus for the baseline filter, baseline and an additional filter based on non scattering CH_4 retrievals, baseline and an additional filter based on non-scattering H_2O retrievals, and the baseline and both CH_4 and H_2O filters.

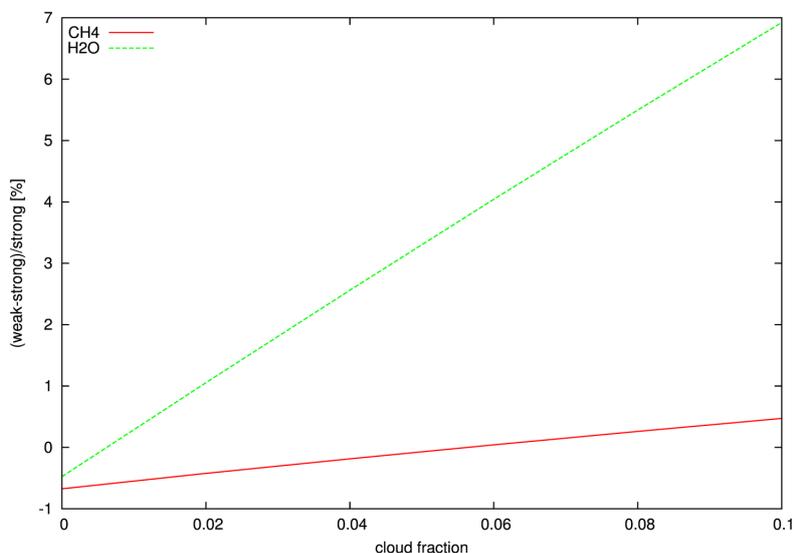


Figure 25: Difference in retrieved columns between weak and strong absorption bands for non-scattering atmosphere for CH_4 (solid) and H_2O (dotted). For the creation of synthetic measurements a cloud with optical thickness 5 was used with cloud top height = 2 km and cloud geometrical thickness = 1 km.

Such a bias correction further enhances the use of S5P XCH_4 data for inverse modelling and other scientific applications.

Apart from forward model errors induced by aerosols, we also studied model errors due to errors in temperature, pressure, and water vapour profiles. Here, we found that fitting a constant shift in the temperature profile is an effective way to correct for errors in the temperature profile. Furthermore, we found that it is

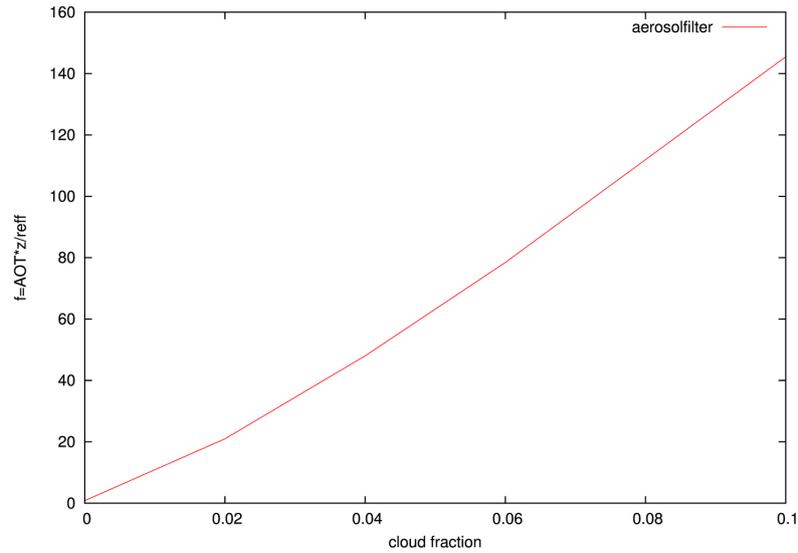


Figure 26: Aerosol filter f value versus cloud fraction. Cases with $f > 110$ will be filtered out, which roughly corresponds to a cloud fraction of 0.08.

