SRON
Netherlands Institute for Space Research

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

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

document number	:	SRON-S5P-LEV2-RP-002
CI identification	:	CI-7430-ATBD
issue	:	1.10
date	:	2018-06-15
status	:	released

Document approval record

	digital signature	
prepared:		Jochen Landgraf c=NL, st=Utrecht, o=SRON, ou=EPS, cn=Jochen Landgraf, email=j.landgraf@sron.nl 2018.06.11 22:01:20 +01'00'
checked:	Michael Buchwitz and Thomas Krings Institute of Environmental Physics (IUP) University Bremen, Germany	
approved PM:	Arnoud Apituley	Digitally signed by Arnoud Apituley Date: 2018.06.14 18:13:24 +02'00'
approved PI:	J. Pepijn Veefkind	Digitally signed by J. Pepijn Veefkind Date: 2018.06.15 13:24:12 +02'00'

Document change record

issue	date	item	comments
0.0.1	2012-07-09	All	Initial draft version
0.0.2	2012-11-23	Appendix A	Appendix added
0.0.2	2012-11-23	Sec. 9	Hardware requirement updated
0.5.0	2013-06-04	all	Document revised according to SRR/PDR review Feb. 2013
0.5.0	2013-06-21	all	Document revised according to internal review June 2013
0.9.0	2013-11-30	all	Document revised according to external review September 2013
0.9.1	2013-12-03	sec. 9	Section 9 'Feasibility' revised
0.10.0	2014-04-15	app. B	SWIR pre-processor introduced
		app. C	Appendix C 'HDO/H_2O' revised
0.11.0	2014-09-30	sec. 8	Section added including tables on algorithm input and output
		app. B	Appendix B revised
0.11.1	2014-12-16	sec. 5.2.1	Cloud filter analysis extended to one year GOSAT data
		sec. 9.1	Estimated computational effort now based on cloud filter performance for one year of GOSAT data (no significant change regarding computational cost)
0.13.0	2015-09-08	sec. 8.3 and 8.4	Update of algorithm input/output satisfying project unit definition. Update for limited release to S5p Validation Team
0.13.0	2015-12-09	all	References updated, typos corrected
1.00	2016-02-05	all	Document revised according to internal review Dec. 2015
1.10	2018-06-15	all	appendix *TROPOMI HDO/H2O retrievals* removed section *Examples of TROPOMI CO data* added

Contents

List of Tables

List of Figures

1 Introduction

1.1 Identification

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

1.2 Purpose and objectives

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

1.3 Document overview

The document is structured as follows: After this introduction, references are provided in Sec. 2 and Sec. 3 contains a list of abbreviations used in this document. Sec. 4 provides a short introduction to satellite remote sensing of atmospheric CO abundance and the heritage of the presented algorithm is summarized. Moreover, we recall the level-2 requirement for the CO column product which represents the underlying criterion for the performance analysis of the presented algorithm. The theoretical concept of the CO retrieval algorithm SICOR is summarized in Sec. 5, comprising a description of the radiative transfer model and the inversion scheme. The parameters to be retrieved, ancillary data and a priori knowledge are discussed including the final data product of the algorithm. Section 7 considers the performance of the retrieval algorithm based on a set of generic measurement ensembles and a geo-physical ensemble of simulated measurements over China. Here, we investigate the CO retrieval noise and CO retrieval biases due to forward model errors, erroneous atmospheric input data and instrument artifacts. Based on this, we evaluate the algorithm performance in the context of the S5P level-1 and 2 requirements. The numerical feasibility is the subject of Sec. 9, which comprises an estimate of the numerical effort, a high level data product description and the spatial data selection criteria of the measurements to be processed. The validation approach of the CO data product is broached in the Sec. 10 and Sec. 12 concludes the document. Section 11 summarizes our analysis of the CO data quality during the S5P commission phase lasting until April, 2018, and Sec. 12 concludes our document.

Additional material is provided in the appendices, where Appendix A discusses in detail the linearized two-stream method and Appendix B describes the SWIR preprocessing module which provides required input to both the SICOR CO algorithm and the RemoTeC CH_4 algorithm (see [RD2]). Finally,

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.

[AD3] GMES Sentinel-5 Precursor – S5p System Requirement Document.
source: ESA; **ref:** S5p-RS-ESA-SY-0002; **date:** 2011.

[AD4] NL TROPOMI L2 data processors: Processor Design Document.
source: KNMI; **ref:** S5P-KNMI-L2-0030-SD; **issue:** 0.0.0; **date:** 2014-05-22.

2.2 Standard documents

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

[SD2] Space Product Assurance – Software Product Assurance.
source: ESA; **ref:** ECSS-E-ST-40C; **date:** 2009.

2.3 Reference documents

[RD1] A. J. Heymsfield and C. M. R. Platt; A Parameterization of the Particle Size Spectrum of Ice Clouds in Terms of the Ambient Temperature and the Ice Water Content. *J. Atmos. Sci.*; **41** (1984), 846; doi:10.1175/1520-0469(1984)041<0846:APOTPS>2.0.CO;2.

[RD2] Algorithm Theoretical Baseline Document for Sentinel-5 Precursor methane retrieval.
source: SRON; **ref:** SRON-S5P-LEV2-RP-001; **date:** 2014.

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

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

[RD5] H.H. Levy; Normal atmosphere: Large radical and formaldehyde concentrations predicted. *Science*; **173** (1971), 141.

[RD6] J.A. Logan, M.J. Prather, S.C. Wofsy *et al.*; Tropospheric chemistry: A global perspective. *J. Geophys. Res.*; **86** (1981), 7210.

[RD7] D.T. Shindell, G. Faluvegi, D.S. Stevenson *et al.*; Multi-model simulations of carbon monoxide: Comparison with observations and projected near-future changes. *J. Geophys. Res.*; **111** (2006), D19306, doi: 10.1029/2006JD007100.

[RD8] D.P. Edwards, L.K. Emmons, D.A. Hauglustaine *et al.*; Observations of carbon monoxide and aerosols from the Terra satellite: Northern Hemisphere variability. *J. Geophys. Res.*; **109** (2004), D24202, doi: 10.1029/2004JD004727.

[RD9] A.M.S. Gloudemans, M.C. Krol, J.F. Meirink *et al.*; Evidence for long-range transport of Carbon Monoxide in the Southern Hemisphere from SCIAMACHY observations. *Geophys. Res. Lett.*; **33** (2006), L16807, doi:10.1029/2006GL026804.

[RD10] A. Hollingsworth, R. J. Engelen, A. Benedetti *et al.*; Toward a Monitoring and Forecasting System For Atmospheric Composition: The GEMS Project. *Bull. Amer. Meteor. Soc.*; **89** (2008), 1147.

[RD11] H.G. Reichle Jr and V.S. Connors; The mass of CO in the atmosphere during October 1984, April 1994, and October 1994. *J. Atmos. Sci.*; **56** (1999), 307.

[RD12] H. Kobayashi, A. Shimota, K. Kondo *et al.*; Development and evaluation of the Interferometric Monitor for Greenhouse Gases: a high throughput Fourier transform infrared radiometer for nadir Earth observations. *Appl. Opt.*; **38** (1999), 6801.

[RD13] M.N. Deeter, L. K. Emmons, G. L. Francis *et al.*; Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument. *J. Geophys. Res.*; **108** (2003), 4399, doi:10.1029/2002JD003186.

[RD14] W.W. McMillan, C. Barnet, L. Strow *et al.*; Daily global maps of carbon monoxide from NASA's Atmospheric Infrared Sounder. *Geophys. Res. Lett.*; **32** (2005), L11801, doi:10.1029/2004GL021821.

[RD15] C. P. Rinsland, M. Luo, J. A. Logan *et al.*; Nadir measurements of carbon monoxide distributions by the Tropospheric Emission Spectrometer instrument onboard the Aura Spacecraft: Overview of analysis approach and examples of initial results. *Geophys. Res. Lett.*; **33** (2006), L22806, doi:10.1029/2006GL027000.

[RD16] S. Turquety, J. Hadji-Lazaro, C. Clerbaux *et al.*; Operational trace gas retrieval algorithm for the Infrared Atmospheric Sounding Interferometer. *J. Geophys. Res.*; **109** (2004), D21301, doi:10.1029/2004JD004821.

[RD17] The potential of MTG-IRS and S4-TIR to detect high pollution events at urban and regional scales. **source:** EUMETSAT; **ref:** EUM/CO/07/4600000447/SAT; **date:** 2009.

[RD18] H. Bovensmann, J. P. Burrows, M. Buchwitz *et al.*; SCIAMACHY: Mission Objectives and Measurement Modes. *Journal of Atmospheric Sciences*; **56** (1999), 127.

[RD19] M.N. Deeter, D. P. Edwards, J. C. Gille *et al.*; CO retrievals based on MOPITT near-infrared observations. *J. Geophys. Res.*; **114** (2009), D04303, doi:10.1029/2008JD010872.

[RD20] H. M. Worden, M. N. Deeter, D. P. Edwards *et al.*; Observations of near-surface carbon monoxide from space using MOPITT multispectral retrievals. *J. Geophys. Res.*; **115** (2010), D18314, doi:10.1029/2010JD014242.

[RD21] P. Veefkind; TROPOMI on the ESA Sentinel-5 Precursor: a GMES mission for Global Observations of the Atmospheric Composition for Climate and Air Quality Applications. *Remote Sens. Envirom.*; **120** (2012), 70.

[RD22] M. Buchwitz, I. Khlystova, H. Bovensmann *et al.*; Three years of global carbon monoxide from SCIAMACHY: comparison with MOPITT and first results related to the detection of enhanced CO over cities. *Atmos. Chem. Phys.*; **7** (2007), 2399.

[RD23] 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.

[RD24] A.M.S. Gloudemans, A.T.J. de Laat, H. Schrijver *et al.*; SCIAMACHY CO over land and oceans: 2003-2007 interannual variability. *Atmos. Chem. Phys.*; **9** (2009), 3799.

[RD25] M. Buchwitz, R. de Beek, S. Noel *et al.*; Atmospheric carbon gases retrieved from SCIAMACHY by WFM-DOAS: version 0.5 CO and CH_4 and impact of calibration improvements on CO_2 retrieval. *Atmos. Chem. Phys.*; **6** (2006), 2727.

[RD26] S. Turquety, C. Clerbaux, K. Law *et al.*; CO emission and export from Asia: an analysis combining complementary satellite measurements (MOPITT, SCIAMACHY and ACE-FTS) with global modeling. *Atmos. Chem. Phys.*; **8** (2008) (17), 5187.

[RD27] A. T. J. de Laat, A. M. S. Gloudemans, I. Aben *et al.*; Global evaluation of SCIAMACHY and MOPITT carbon monoxide column differences for 2004-2005. *J. Geophys. Res.*; **115** (2010), D06307, doi:10.1029/2009JD012698.

[RD28] B. Dils, M. De Maziere, J. F. Muller *et al.*; Comparisons between SCIAMACHY and ground-based FTIR data for total columns of CO, CH_4, CO_2 and N_2O. *Atmos. Chem. Phys.*; **6** (2006), 1953.

[RD29] A. T. J. de Laat, A. M. S. Gloudemans, H. Schrijver *et al.*; Validation of five years (2003-2007) of SCIAMACHY CO total column measurements using ground-based spectrometer observations. *Atmos. Meas. Tech.*; **3** (2010), 1457.

[RD30] A. T. J. de Laat, A. M. S. Gloudemans, I. Aben *et al.*; Scanning Imaging Absorption Spectroscopy for Atmospheric Chartography carbon monoxide total columns: Statistical evaluation and comparison with chemistry transprot model results. *J. Geophys. Res.*; **112** (2007), D12310, doi:10.1029/2006JD008256.

[RD31] GMES Sentinels 4 and 5 mission requirements document.
source: ESA; **ref:** EOP-SMA/1507/JL-dr; **date:** 2011.

[RD32] J. Vidot, J. Landgraf, O.P. Hasekamp *et al.*; Carbon monoxide from shortwave infrared reflectance measurements: A new retrieval approach for clear sky and partially cloudy atmospheres. *Remote Sens. Envirom.*; **120** (2012), 255.

[RD33] M. Krol, S. Houweling, B. Bregman *et al.*; The two-way nested global chemistry-transport zoom model TM5: algorithm and applications. *Atmos. Chem. Phys.*; **5(2)** (2005), 417.

[RD34] C.D. Rodgers; *Inverse Methods for Atmospheres: Theory and Practice*; volume 2 (World Scientific, 2000).

[RD35] A.M.S. Gloudemans, H. Schrijver, O.P. Hasekamp *et al.*; Error analysis for CO and CH_4 total column retrieval from SCIAMACHY 2.3μm spectra. *Atmos. Chem. Phys.*; **8** (2008), 3999.

[RD36] W.G. Zdunkowski, Welch R.M. and Korb G.; An Investigation of the Structure of Typical Two-stream-methods for the Calculation of Solar Fluxes and Heating Rates in Clouds. *Contrib. Atmos. Phys.*; **53** (1979), 147.

[RD37] W. J. Wiscombe and G. W. Grams; The backscattered fraction in two-stream approximations. *Journal of Atmospheric Sciences*; **33** (1976), 2440.

[RD38] G.I. Marchuk; Equation for the Value of Information from Weather Satellites and Formulation of Inverse Problems. *Cosmic Res.*; **2** (1964), 394.

[RD39] M.A. Box, S.A.W. Gerstl and C. Simmer; Application of the adjoint formulation to the calculation of atmospheric radiative effects. *Beitr. Phys. Atmos.*; **61** (1988), 303.

[RD40] 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.

[RD41] O.P. Hasekamp and J. Landgraf; A linearized vector radiative transfer model for atmospheric trace gas retrieval. *J. Quant. Spectrosc. Radiat. Transfer*; **75** (2002).

[RD42] V.V. Rozanov, D. Diebel, R.J.D. Spurr *et al.*; GOMETRAN: A radiative transfer model for the satellite project GOME, the plan-parallel version. *J. Geophys. Res.*; **102** (1997), 16,683.

[RD43] V.V. Rozanov, T. Kurosu and J.P. Burrows; Retrieval of Atmospheric Constituents in the UV-Visible: A new Quasi-Analytical Approach for the Calculation of Weighting Functions. *J. Quant. Spectrosc. Radiat. Transfer*; **60** (1998), 277.

[RD44] H. H. Walter, J. Landgraf and O. P. Hasekamp; Linearization of a pseudo-spherical vector radiative transfer model. *J. Quant. Spectrosc. Radiat. Transfer*; **85** (2004), 251.

[RD45] H. H. Walter and J. Landgraf; Towards linearization of atmospheric radiative transfer in spherical geometry. *J. Quant. Spectrosc. Radiat. Transfer*; **95** (2005), 175.

[RD46] H. H. Walter, J. Landgraf, F. Spada *et al.*; Linearization of a radiative transfer model in spherical geometry. *J. Geophys. Res.*; **111** (2006), 10.1029/2005JD007014.

[RD47] G. I. Bell and S. Glasstone; *Nuclear Reactor Theory* (Van Nostrand Reinhold Company, New York, 1970).

[RD48] J. Lewins; *Importance, the adjoint function* (Pergamon Press, Oxford, England, 1965).

[RD49] E. A. Ustinov; Inverse problem of photometric observation of solar radiation reflected by an optically dense planetary atmosphere. Mathematical methods and weighting functions of linearized inverse problem. *Cosmic Res.*; **29** (1991), 519.

[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; Analysis of discrete ill posed problems by means of the L-curve. *SIAM Rev.*; **34** (1992), 561.

[RD53] P.C. Hansen and D.P. O'Leary; The use of the L-curve in the regularization of discrete ill posed problems. *SIAM J. Sci. Comput.*; **14** (1993), 1487.

[RD54] T Borsdorff, O.P. Hasekamp, Wassmann A. *et al.*; Trace gas column retrieval and its appropriate regulatrization scheme. Manufscript in preparation.

[RD55] A. Butz, A Galli, O.P. Hasekamp *et al.*; ROPOMI aboard Sentinel-5 Precursor: Prospective performance of CH4 retrievals for aerosol and cirrus loaded atmospheres. *Remote Sens. Envirom.*; **120** (2012), 267.

[RD56] W.H. Press, S.A. Teukolsky, W.T. Vetterling *et al.*; *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge University Press, 1992); 2nd edition.

[RD57] L.S. Rothman, I.E. Gordon, A. Barbe *et al.*; The HITRAN 2008 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transfer*; **110** (2009), 533.

[RD58] OMI aerosol retrieval algorithm.
source: KNMI; **ref:** ATBD-OMI-03; **date:** 2001.

[RD59] J.M.J. Schepers, D.and aan de Brugh, Ph. Hahne, A. Butz *et al.*; LINTRAN v2.0: A linearised vector radiative transfer model for efficient simulation of satellite-born nadir-viewing reflection measurements of cloudy atmospheres. *J. Quant. Spectrosc. Radiat. Transfer*; **149** (2015), 247.

[RD60] S. Chandrasekhar; *Radiative transfer* (Dover Publications, Inc., New York, 1960).

[RD61] 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.

[RD62] 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.

[RD63] A Marshak, A. Davis, W. Wiscombe *et al.*; The Verisimilitude of the Independent Pixel Approximation Used in Cloud Remote Sensing. *Remote Sens. Envirom.*; **52** (1995), 71.

[RD64] Instrument noise model for the Sentinel 5 SWIR bands.
source: Netherlands Insitute for Space Research, SRON; **ref:** SRON-TROPSC-TN-2011-002; **date:** 2011.

[RD65] U.S. Standard Atmosphere, 1976.
source: National Oceanic and Atmospheric Administration; **ref:** NOAA-S/T76-1562.

[RD66] Observation Techniques and Mission Concepts for Atmospheric Chemistry (CAMELOT).
source: European Space Agency; **ref:** 20533/07NL/HE.

[RD67] A. E. Dessler and P. Yang; The Distribution of Tropical Thin Cirrus Clouds Inferred from Terra MODIS Data. *Journal of Climate*; **16** (2003), 1241.

[RD68] P. Yang, B.-C. Gao, B.A. Baum *et al.*; Radiative properties of cirrus clouds in the infrared (8–13 μm) spectral region. *J. Quant. Spectrosc. Radiat. Transfer*; (2001).

[RD69] B. Mijling and R.J. van der A; Using daily satellite observations to estimate emissions of short-lived air pollutants on a mesoscopic scale. *J. Geophys. Res.*; (2012).

[RD70] T.P.C. van Noije, H.J. Eskes, F.J. Dentener *et al.*; Multi-model ensemble simulations of tropospheric NO2 compared with GOME retrievals for the year 2000. *Atmos. Chem. Phys.*; **6** (2006), 2943.

[RD71] J. F. Meirink, P. Bergamaschi and M. C. Krol; Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: Method and comparison with synthesis inversion. *Atmos. Chem. Phys.*; **8** (2008), 6341.

[RD72] P. Bergamaschi, M. Krol, F. Dentener *et al.*; Inverse modelling of national and european CH_4 emissions using the atmospheric zoom model TM5. *Atmos. Chem. Phys.*; **5** (2005), 2431.

[RD73] P. Bergamaschi, C. Frankenberg, J.-F. Meirink *et al.*; Inverse modeling of global and regional CH_4 emissions using SCIAMACHY satellite retrievals. *J. Geophys. Res.*; **114** (2009), doi:10.1029/2009JD012287.

[RD74] A. Butz, S. Guerlet, O. Hasekamp *et al.*; Toward accurate CO_2 and CH_4 observations from GOSAT. *Geophys. Res. Lett.*; **38** (2011), doi:10.1029/2011GL047888.

[RD75] 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/H_2O 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/`.