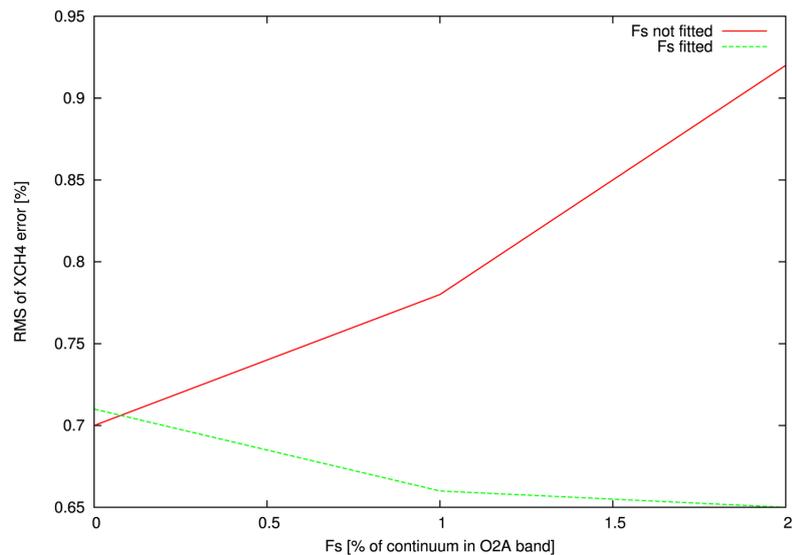


Figure 27: RMS X_{CH_4} error as a function of fluorescence emission for retrievals with and without fluorescence included in the fit.

important that the surface pressure should be known within about 2 hPa in order to avoid the need to fit this parameter (which strongly correlates with aerosol).

Concerning instrument errors, we found that the most critical error source is an error in the ISRF in the SWIR band. If the FWHM of the ISRF in the SWIR has an error of 1%, this may cause errors as large as 0.5% on X_{CH_4} , which means that the requirement formulated in the SRD [RD67] is not strict enough. Therefore it is extremely important that the ISRF is well calibrated. Furthermore, it is important that the offset on the measured spectrum is small, as the measurement does not contain sufficient information to correct for such an error in the retrieval algorithm.

8 Validation

Please refer to the Mission Performance Centre (MPC) validation reports that are regularly updated on <https://mpc-vdaf.tropomi.eu/index.php/methane>.

9 Examples of TROPOMI CH₄ data

After the successful launch of TROPOMI on October 13th, 2017 as the single payload of ESA's Sentinel-5 Precursor (S-5P) satellite, first calibrated SWIR radiance data were received at November 9th, 2017. The data quality was already sufficient to process the CH₄ total column product with remarkable accuracy, as presented in [RD53]. This section shows first examples of the operational TROPOMI CH₄ product and summarizes preliminary data validation results.

Figure 28 shows first results of global CH₄ observations of TROPOMI averaged from November 28th 2018 to January 16th 2019. For the retrieval, we process only clear-sky observations with a solar zenith angle (SZA) < 70°, a viewing zenith angle (VZA) < 60° and a surface albedo > 0.02. Here only data are used with a quality assurance value (qa_value) > 0.5. The data selection approach is summarized in Section 6.2, and the specifics on the recommendations for data usage and quality assurance are provided in the Methane Product Readme File [RD64].