[RD76] Sentinel-5 precursor/TROPOMI: Level 2 Product User Manual Carbon Monoxide.
source: SRON; **ref:** SRON-S5P-LEV2-MA-002; **issue:** 0.8.1; **date:** 2015-012-08.

[RD77] 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.

[RD78] C. P. Rinsland, E. Mahieu, R. Zander *et al.*; Free tropospheric CO, C_2H_6, and HCN above central Europe: Recent measurements from the Jungfraujoch station including the detection of elevated columns during 1998. *J. Geophys. Res.*; **105** (2000), 24235; doi:10.1029/2000JD900371.

[RD79] R. Sussmann and T. Borsdorff; Technical Note: Interference errors in infrared remote sounding of the atmosphere. *Atmospheric Chemistry & Physics*; **7** (2007), 3537.

[RD80] A. Karon, C. Sweenney, P. Tans *et al.*; AirCore: An Innovative Atmospheric Sampling System. *J. Atm. Ocean Tech.*; **27** (2010), 1839; doi:doi:10.1175/2010JTECHA1448.1.

[RD81] T. Borsdorff, J. Aan de Brugh, H. Hu *et al.*; Measuring Carbon Monoxide With TROPOMI: First Results and a Comparison With ECMWF-IFS Analysis Data. *Geophys. Res. Let.*; **45** (2018), 2826; doi:10.1002/2018GL077045.

[RD82] T. Borsdorff, J. Aan de Brugh, H. Hu *et al.*; Mapping carbon monoxide pollution from space down to city scales with daily global coverage. *Atmos. Meas. Tech. Discuss. (in review)*; **45** (2018), 2826; doi:10.5194/amt-2018-132.

[RD83] A. Inness, A.-M. Blechschmidt, I. Bouarar *et al.*; Data assimilation of satellite-retrieved ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS. *Atmospheric Chemistry and Physics*; **15** (2015) (9), 5275; doi:10.5194/acp-15-5275-2015. URL `https://www.atmos-chem-phys.net/15/5275/2015/`.

[RD84] F. Hase, T. Blumenstock, S. Dohe *et al.*; TCCON data from Karlsruhe, Germany, Release GGG2014R1. http://dx.doi.org/10.14291/tccon.ggg2014.karlsruhe01.R1/1182416 (2014); doi:10.14291/tccon.ggg2014.karlsruhe01.R1/1182416. TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

[RD85] R. Sussmann and M. Rettinger; TCCON data from Garmisch, Germany, Release GGG2014R0. http://dx.doi.org/10.14291/tccon.ggg2014.garmisch01.R0/1149299 (2014); doi:10.14291/tccon.ggg2014.garmisch01.R0/1149299. TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

[RD86] P. O. Wennberg, D. Wunch, Y. Yavin *et al.*; TCCON data from Jet Propulsion Laboratory, Pasadena, California, USA, Release GGG2014R0. http://dx.doi.org/10.14291/tccon.ggg2014.jpl01.R0/1149163 (2014); doi:10.14291/tccon.ggg2014.jpl01.R0/1149163. TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

[RD87] P. O. Wennberg, C. Roehl, J.-F. Blavier *et al.*; TCCON data from Jet Propulsion Laboratory, Pasadena, California, USA, Release GGG2014R0. http://dx.doi.org/10.14291/tccon.ggg2014.jpl02.R0/1149297 (2014); doi:10.14291/tccon.ggg2014.jpl02.R0/1149297. TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

[RD88] T. Blumenstock, F. Hase, M. Schneider *et al.*; TCCON data from Izana, Tenerife, Spain, Release GGG2014R0. http://dx.doi.org/10.14291/tccon.ggg2014.izana01.R0/1149295 (2014); doi:10.14291/tccon.ggg2014.izana01.R0/1149295. TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

[RD89] D. G. Feist, S. G. Arnold, N. John *et al.*; TCCON data from Ascension Island, Saint Helena, Ascension and Tristan da Cunha, Release GGG2014R0. http://dx.doi.org/10.14291/tccon.ggg2014.ascension01.R0/1149285 (2014); doi:10.14291/tccon.ggg2014.ascension01.R0/1149285. TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

[RD90] M. De Maziere, M. K. Sha, F. Desmet *et al.*; TCCON data from Reunion Island (La Reunion), France, Release GGG2014R0. http://dx.doi.org/10.14291/tccon.ggg2014.reunion01.R0/1149288 (2014); doi:10.14291/tccon.ggg2014.reunion01.R0/1149288. TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

[RD91] V. Sherlock, B. Connor, J. Robinson *et al.*; TCCON data from Lauder, New Zealand, 125HR, Release GGG2014R0. http://dx.doi.org/10.14291/tccon.ggg2014.lauder02.R0/1149298 (2014); doi:10.14291/tccon.ggg2014.lauder02.R0/1149298. TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

[RD92] F. Kasten and T. Young; Revised optical air mass tables and approximation formula. *Appl. Opt.*; **28** (1989), 4735.

2.4 Electronic references

There are no electronic references

3 Terms, definitions and abbreviated terms

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

3.1 Acronyms and abbreviations

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

S5P	Sentinel-5 Precursor
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Chartography
SICOR	Shortwave Infrared CO Retrieval
SMOW	Standard Mean Ocean Water
SNR	Signal-to-Noise Ratio
SPEC	Standard Performance Evaluation Corporation
SSD	Spectral Sampling Distance
SWIR	Shortwave Infrared
SZA	Solar Zenith Angle
TCCON	Total Carbon Column Observing Network
TES	Tropospheric Emission Spectrometer
TM4	Transport Model 4
TM5	Transport Model 5
TOA	Top Of model Atmosphere
TROPOMI	Tropospheric Monitoring Instrument
VIIRS	Visible Infrared Imager Radiometer Suite
VZA	Viewing Zenith Angle
WFM-DOAS	Weighting Function Modified-Differential Optical Absorption Spectroscopy

4 Remote Sensing of Carbon Monoxide

Carbon monoxide (CO) is an important atmospheric trace gas and in certain urban areas, it is a major atmospheric pollutant. Measurements of its global abundance improve our understanding of tropospheric chemistry and atmospheric long range transport [RD5, RD6, RD7, RD8]. Main sources of CO are combustion of fossil fuels, biomass burning, and atmospheric oxidation of methane and other hydrocarbons. Whereas fossil fuel combustion is the main source of CO at Northern mid-latitudes, the oxidation of isoprenes and biomass burning play an important role in the tropics. Due to the long lifetime of methane (CH_4), its oxidation provides a close-to uniform background on the global CO distribution. The most important sink of CO is its reaction with the hydroxyl radical OH. The CO lifetime is several weeks to several months, which makes CO a good tracer to study long range transport processes (e.g. [RD9]). Moreover, CO is one of the highest priority chemical species measured by the Monitoring Atmospheric Composition and Climate (MACC) project of the Global Monitoring for Environment and Security (GMES) program [RD10].

The first spaceborne measurements of CO were performed with the MAPS (Measurement of Air Pollution from Satellites) instrument during four flights of the space shuttle between 1981 and 1999 [RD11], and with the IMG (Interferometric Monitor for Greenhouse gases) instrument onboard ADEOS (Advanced Earth Observing System) in 1996 and 1997 [RD12]. Since 2000, long-term global data sets of CO are provided by MOPITT (Measurements of Pollution in the Troposphere) instrument (e.g. [RD13]), which measures in two spectral ranges at 2.3 μm and 4.7 μm, respectively, using correlation radiometry. At present, the operational CO MOPITT data product relies only on measurements around the fundamental 1-0 CO absorption band at 4.7 μm. This spectral range is also employed by three other spaceborne spectrometers which measure the infrared brightness of the Earth's surface and the atmosphere: (1) AIRS (Atmospheric Infrared Sounder [RD14]) launched in 2002 onboard the Aqua satellite, (2) TES (Tropospheric Emission Spectrometer [RD15] launched in 2004 onboard the Aura satellite and (3) IASI (Infrared Atmospheric Sounding Interferometer[RD16]) onboard a series of three METOP (Meteorological Operational) satellites. Generally, these thermal infrared measurements exhibit peak sensitivity to CO in the middle troposphere and are thus well suited to study long range atmospheric transport. However, depending on spectral resolution and the thermal contrast in the lower troposphere, the measurements show also sensitivity to CO in the lower troposphere [RD17].

CO total columns with sensitivity to the tropospheric boundary layer can be inferred from sunlight reflected by the Earth atmosphere in the 2.3 μm spectral range of the shortwave infrared (SWIR) part of the solar spectrum. The first overtone 2-0 absorption band of CO is situated between 2305 nm and 2385 nm. For clear sky measurements, this spectral range is subject to little atmospheric scattering and most of the measured light is thus reflected by the Earth's surface. SWIR measurements are therefore sensitive to the integrated amount of CO along the light path, including the contribution of the planetary boundary layer. This makes the SWIR spectral range particularly suitable for detecting surface sources of CO. Since the launch of SCIAMACHY (Scanning Imaging Absorption Spectrometer for Atmospheric Chartography [RD18] in the year 2002 on the Envisat satellite, a continuous time series of global CO SWIR measurements is available. Moreover, first results were reported recently using MOPITT measurements in the SWIR [RD19] and a combination of SWIR and the 4.7 μm CO absorption band [RD20]. However, new space-borne instrumentation is required to ensure continuity of SWIR measurements of CO in the future. In this respect, the Sentinel 5 Precursor mission (S5P [RD21]) provides the opportunity to extend this unique long-term global data set of CO using the same type of measurement. On top of that, the major scientific objective of these mission is to participate in the development and improvement of air quality model processes and data assimilation in support of operational services, including air quality forecasting and protocol monitoring. This scientific objective requires the combination of new instrumentation with better radiometric performances and higher spatial sampling, and the development of fast and accurate retrieval algorithms.