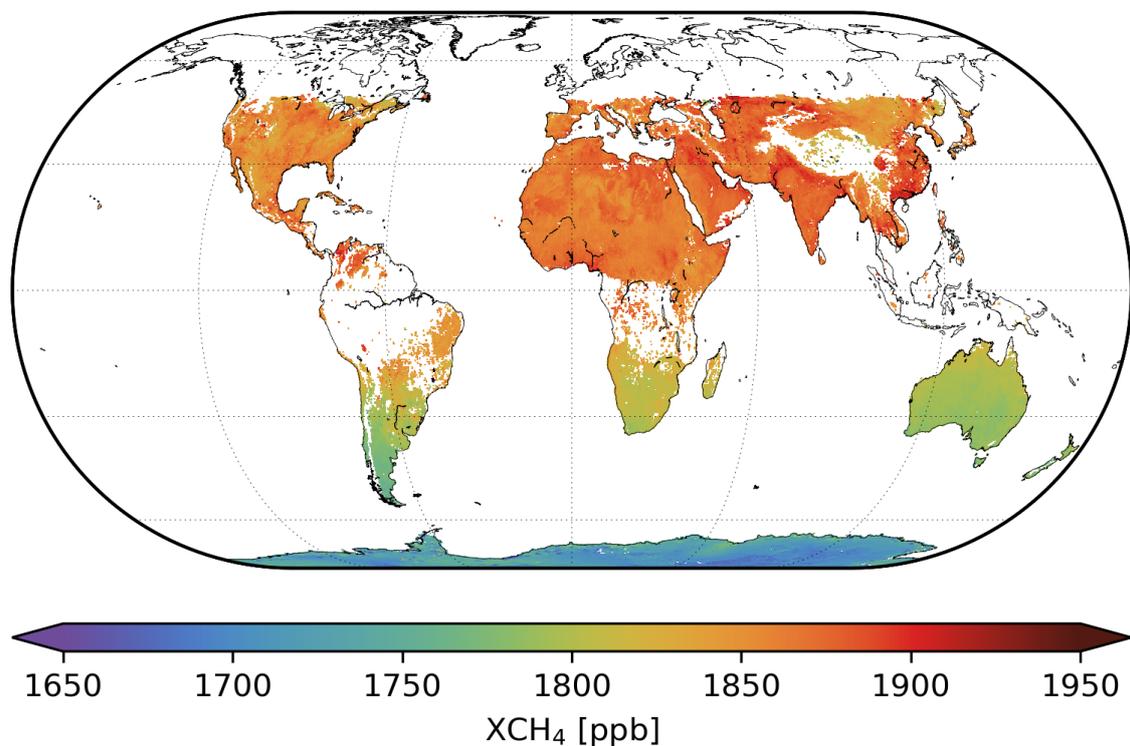


Figure 28: CH₄ total column mixing ratio bias corrected of TROPOMI averaged from November 28th 2018 to January 16th 2019

10 Conclusions

This document describes the algorithm for methane retrieval from the S5P instrument. The heritage of the algorithm lies in earlier algorithm developments for retrieval of CO₂ columns from OCO (launch failed February 2009) and CO₂ and CH₄ from GOSAT (successfully launched January 2009), described by Butz et al. 2009 [RD20]. Tests of the retrieval algorithm have been performed on synthetic GOSAT data [RD9], and real GOSAT data [RD23, RD16].

In order to account for the effect of aerosols and cirrus, the developed algorithm retrieves the methane column simultaneously with the aerosol/cirrus amount (column integrated particle number concentration), a parameter related to the particle size distribution, and a parameter describing the height distribution. Here, the particle size distribution is described by a power-law function [RD40], which only has two free parameters (related to amount and size). The choice of aerosol/cirrus parameters reflects the information content of the measurements as close as possible. The retrieval algorithm uses the level 1b reflectance measurements in the Short Wave Infra-Red (SWIR) band and additionally in the NIR band between 757-774 nm (O₂ A band). Additional fit parameters are the surface albedo and its third order spectral dependence in the two bands, and the total columns of carbon-monoxide and water vapor, respectively. In order to obtain a proper characterization of the retrieved methane column, it is important to first retrieve a vertical profile (layer averaged number density in different layers of the model atmosphere) and use this retrieved vertical profile to calculate the vertical column. Here, we choose to provide the vertical column as a product, and not the full profile, because the Degree of Freedom for Signal (DFS) of the retrieved methane profile is about 1. The inversion is performed using Phillips-Tikhonov regularization in combination with a reduced step size Gauss-Newton iteration scheme.

The forward model of the retrieval algorithm uses online radiative transfer calculations, fully including multiple scattering. Here, the radiative transfer model developed Landgraf et al. [RD45], and Hasekamp et al. 2002, 2005 [RD68, RD46] is being used. This model uses the Gauss-Seidel iterative method to solve the radiative transfer equation in a plane-parallel, vertically inhomogeneous atmosphere. To avoid time consuming line-by-line calculations we employ the linear-*k* method developed by Hasekamp and Butz 2008 [RD27]. Absorption cross sections of the relevant atmospheric trace gases are tabulated in a lookup table as function of pressure and temperature. Optical properties of aerosols are also calculated from lookup tables as described by Dubovik 2006 [RD42]. The linear *k*-binning method in combination with other speed optimizations allow us to perform the methane retrievals with online RT calculations within ~7-10 seconds for a single retrieval. This makes the algorithm feasible for operational S5P data processing.

To test the developed algorithm in the pre-launch phase we generated an ensemble of simulated measurements that covers the range of scenes that are encountered by the S5P instrument. This includes a realistic and detailed description of aerosol and cirrus properties, in combination with a surface albedo. Aerosol, cirrus and surface properties are taken from model (ECHAM5-HAM) and satellite measurements (MODIS, CALIOP/CALIPSO). Overall, the developed algorithm does well in correcting for the effect of aerosols and cirrus on the retrieved XCH₄. For the ensemble of synthetic measurements 85% of the cases has an XCH₄ error < 0.5%. To achieve this a strict filtering for cirrus clouds is needed which can be performed based on non-scattering retrievals of CH₄ and H₂O from weak and strong absorption bands in the SWIR. These retrievals are provided by the CO algorithm SICOR [RD58]. It should be noted that the filters based on S5P itself cannot replace the VIIRS cloud mask for water clouds. Apart from a strict filtering, we have performed a bias correction of the XCH₄ data (see Sect. 5.6), similar to what we did for GOSAT [RD28]. Such a bias correction further enhances the use of S5P XCH₄ data for inverse modelling and other scientific applications.

Apart from forward model errors induced by aerosols, we also studied model errors due to errors in temperature, pressure, and water vapour profiles. Here, we found that fitting a constant shift in the temperature profile is an effective way to correct for errors in the temperature profile. Furthermore, we found that it is important that the surface pressure should be known within about 2 hPa in order to avoid the need to fit this parameter (which strongly correlates with aerosol).

Concerning instrument errors, we found that the most critical error source is an error in the ISRF in the SWIR band. If the FWHM of the ISRF in the SWIR has an error of 1%, this may cause errors as large as 0.5% on XCH₄. Furthermore, it is important that the offset on the measured spectrum is small, as the measurement does not contain sufficient information to correct for such an error in the retrieval algorithm.

Based on the analysis of the first CH₄ measurements during the first months of the Sentinel 5 Precursor commissioning phase [RD53] and on the validation of the first months of data from the offline operational data product (Sect. 9), we conclude that the mission provides a reliable high-quality CH₄ data product with daily global coverage. Preliminary validations with ground based TCCON observations indicated a CH₄ data quality that already complies with the mission requirements.

11 Appendix B: Description of test cases

The simulated measurements are spectra of backscattered sunlight including noise based on the TROPOMI requirements as described in the TROPOMI SRD. For the O₂ A band in the TROPOMI NIR channel and the TROPOMI SWIR band signal-to-noise ratios of 500 and 120 in the continuum, respectively, have been adopted. Within the O₂ A band, the noise scales with the square root of the intensity, implying shot noise is the dominant noise contribution. For the TROPOMI SWIR band, a noise floor has been taken into account in addition to shot noise.

The forward model used to calculate the simulated measurements is the same as that in the retrieval method except for the treatment of aerosols and cirrus clouds which is highly simplified in the retrieval forward model in comparison with the true variability of aerosols and cirrus clouds in the atmosphere. For the simulation of measurements, a refined model is used that takes into account seven different aerosol modes, cirrus clouds, and arbitrary height distributions of the scatterers.

The aerosol input parameters for the generation of the ensemble simulations are derived from a run of the ECHAM5-HAM global aerosol model [RD41] for the year 2015 according to the A1B scenario defined by the Intergovernmental Panel on Climate Change. ECHAM5-HAM provides estimates of aerosol particle number concentrations and size distribution parameters for seven modes of a lognormal size distribution in 19 atmospheric layers between the ground and 10 hPa pressure on a 3° × 3° latitude by longitude grid. Each mode is a composite of five aerosol chemical types, i.e. sulfate, black carbon, organic matter, sea salt, and mineral dust. For each mode and considered spectral window, we calculate an average refractive index weighted by the relative mass contribution of the individual aerosol chemical type. Given size distribution and mode wise refractive index, the aerosol scattering and absorption OT and the aerosol scattering phase matrix are computed for the 19 ECHAM5-HAM layers individually and then interpolated to the simulation forward model height grid. The AOT from the ECHAM5-HAM model is scaled to MODIS observations (available at <http://modis-atmos.gsfc.nasa.gov/index.html>) where these are available. Figure 29 shows the distribution of AOT in the synthetic ensemble for the different seasons. Only data over land and with solar zenith angles below 70° are considered.

For cirrus clouds, we use the model developed by Hess et al. 1994 and 1998 [RD69, RD63] to calculate the phase matrix and the scattering and absorption OT of hexagonal ice crystals from a total column number density of cirrus. The ice crystals are assumed to be randomly oriented in space and to exhibit columnar shape. With a being the length of the side of the hexagon and c being the length of the column, cirrus optical properties are computed for eight crystal size pairs (a, c) between (1.4 μm, 3.5 μm) and (110 μm, 1300 μm) essentially following table 1 in Hess et al. 1994 [RD69]. Since model calculations for perfect hexagonal ice crystals exhibit unrealistically pronounced halos of the phase matrix at 22° and 44° scattering angle, we apply a smoothing correction suggested by Hess et al. 1998 [RD63]. Given the database of crystal size pairs (a, c) , we aggregate cirrus optical properties according to a power law size distribution in analogy to the study by Heymsfield et al. 1984 [RD70].