4.1 Algorithm heritage

Several fast algorithms are used to retrieve CO column from SCIAMACHY SWIR measurements, including the Weighting Function Modified-Differential Optical Absorption Spectroscopy (WFM-DOAS) approach ([RD22] and references therein) the Iterative Maximum A Posteriori (IMAP) approach [RD23], and the Iterative Maximum Likelihood Method (IMLM) approach ([RD24], and references therein). These algorithms retrieve vertically integrated CO column density over land and above clouds over oceans. Over ocean, the surface albedo is too low to retrieve CO under clear sky conditions. For numerical efficiency, scattering by aerosols and clouds is not considered in the radiative transfer of the IMLM retrieval approach. Both Buchwitz et al. [RD25] and Gloudemans et al. [RD24] use a priori methane information to characterise the light path through the atmosphere. The different approaches allow the detection of strongly polluted areas, the seasonal variability

of global atmospheric CO and the long range transport of CO. The retrievals have been compared with chemical transport models (CTM, e.g. [RD26, RD24]), with MOPITT retrievals (e.g. [RD22, RD27]) and ground-based FTIR (Fourier Transform Infrared) measurements [RD28, RD29]. Since CO is a weak absorber and SCIAMACHY exhibits a low signal-to-noise ratio in the 2.3 μm region, the retrieval of single-sounding CO column density is subject to large retrieval noise, typically on the order of 10-100 % [RD30]. Hence, to compare SCIAMACHY retrievals with ground-based measurements, the averaging of multiple SCIAMACHY measurements is needed. De Laat et al. [RD29] demonstrated that, at the Northern latitudes and for a typical sampling area of $8^o \times 8^o$, the absolute difference between the mean IMLM SCIAMACHY CO retrievals and ground-based measurements are close to or fall within the 2σ precision of 2×10^{17} molec./cm^2. Thus on the Northern Hemisphere, SCIAMACHY can observe CO adequately for the given sampling size. At mid to high latitudes of the Southern Hemisphere, the IMLM algorithm systematically underestimates the ground based measurements by $1 - 5 \times 10^{17}$ molec./cm^2. Reasons for this bias are currently under investigation. Next to the large sampling areas and the Southern Hemisphere bias, one major limitation of the current SWIR CO retrieval product is that for the individual soundings, no information is provided on the vertical sensitivity of the CO column retrieval, the so-called column averaging kernel. Column averaging kernels are essential information to validate SWIR CO retrievals with other independent retrievals or CTM simulations. Last but not least, it is also an important piece of information for data assimilation in air quality forecasting employing variational schemes.

Based on SCIAMACHY heritage, we present a modified SWIR CO retrieval (SICOR) approach for S5P, which can be applied to clear sky observations over land and cloudy observations over both land and oceans. The algorithm is based on the heritage of the CO retrieval algorithm for SCIAMACHY [RD24] but improves current CO retrieval approaches for cloudy and aerosol loaded atmospheres. A typical SWIR spectrum is illustrated on the top panel of Fig. 1. It shows the total transmittance of solar light along its path from the Sun to the surface to the satellite. The transmittance is simulated using the Beer's extinction law. In the band, the relevant absorbing species are H_2O, CO and CH_4, with the optical depth of CO generally much smaller than those of H_2O and CH_4.

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

4.2 Carbon Monoxide level-2 requirements

To improve our present knowledge on CO on a global scale, satellite measurements of the total CO column are needed within an accuracy of $< 15\%$ and a precision with $\leq$ 10 % even for background CO abundance and low surface reflection in the shortwave infrared spectral range [RD21, RD31]. For the CO error budget, we assume that instrument and forward model errors contribute equally to the error budget and that all error terms add up quadratically. Herewith, both instrument and forward model errors must not exceed 8 %. These CO level-2 requirements should be considered as thresholds.

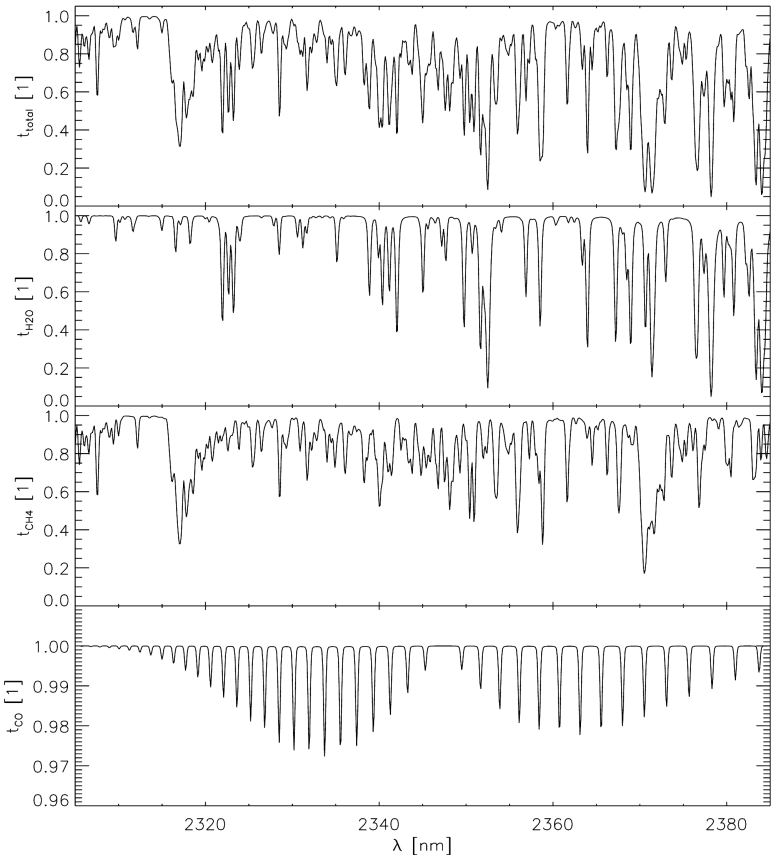

Figure 1: SWIR spectral transmittance along the light path of the solar beam reflected by the Earth surface into the instrument viewing direction. Simulations are performed for viewing zenith angle $(\mathrm{VZA}) = 0°$, and a solar zenith angle $(\mathrm{SZA}) = 30°$, and by assuming a US standard atmospheric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

5 Algorithm Description

To retrieve atmospheric CO abundances from TROPOMI SWIR measurements as a near real-time (NRT) data product, we follow an algorithm concept which is summarised in Fig. 2. It is a further development of the retrieval method of Vidot et al. [RD32] and it includes a significant improvement with respect to the description of atmospheric scattering. The retrieval algorithm requires several input fields:

- The measured Earth radiance and solar irradiance spectra including noise estimate, solar and viewing geometry, and information of geo-location.
- ECMWF temperature, water vapor and pressure profiles and geo-potential height.
- An estimate of the CH_4 field using a chemistry transport model, e.g. Transport Model 5 (TM5, [RD33]).
- An estimate of the CO column from a chemistry transport model (e.g. TM5).

The retrieval is performed in two steps: first as part of the SWIR preprocessing module (see Appendix B), the vertically integrated amount of methane is retrieved from a dedicated fit window of the SWIR band between 2315 and 2324 nm using a non-scattering radiative transfer model. The extent of lightpath shortening and enhancement due to atmospheric scattering by clouds and aerosols can be indicated by comparing the retrieved CH_4 column with a priori knowledge. If the difference ΔCH_4 exceeds a certain threshold, observations are strongly contaminated by clouds and are rejected. In a second step, the SICOR full physics retrieval approach is used to infer CO columns from the adjacent spectral window, 2324-2338 nm. Here, the methane absorption features are used to infer information on atmospheric scattering by clouds and aerosols, which passed the cloud filter, together with the atmospheric CO and H_2O abundances, surface albedo and a spectral calibration of the reflectance spectrum. The scattering layer has a triangular height distribution of fixed geometrical thickness, and its optical depth and height are parameters to be retrieved. This step of the retrieval relies on accurate a priori knowledge of CH_4 which will be provided within an accuracy of ± 3 % by a dedicated methane forecast using the TM5 atmospheric transport model. The atmospheric scattering is described by a two-stream radiative transfer model. Finally, the retrieval product consists of a CO column estimate including a column averaging kernel and a random error estimate. A detailed description of input and output fields are given in Sec. 9.

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

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \mathbf{e}_\mathbf{y} \, . \tag{1}$$

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

$$\mathbf{F}(\mathbf{x}, \mathbf{b}) = \mathbf{F}(\mathbf{x}_o, \mathbf{b}) + \frac{\partial \mathbf{F}}{\partial \mathbf{x}}(\mathbf{x}_o, \mathbf{b})\{\mathbf{x} - \mathbf{x}_o\} + O(\mathbf{x}^2) \, , \tag{2}$$

where $O(\mathbf{x}^2)$ indicates second and higher order contributions of the expansion. This iteration approach shows satisfying convergence properties for moderately non-linear problems [RD34]. In the following sub-sections, both the SICOR inversion approach and the forward model are described in more depth.