The total COT and the height and thickness of the cirrus layer are derived from the CALIOP/CALIPSO monthly mean data of the year 2007 [RD71]. We considered the level 2 cloud product data at 5 km resolution, provided by the ICARE Data and Services Center, <http://www.icare.univ-lille.fr>. Only cases with COT < 0.4 were considered for averaging and for calculation of the synthetic spectra, because scenes with higher OT will be identified as cloudy by the cloud screening algorithm of the VIIRS instrument. Figure 30 shows the global distribution of COT derived from CALIOP/CALIPSO data.

The surface albedo in the TROPOMI SWIR band is based on SCIAMACHY retrievals by Schrijver et al. 2009 [RD72]. Where no SCIAMACHY data are available, the MODIS LAND product is used, after scaling the measured albedo in band 7 (2.105-2.155 μm) by a factor of 0.7. This factor has been derived by estimating the ratio of the 2.3 μm/2.1 μm surface albedo for different vegetation types in the ASTER data base. The global distribution of the surface albedo used in the ensemble is shown in Figure 31. The albedo in the TROPOMI NIR band is taken from MERIS data, publicly available at www.temis.nl/data/meris.html [RD73].

The atmospheric conditions (pressure, temperature, humidity) are derived from the ECMWF (European Centre for Medium-Range Weather Forecasts) ERA-Interim analysis provided 6-hourly on a 1.5° × 1.5° latitude × longitude grid (available at <http://data-portal.ecmwf.int/data/d/interimdaily/>). The amount of CH₄ and CO is taken from the TM4 model [RD7]. Finally, atmospheric input, surface albedo, and cirrus and aerosol optical properties are combined to create a global ensemble representative for January, April, July, and October. One simulated measurement is created for each 3° × 3° gridbox.

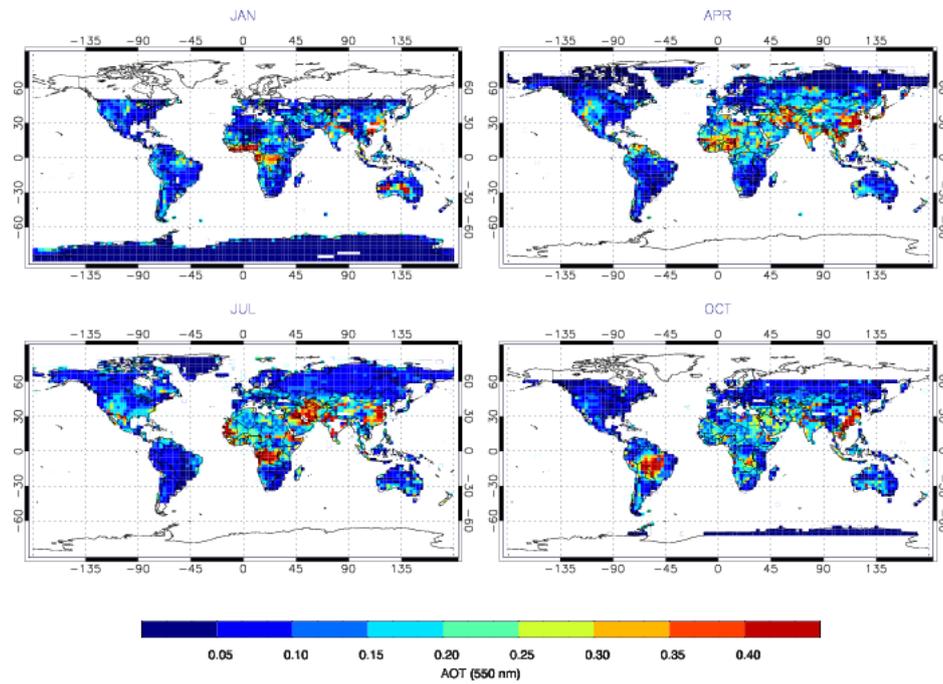


Figure 29: Global distribution of Aerosol Optical Thickness (AOT) adopted for the synthetic ensemble. January (top left), April (top right), July (bottom left), and October (bottom right).

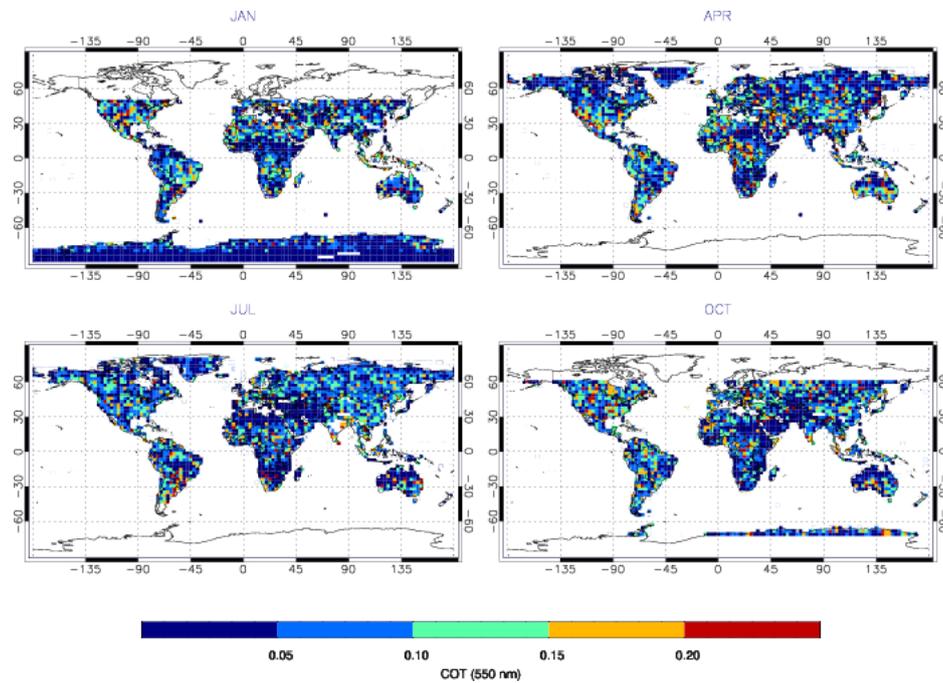


Figure 30: Global distribution of Cirrus Optical Thickness (COT) as used in the ensemble calculations.

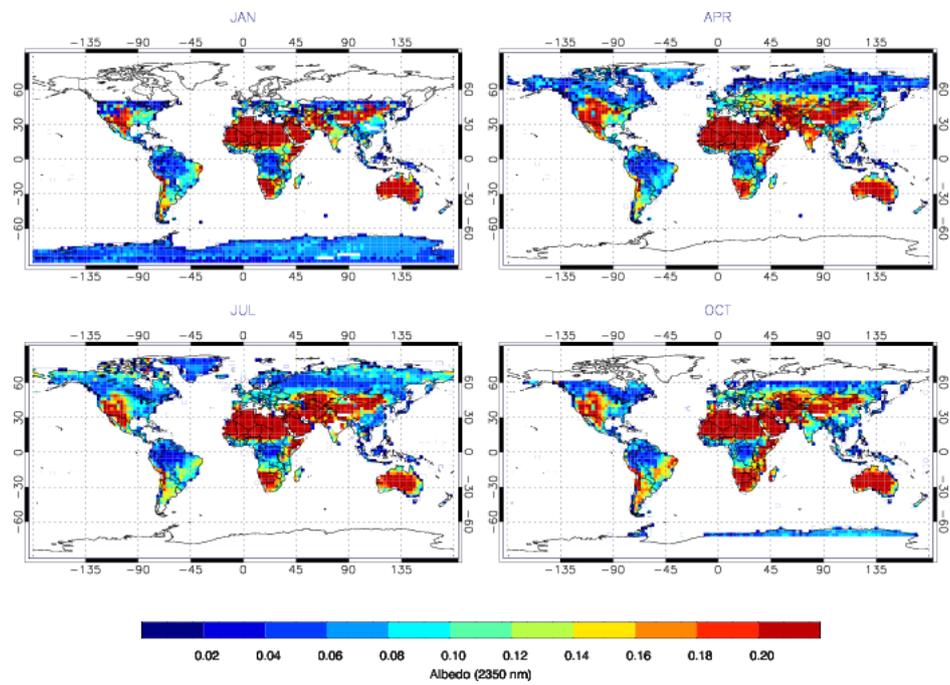


Figure 31: Surface albedo maps in the SWIR range for the synthetic ensemble.