5.1 Forward model

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

$$I_i^{\mathrm{TOA}} = s * I^{\mathrm{TOA}} = \int s(\lambda_i, \lambda) \, I^{\mathrm{TOA}}(\lambda) d\lambda \, , \tag{3}$$

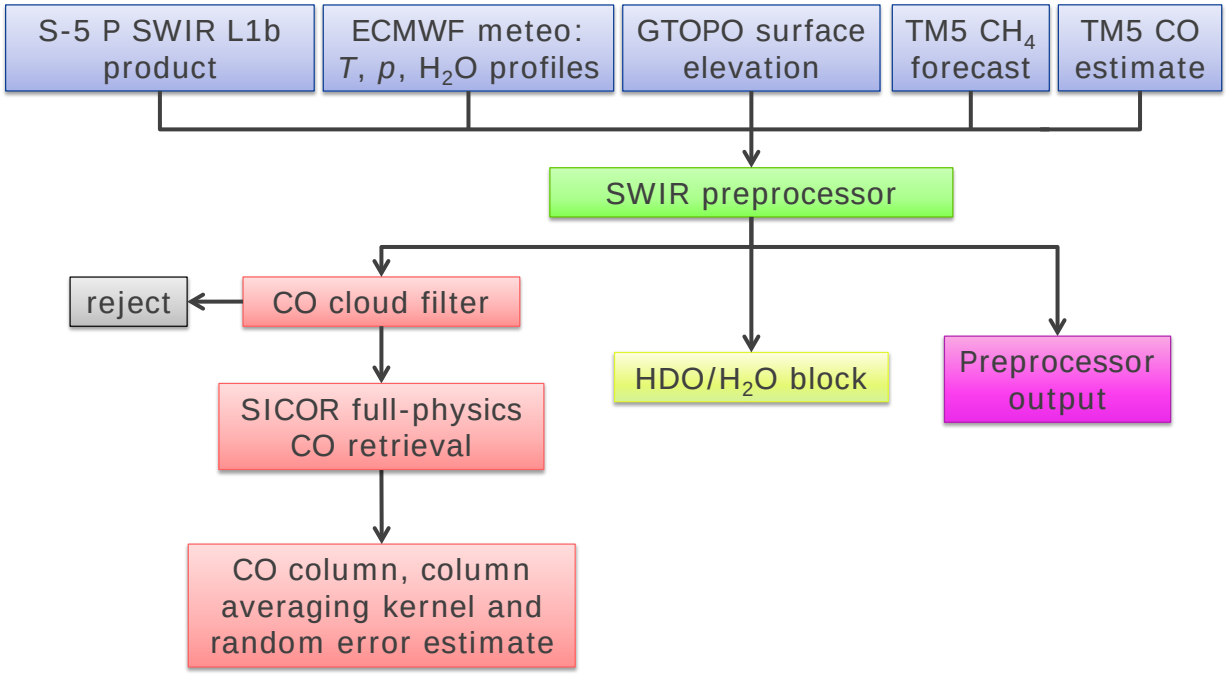

Figure 2: Overall structure of the CO retrieval algorithm. For the structure of the SWIR pre-processing algorithm see Appendix B.

and a corresponding expression for the solar spectrum. Here, s is the ISRF, λ_i is the wavelength assigned to the spectral pixel i of the measurement and I_i^{TOA} is corresponding radiance simulated by a radiative transfer model. The reflectance is the ratio of the radiance and the solar spectrum. Thus

$$F_i = r_i = \frac{s * I^{\mathrm{TOA}}}{s * F_o} \,, \tag{4}$$

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

$$\frac{dr_i}{d(\Delta\lambda)} = \frac{(s' * I^{\mathrm{TOA}})(s * F_o) - (s' * F_o)(s * I^{\mathrm{TOA}})}{(s * F_o)^2} \,. \tag{5}$$

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

State-of-the-art radiative transfer models account for multiple scattering in multiple propagation directions (streams) including the polarisation of light. For our application, the computational effort of such simulations is high and thus, it is not feasible to use these models for operational data processing of the CO column density. So, approximation methods have to be used. In this respect, we propose to ignore atmospheric Rayleigh scattering which contributes less than < 0.15 % to the total signal [RD35] and to use a numerically efficient two-stream radiative transfer model where the effect of clouds and aerosols are represented in the model atmosphere by a triangular scattering layer with a center height z_{scat} and a fixed width. Singly scattered light is treated exactly by the model, whereas multiple scattered light is described in an approximated manner.

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

The generalised two-stream method

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

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

$$\begin{pmatrix} S_{n+1} \\ F^{\downarrow}_{n+1} \\ F^{\uparrow}_{n} \end{pmatrix} = \begin{pmatrix} a_{1,n} & 0 & 0 \\ a_{2,n} & a_{4,n} & a_{5,n} \\ a_{3,n} & a_{5,n} & a_{4,n} \end{pmatrix} \begin{pmatrix} S_n \\ F^{\downarrow}_{n} \\ F^{\uparrow}_{n+1} \end{pmatrix} \tag{6}$$

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

$$\begin{aligned} S_1 &= \mu_o F_o \\ F^{\downarrow}_1 &= 0 \\ F^{\uparrow}_{N+1} &= A_s (F^{\downarrow}_{N+1} + S_{N+1}) \,. \end{aligned} \tag{7}$$

where A_s is the surface albedo and $\mu_o = \cos(\Theta_o)$ with the solar zenith angle Θ_o. In the following, we discuss the practical improved flux method (PIFM) of Zdunkowski at al. [RD36] for a three layer model atmosphere to simplify matters. The approach can be generalised in a straight forward manner to a N-layer model atmosphere, which is the basis for the software implementation. Combining the internal and external boundary constraints for the three layer system, we obtain the matrix equation

$$\mathbf{MF} = \mathbf{C} \tag{8}$$

with the sparse matrix

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -a_{3,1} & -a_{5,1} & 1 & 0 & 0 & -a_{4,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ -a_{1,1} & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -a_{2,1} & -a_{4,1} & 0 & 0 & 1 & -a_{5,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -a_{3,2} & -a_{5,2} & 1 & 0 & 0 & -a_{4,2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -a_{1,2} & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -a_{2,2} & -a_{4,2} & 0 & 0 & 1 & -a_{5,2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -a_{3,3} & -a_{5,3} & 1 & 0 & 0 & -a_{4,3} \\ 0 & 0 & 0 & 0 & 0 & 0 & -a_{1,3} & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -a_{2,3} & -a_{4,3} & 0 & 0 & 1 & -a_{5,3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -A_s & -A_s & 1 \end{pmatrix}, \tag{9}$$

the flux vector

$$\mathbf{F} = \left(\; S_1, \quad F_1^{\downarrow}, \quad F_1^{\uparrow}, \quad S_2, \quad F_2^{\downarrow}, \quad F_2^{\uparrow}, \quad S_3, \quad F_3^{\downarrow}, \quad F_3^{\uparrow} \quad S_4, \quad F_4^{\downarrow}, F_4^{\uparrow} \; \right)^T \tag{10}$$

and the right hand side

$$\mathbf{C} = \left(\; \mu_o F_o, \quad 0, \quad 0, \quad 0, \quad 0, \quad 0, \quad 0, \quad 0, \quad 0, \quad 0, \quad 0, \quad 0 \; \right)^T . \tag{11}$$

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

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

$$I^{\mathrm{TOA}}(\mu_v) = \frac{F_{N+1}^{\uparrow}}{\pi} \exp(-\tau_{\mathrm{tot}}/\mu_v) + \frac{1}{\mu_v} \int_0^{\tau_{\mathrm{tot}}} d\tau \; J(\tau,\mu_v) \; \exp(-\tau/\mu_v) \, , \tag{12}$$

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

$$J(\tau,\mu_v) = \frac{\omega(\tau)}{2} \int_{-1}^{+1} d\mu' P(\tau;\mu_v,\mu') I(\tau,\mu') + \frac{\omega(\tau)}{4\pi} P(\tau;\cos\Theta_v) F_o \exp(-\tau/\mu_o) \, . \tag{13}$$

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

$$F^{\downarrow\uparrow} = 2\pi \int_o^{\pm 1} d\mu \; \mu \; I(\tau,\mu) = 2\pi \mu_{\downarrow\uparrow} I_{\downarrow\uparrow} \tag{14}$$

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

$$\beta(\tau,\mu) = \frac{1}{2} \int_{-1}^{0} d\mu' P(\tau;\mu,\mu') = 1 - \frac{1}{2} \int_0^1 d\mu' P(\tau;\mu,\mu') \quad \text{for } \mu > 0 \, , \tag{15}$$

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

$$\bar{\beta}(\tau) = \frac{1}{2} \int_0^1 d\mu \int_{-1}^{0} d\mu' P(\tau;\mu,\mu') \tag{16}$$

(see e.g. Wiscombe and Grams, 1976, [RD37]). The expansion of the phase function in Legendre polynomials P_l,

$$P(\tau;\mu,\mu') = \sum_{l=0}^{\infty} \chi_l(\tau) P_l(\mu) P_l(\mu') \tag{17}$$

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

$$\beta(\tau,\mu) = \sum_{l=0}^{\infty} \frac{\chi_l(\tau)}{2} P_l(\mu) \int_{-1}^{0} d\mu' P_l(\mu') \tag{18}$$

and

$$\bar{\beta}(\tau) = \sum_{l=0}^{\infty} \frac{\chi_l(\tau)}{2} \int_0^1 d\mu P_l(\mu) \int_{-1}^0 d\mu' P_l(\mu') \,. \tag{19}$$

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

$$J(\tau,\mu_v) = \frac{\omega}{2\pi}\{U_\uparrow[1-\bar{\beta}(\tau)]F^\uparrow(\tau) + U_\downarrow\bar{\beta}(\tau)F^\downarrow(\tau)\} + \frac{\omega}{4\pi\mu_o}P(\cos(\Theta_v))S_o\exp(-\tau/\mu_o) \,. \tag{20}$$

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

$$F^{\downarrow\uparrow}(\tau) = \frac{F_i^{\downarrow\uparrow} + F_{i+1}^{\downarrow\uparrow}}{2} \quad \text{for } \tau_i < \tau < \tau_{i+1} \,. \tag{21}$$

Thus, we obtain

$$\begin{aligned}\int_0^{\tau_{\rm tot}} d\tau\, J(\tau,\mu_v)\, \exp(-\tau/\mu_v) = \sum_{n=1}^{N} \frac{\omega_n}{4\pi\mu_o}\tilde{\mu}P_n t_n(\tilde{\mu})S_{n-1} + \\ U^\uparrow \sum_{n=1}^{N} \frac{\omega_n}{4\pi}(1-\beta_n)t_n(\mu_v)[F_n^\uparrow + F_{n+1}^\uparrow] + U^\downarrow \sum_{n=1}^{N} \frac{\omega_n}{4\pi}\beta_n t_n(\mu_v)[F_n^\downarrow + F_{n+1}^\downarrow]\end{aligned} \tag{22}$$

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

$$t_n(\mu) = \exp(-\frac{\tau_{n-1}}{\mu_v})[1-\exp(-\frac{\Delta\tau_n}{\mu})] \qquad \text{for } n = 1,\cdots,N \,. \tag{23}$$

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

$$I^{\rm TOA} = \mathbf{R}^T\mathbf{F} \tag{24}$$

with the so-called response vector

$$\mathbf{R} = \frac{1}{4\pi}(u_1, v_1, w_1, \cdots, u_{N+1}, v_{N+1}, w_{N+1}) \tag{25}$$

and its coefficients

$$\begin{aligned} u_n &= \frac{\omega_n}{4\pi\mu_o}\tilde{\mu}P_n t_n(\tilde{\mu}) \\ v_n &= \frac{U^\downarrow}{4\pi}\left[\omega_n\beta_n t_n(\mu_v) + \omega_{n-1}\beta_{n-1}t_{n-1}(\mu_v)\right] \\ w_n &= \frac{U^\uparrow}{4\pi}\left[\omega_n(1-\beta_n)t_n(\mu_v) + \omega_{n-1}(1-\beta_{n-1})t_{n-1}(\mu_v)\right] \end{aligned} \tag{26}$$

with $t_0 = t_{N+1} = 0$. Equation (24) is an essential step towards the linearisation of the model with respect to absorption and scattering properties of the model atmosphere, which is discussed in the following.

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

The linearisation of the two-stream method is less evident than the linearisation of the transmission model. Due to multiple scattering, the functional dependence of I^{TOA} on optical depth, single-scattering albedo and on the scattering phase function characteristics is not obvious (see Eqs. (8) and (24)). In this context, the forward-adjoint perturbation theory provides a method, which is frequently used to linearize multiple scattering radiative transfer simulations with respect to absorption and scattering properties of the model atmosphere (e.g. [RD38, RD39, RD40, RD41, RD42, RD43, RD44, RD45, RD46]). Here, we summarise the main features of the perturbation theory, which is subsequently used to motivate our approach to linearize the two-stream method.

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

$$\mathbf{L}I = Q\,, \tag{27}$$

where

$$\mathbf{L} = \int\limits_{4\pi} d\tilde{\Omega} \left\{ \left[\mu \frac{\partial}{\partial z} + \beta_e(r) \right] \delta(\mathbf{\Omega} - \tilde{\mathbf{\Omega}}) - \frac{\beta_s(r)}{4\pi} P(r, \tilde{\mathbf{\Omega}}, \mathbf{\Omega}) + R_s(\tilde{\mathbf{\Omega}}, \mathbf{\Omega}) \delta(z) \Theta(\mu) |\mu| \Theta(-\tilde{\mu}) |\tilde{\mu}| \right\}, \tag{28}$$

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

$$\delta(\mathbf{\Omega} - \tilde{\mathbf{\Omega}}) = \frac{1}{\sin\theta} \delta(\theta - \tilde{\theta})\, \delta(\varphi - \tilde{\varphi}). \tag{29}$$

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

$$Q(z, \mathbf{\Omega}) = \mu_0 F_0 \delta(z - z_{\mathrm{TOA}})\, \delta(\mathbf{\Omega} - \mathbf{\Omega}_0)\,. \tag{30}$$

Here, z_{TOA} is the height of the upper boundary of the model atmosphere and $\mathbf{\Omega}_0 = (\theta_0, \varphi_0)$ describes the direction of the solar beam. θ_0 is the local solar zenith angle and φ_0 the local solar azimuth angle. Because surface reflection is included in the radiative transfer operator $\mathbf{L}$ and the incoming solar source is described by the source Q, the radiation field I is subject to the free surface boundary conditions of no incoming radiation at the top and bottom of the model atmosphere.

In the context of satellite remote sensing, one is interested in the simulation of the intensity at the top of the model atmosphere in the viewing direction of the instrument, thus $I^{\mathrm{TOA}} = I(\mathbf{z}_{\mathrm{TOA}}, \mathbf{\Omega}_{\mathrm{v}})$. Here, $\mathbf{\Omega}_{\mathrm{v}} = (\theta_{\mathrm{v}}, \varphi_{\mathrm{v}})$ describes the viewing direction of the satellite instrument. For the forward-adjoint perturbation theory, it is crucial to express the observation I^{TOA} by the inner product of the form

$$I^{\mathrm{TOA}} = \langle R | I \rangle = \int\limits_{o}^{z_{TOA}} dz \int\limits_{4\pi} d\Omega\; R(z, \mathbf{\Omega}) I(z, \mathbf{\Omega})\,. \tag{31}$$

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

$$R(z, \mathbf{\Omega}) = \delta(z - z_{TOA}) \delta(\mathbf{\Omega} - \mathbf{\Omega}_{\mathrm{v}})\,. \tag{32}$$

Additionally, the linearisation of radiative transfer with the forward-adjoint perturbation theory requires the solution of the adjoint transfer equation

$$\mathbf{L}^{\dagger} I^{\dagger} = R \,, \tag{33}$$

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

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

$$I^{\mathrm{TOA}} = \langle I^{\dagger} | Q \rangle \,. \tag{34}$$

This equation shows the physical meaning of the adjoint field: $I^{\dagger}$ can be interpreted as the importance of light [RD47, RD48] within the atmosphere for a given measurement.

With the solution of the forward and the adjoint transfer problem I_0 and $I_0^{\dagger}$ for an atmospheric state $\mathbf{x}_o$, it is possible to calculate the derivatives of I^{TOA} with respect to a optical parameter x at x_o by

$$\frac{\partial I^{\mathrm{TOA}}}{\partial x}(x_0) = \langle I_0^{\dagger} | \mathbf{L}' I_0 \rangle + \langle I_0^{\dagger} | Q' \rangle + \langle R' | I_0 \rangle \tag{35}$$

with the partial derivative

$$\mathbf{L}' = \lim_{\Delta x \to 0} \frac{\Delta \mathbf{L}}{\Delta x} \,, \tag{36}$$

$R' = \partial R / \partial x$ and $Q' = \partial Q / \partial x$ (e.g. [RD49, RD44]).

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

$$\langle \mathbf{x} | \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} \,. \tag{37}$$

Moreover, the adjoint matrix $\mathbf{M}^{\dagger}$ is given by its transposed, M^T, and so the response vector $\mathbf{R}$ defines the adjoint problem

$$\mathbf{M}^T \mathbf{F}^{\dagger} = \mathbf{R} \,, \tag{38}$$

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

$$\frac{\partial I^{\mathrm{TOA}}}{\partial x} = \mathbf{F}^{\dagger,T} (\mathbf{M}' \mathbf{F}) + \mathbf{F}^{\dagger,T} \mathbf{C}' + \mathbf{R}'^{T} \mathbf{F} \tag{39}$$

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

$$\frac{\partial I^{\mathrm{TOA}}}{\partial x} = \mathbf{F}^{\dagger,T} (\mathbf{M}' \mathbf{F}) + \mathbf{R}'^{T} \mathbf{F} \,. \tag{40}$$

In general, x represents any atmospheric parameter which influences the solution $\mathbf{F}$. In our case, this is the optical depth $\Delta\tau_n$, the single-scattering albedo ω_n and the surface albedo A_s. Equation (40) can be numerically implemented in a straight forward manner, which is described in detail in Appendix A.

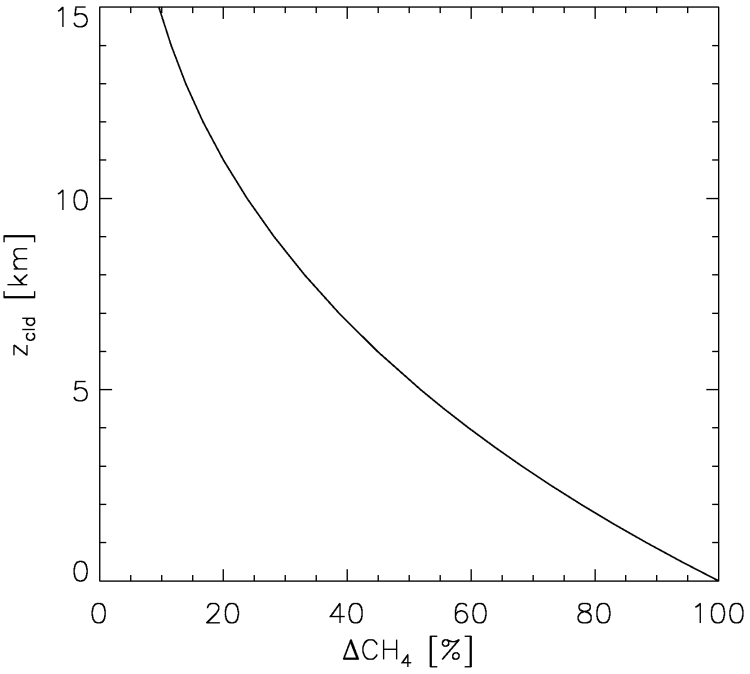

Figure 3: Relative CH_4 column above a cloud top height z_{cld} with respect to the total column amount, using the US standard model atmosphere.

5.2 Inversion

5.2.1 Methane cloud filter

Assuming accurate a priori knowledge, one can use methane, which is retrieved for a non-scattering model atmosphere, to filter for the presence of optically thick water clouds. A corresponding non-scattering retrieval algorithm is described in Appendix B as part of the SWIR pre-processing. To illustrate the functionality of such a cloud filter, Fig. 3 shows the amount CH_4 above a cloud with a cloud top height z_{cld}. Above 2, 5 and 10 km about 80, 50 and 20 % of the total amount of methane is located. For an optically thick cloud, the indicated fraction of methane corresponds to the methane column which can be retrieved from SWIR satellite measurements assuming a non-scattering atmosphere. The situation differs when we consider optically thin clouds. Here, light can penetrate the cloud and photon trapping between the cloud and the Earth surface may cause an effective pathlength enhancement with respect to a clear-sky situation. In this case, a non-scattering retrieval of CH_4 will overestimate the true CH_4 column.

Figure 4 shows the probability density function (PDF) of the non-scattering CH_4 retrieval error and its cumulative distribution (CPDF) for one year of GOSAT data (2010). As references we use collocated CH_4 columns from TM5 model simulations. The maximum of the PDF around small retrieval error indicates scenes, which are affected little by clouds. This maximum is present in both the ocean and land PDF. For about 80 % of all observations, the methane abundance is underestimated by the non-scattering retrieval due to the presence of optically thick clouds. Here, the ocean PDF shows a relatively high probability of retrieval errors between -20 % and -5 % due to the presence of low stratiform clouds over ocean. For land pixels, this type of cloudiness occurs less frequent. Finally, 20 % of all cases show an overestimation of methane by the non-scattering retrieval indicating an effective pathlength enhancement. Overall, the figure clearly demonstrates the ability of the difference between the non-scattering retrieval and methane a priori knowledge for cloud filtering of the TROPOMI dataset.

To implement a CH_4 non-scattering cloud filter based on this principle, we employ TROPOMI SWIR measurements in the spectral range 2315-2324 nm with strong methane absorption (see Fig. 1). We fit the total column of CH_4, CO, H_2O and HDO together with a surface albedo A_s, its linear dependence on wavelength, and a spectral offset of the measurements. For the inversion, a standard least squares approach is used in combination with the transmission radiative transfer model described in Section B. For data selection, we

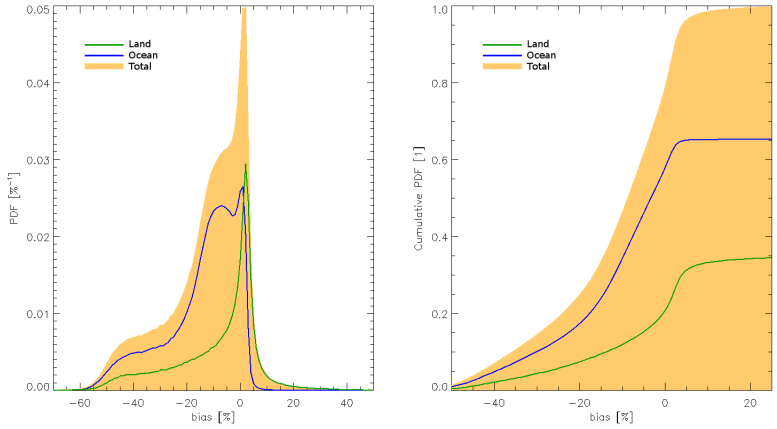

Figure 4: Probability density function (left panel) and cumulative probability density function (right panel) of the non-scattering methane error for one year of GOSAT observations (2010) with respect to corresponding TM5 model simulations. The figure differentiates the contribution of ocean and land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

accept all cases with a non-scattering CH_4 error $\Delta CH_4 > 25$ %, and so we filter out observations with optically thick and high clouds.

5.2.2 Tikhonov regularisation for CO column retrieval

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

For this purpose, we employ the Tikhonov regularisation technique [RD50, RD51] embedded in a Gauss-Newton iteration scheme. For each iteration step, the solution is given by the least square solution with an additional side constraint,

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \left\{ \|\mathbf{S}_\mathbf{y}^{-1/2}(\mathbf{F}(\mathbf{x}) - \mathbf{y})\|^2 + \gamma^2 \|\mathbf{L}_1 \mathbf{x}\|^2 \right\} . \tag{41}$$

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

$$\mathbf{x}_{CO} = \rho_{CO} / \rho_{CO}^{ref} . \tag{42}$$

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

$$\mathbf{x} = (\mathbf{x}_{CO}, c_{H_2O}, c_{HDO}, A_{surf}, \Delta A_s, z_{scat}, \tau_{scat}, \Delta\lambda) . \tag{43}$$

$\mathbf{L}_1$ describes the operator for the vertical derivative of the CO column.

$$\mathbf{L}_1 \mathbf{x} = \frac{d\mathbf{x}_{CO}(z)}{dz} \tag{44}$$

For an equidistant model grid with layer thickness Δz, the discrete representation of the first derivative of the CO column is

$$\mathbf{L}_1\mathbf{x} = \frac{1}{\Delta z}\begin{pmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \end{pmatrix}\mathbf{x}_{\mathrm{CO}} \,. \tag{45}$$

The choice of the regularisation parameter γ is of crucial importance for the inversion. If γ is chosen too large, the noise contribution to the solution of the measurement is low, but the least squares residual norm deviates significantly from its minimum. On the other hand, if γ is chosen too small, the measurement is fitted well but the solution norm is high, and so, the solution is overwhelmed by noise. The part of the profile that is dominated by noise defines the effective null-space of the problem. Thus, γ should be chosen such that the two minimisations are well balanced. For atmospheric profile retrievals, it has been demonstrated in several studies that an appropriate value for γ can be found using the L-curve method [RD52, RD53]. However, the method becomes less stable when the information on the vertical distribution of the trace gas is limited as in our case, where about one independent piece of information of the CO profile can be retrieved. Therefore, we decided to regularize the solution in Eq. (41) with respect to its degree of freedom for signal (DFS) of the retrieved CO profile. For Eq. (41), this corresponds to a regularisation parameter $\gamma \to \infty$ (e.g. [RD54]).

For a forward model which is linearized around the state vector $\mathbf{x}_o$,

$$\mathbf{F}(\mathbf{x},\mathbf{b}) = \mathbf{F}(\mathbf{x}_o,\mathbf{b}) + \mathbf{K}(\mathbf{x}-\mathbf{x}_o) \,, \tag{46}$$

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

$$\hat{\mathbf{x}} = \mathbf{G}\tilde{\mathbf{y}} \tag{47}$$

with the gain matrix

$$\mathbf{G} = \left(\mathbf{K}^T\mathbf{S}_\mathbf{y}^{-1}\mathbf{K} + \gamma^2\mathbf{L}_1^T\mathbf{L}_1\right)^{-1}\mathbf{K}^T\mathbf{S}_\mathbf{y}^{-1} \tag{48}$$

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

$$\mathbf{A} = \frac{\partial\hat{\mathbf{x}}}{\partial\mathbf{x}_{\mathrm{true}}} = \mathbf{G}\mathbf{K} \,, \tag{49}$$

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

$$\hat{\mathbf{x}} = \mathbf{A}\mathbf{x}_{\mathrm{true}} + \mathbf{e}_\mathbf{x} \,, \tag{50}$$

where $\mathbf{e}_\mathbf{x} = \mathbf{G}\mathbf{e}_\mathbf{y}$ represents the error on the retrieved CO profile caused by the error $\mathbf{e}_\mathbf{y}$. For a proper error characterisation of the inversion, the retrieval error covariance matrix $\mathbf{S}_\mathbf{x}$ is needed. This can be calculated from the measurement covariance matrix $\mathbf{S}_\mathbf{y}$ by

$$\mathbf{S}_\mathbf{x} = \mathbf{G}\mathbf{S}_\mathbf{y}\mathbf{G}^T \,. \tag{51}$$

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

$$\chi^2_n = \frac{1}{N}\sum_i \frac{(y_i - F_i(\mathbf{x}_n))^2}{e_i^2} \,. \tag{52}$$

Here e_i is the standard deviation (error) of the i^{th} measurement, $e_i^2 = \mathbf{S}_{\mathbf{y},ii}$. Convergence is achieved when $|\chi_n^2 - \chi_{n-1}^2| < \epsilon$ for a given choice of ϵ. The value of ϵ can only be determined in a reliable manner using real measurements during the mission consolidation phase. Furthermore, a minimum number of iterations is always performed. If convergence is not achieved within a certain number of iterations, the inversion is terminated and the measurement is rejected.

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

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

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

$$\hat{c} = \mathbf{A}_c\mathbf{x}_{\text{true}} + \mathbf{e}_c \,, \tag{54}$$

where $\mathbf{A}_c = \mathbf{CA}$ is the column averaging kernel, and $\mathbf{e}_c$ is the error on the retrieved column. The retrieval noise on the retrieved column is given by the standard deviation

$$\sigma_c = \sqrt{\mathbf{C}\mathbf{S}_\mathbf{x}\mathbf{C}^T} \,. \tag{55}$$

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

Borsdorff et al. [RD54] showed that for the definition of the CO state vector $\mathbf{x}_{\text{CO}}$ in Eq. (42) and for a regularisation parameter $\gamma \to \infty$, the solution of the Tikhonov minimisation problem (41) is identical to an unregularised least squares approach,

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \|\mathbf{S}_\mathbf{y}^{-1/2}(\mathbf{F}(\mathbf{x}) - \mathbf{y})\|^2 \,, \tag{56}$$

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

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

$$K_i^{\text{col}} = \frac{\partial F_i}{\partial c} = \sum_j K_{ij}^{\text{prof}} \frac{\rho_j^{\text{ref}}}{c^{\text{ref}}} \tag{57}$$

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

$$\hat{\mathbf{x}} = \mathbf{G}_{\text{lsq}}\mathbf{y} \tag{58}$$

with the gain matrix

$$\mathbf{G}_{\text{lsq}} = (\mathbf{K}_{\text{lsq}}^T \mathbf{S}_y^{-1} \mathbf{K}_{\text{lsq}})^{-1} \mathbf{K}_{\text{lsq}}^T \mathbf{S}_y^{-1} \,. \tag{59}$$

To infer the column averaging kernel defined in Eq. (50) from the least squares solution, we follow the approach by [RD54]: We extract the gain vector $\mathbf{g}^{\text{col}}$ from the gain matrix of the least-squares fit $\mathbf{G}_{\text{lsq}}$ which belongs to the fitted CO column. Interpreting the profile scaling approach as a particular case of a regularised profile retrieval using Tikhonov regularisation of the first order with an infinite strong regularisation, [RD54] showed that the total column averaging kernel can be calculated as

$$\mathbf{A}^{\text{col}} = \mathbf{g}^{\text{col}}\mathbf{K}^{\text{prof}} \,. \tag{60}$$

Here, the spatial sampling of the column averaging kernel depends on the corresponding sampling of $\mathbf{K}^{\text{prof}}$. For the full mathematical proof, the reader is referred to [RD54].

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

$$e_n = (\mathbf{C} - \mathbf{A}_c)\rho \tag{61}$$

This is the so-called null space contribution of a profile ρ and is also known as smoothing error of the retrieval [RD34]. [RD54] have shown that the null space contribution of the reference profile ρ_{ref} is zero, $\mathbf{A}^{\text{col}}\rho_{\text{ref}} = \mathbf{C}\rho_{\text{ref}}$. So for the particular case that the relative profile used for the scaling is correct, the retrieved column can be interpreted as an estimate of the true column, as expected. In all other cases, equation (60) and (61) have to be accounted for a proper interpretation of the retrieval product.

5.2.3 Tikhonov regularisation for stability

The solution of the minimisation problem (Eq. (56)) is vulnerable to state variables to which the spectrum is insensitive. For example, in case of a measurement over a scene which is fully overacted by a optically thick cloud, the measurement has no sensitivity to the surface albedo. In this case, one or more eigenvalues of $\mathbf{K}^T\mathbf{K}$ approach zero, leading to extremely large values in $\mathbf{G}$. Therefore, a side constraint similar to that of Eq. (41) is desirable. However, a first order derivative matrix $\mathbf{L}_1$ with this state vector makes no sense, because the state vector does not contain vertically resolved columns.

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

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \left\{ ||\mathbf{S}_{\mathbf{y}}^{-1/2}(\mathbf{F}(\mathbf{x}) - \mathbf{y})||^2 + \gamma^2 ||\mathbf{W}\mathbf{x}||^2 \right\} \tag{62}$$

with state vector $\mathbf{x}$ defined as

$$\mathbf{x} = (c_{\text{CO}}, c_{\text{H}_2\text{O}}, c_{\text{HDO}}, A_{\text{surf}}, \Delta A_s, z_{\text{scat}}, \tau_{\text{scat}}, \Delta\lambda) \, . \tag{63}$$

and $\mathbf{W}$ is defined as

$$\begin{aligned} W_{ij} &= 1 \quad \text{for } i = j \in (A_{\text{surf}}, \Delta A_s, z_{\text{scat}}, \tau_{\text{scat}}) \\ &= 0 \quad \text{otherwise} \, . \end{aligned} \tag{64}$$

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

$$\mathbf{G} = \left(\mathbf{K}^T \mathbf{S}_{\mathbf{y}}^{-1} \mathbf{K} + \gamma^2 \mathbf{W}^T \mathbf{W} \right)^{-1} \mathbf{K}^T \mathbf{S}_{\mathbf{y}}^{-1} \, . \tag{65}$$

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

5.2.4 Step control

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

This issue is treated by introducing a step size reduction factor λ (also explained by Butz et al. [RD55]). The state retrieved during iteration i is the weighted mean of the state retrieved during iteration $i-1$ and the calculated state during iteration i, where the reduction factor λ determines the weights:

$$\mathbf{x}_i = \frac{1}{1+\lambda}\hat{\mathbf{x}}_i + \frac{\lambda}{1+\lambda}\mathbf{x}_{i-1} \, . \tag{66}$$

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

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

5.2.5 Unphysical values

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

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

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

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

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

5.3 Numerical Implementation and Data Product

Based on SCIAMACHY heritage and numerical experience, we calculate the gain matrix $\mathbf{G}$ (see Eq. 65) using Cholevsky decomposition, which is roughly twice as efficient as the LU decomposition [RD56] with good numerical stability.

Furthermore, the numerical performance of the retrieval can be improved by optimising the vertical layering of the model grid used for the radiative transfer simulation. The layering determines the numerical speed of the simulation because it is directly linked to the dimension of matrix equation (8). To optimize the numerical efficiency of the two-stream solver, we first calculate absorption cross sections on a 1 km vertical grid to account for pressure and temperature dependence of atmospheric absorption in the SWIR spectral range. Because Rayleigh scattering is ignored in the radiative transfer simulation, all scattering takes place within the atmospheric scattering layer and between the scattering layer and the model Earth surface. Thus for the simulation, we can combine the atmosphere layers above the scattering layer and between this layer and the surface to one layer each by integrating the optical depth (see Fig. 5). This significantly reduces the number of vertical layers in the radiative transfer simulation (typically to less than 10), depending of the number of internal layers covered by the scattering layer. Although the radiative transfer simulations are performed on an aggregated vertical grid, the forward model Jacobian can be calculated on the original 1 km vertical grid taking into account the layer integration of optical depth. Based on this optimisation, the numerical performance of the retrieval will be discussed in Sec. 9.1.

5.4 Molecular optical properties

The molecular absorption cross sections of CH_4, H_2O, HDO and CO are pre-calculated from the HITRAN 2008 spectroscopic database [RD57] as a function of pressure, temperature and for a spectral sampling of $5 \cdot 10^{-3}$

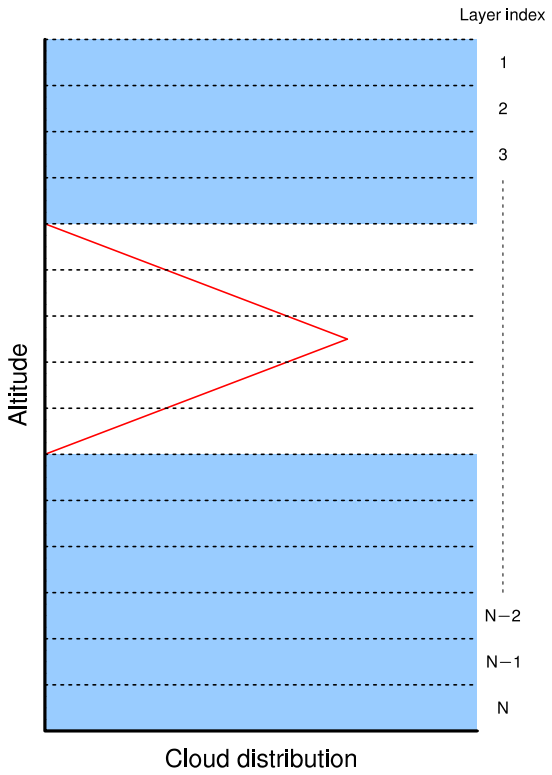

Figure 5: Optimisation of the vertical layering of the two-stream radiative transfer simulation based on the initial grid. The internal grid is needed to account for pressure and temperature dependence of atmospheric absorption. For the radiative transfer, layers above and below the scattering layer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

cm^{-1}. The cross sections are stored in lookup table files (NetCDF format). Here, the format of the lookup table files are identical for the CO and CH_4 retrieval algorithms to guarantee flexibility with respect to updates of the SWIR spectroscopy. For each individual retrieval with a corresponding pressure and temperature profile, cross sections are calculated by a bilinear interpolation on the pressure and temperature grid of the lookup table.

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

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

As a first definition of effective cross section we considered an average with a triangular weight function around the sample wavenumber k_i:

$$\sigma_i = \frac{\int T_i(k)\sigma(k)dk}{\int T_i(k)dk}\,, \tag{67}$$

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

$$\begin{aligned} T_i(k) &= k - k_{i-1} \quad \text{for } k_{i-1} < k < k_i \\ &= k_{i+1} - k \quad \text{for } k_i < k < k_{i+1} \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{68}$$

assuming an equidistant wavenumber grid (k_i) in the model. This approach, however, tends to overestimate the representative cross section systematically. The reason is that not the cross sections themselves should be

averaged, but their effects on the measured quantity, the reflectance. The reflectance is a convex (e^{-x} shape) function of the cross sections, and therefore, the reflectance for an average cross section is lower than the average reflectance for the cross sections. Therefore, conventional triangle convolution tends to overestimate the absorbing power of the molecules, leading to underestimations in the retrievals.

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

$$\sigma_i = \sqrt[m]{\frac{\int T_i(k)\sigma^m(k)dk}{\int T_i(k)dk}} \,. \tag{69}$$

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

5.5 Micro-physical properties of the scattering layer

The micro-physical properties of the scattering layer are of minor relevance for this study. This has been demonstrated by using the single-scattering albedo and the phase function for a highly absorption dust aerosol (type HA DUST 14 of Torres et al. [RD58]), although in all test scenarios measurements were calculated for water and ice clouds. Because of the minor relevance of the micro-physical properties of the scattering layers, these properties have been simplified to a single value for the single-scattering albedo and the asymmetry parameter and a simplified wavelength-dependence of the extinction optical thickness based on the Ångström parameter:

$$\tau(\lambda) = \tau(\lambda_0)\left(\frac{\lambda}{\lambda_0}\right)^{-\alpha} \,. \tag{70}$$

Here, λ_0 is a fixed reference wavelength and α is the Ångström parameter. The phase function for single scattering is approximated with the Henyey-Greenstein phase function, depending only on the asymmetry parameter. The effect of different scattering characteristics will be investigated in a later phase.

5.6 State vector, ancillary parameters and a priori knowledge

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

1. CO profile.
2. CH_4 profile.
3. H_2O profile.
4. HDO profile.
5. Geo-location of the SWIR measurement.
6. Temperature profile.
7. Surface elevation on SWIR pixel level.

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

Table 1: Setup of state vector $\mathbf{x}$. Here, CO, CH_4, H_2O and HDO indicates the total column retrieval of the trace gases, A_s and ΔA_s are the Lambertian surface albedo and its slope, z_{scat} represents the center height of the scattering layer, τ_{scat} the total optical thickness of the layer, and $\Delta\lambda$ is a spectral measurement offset of the measurement.

	Methane pre-fit	CO final retrieval
CO	x	x
CH_4	x	-
H_2O	x	x
HDO	x	x
A_s	x	x
ΔA_s	x	x
z_{scat}	-	x
τ_{scat}	-	x
$\Delta\lambda$	x	x

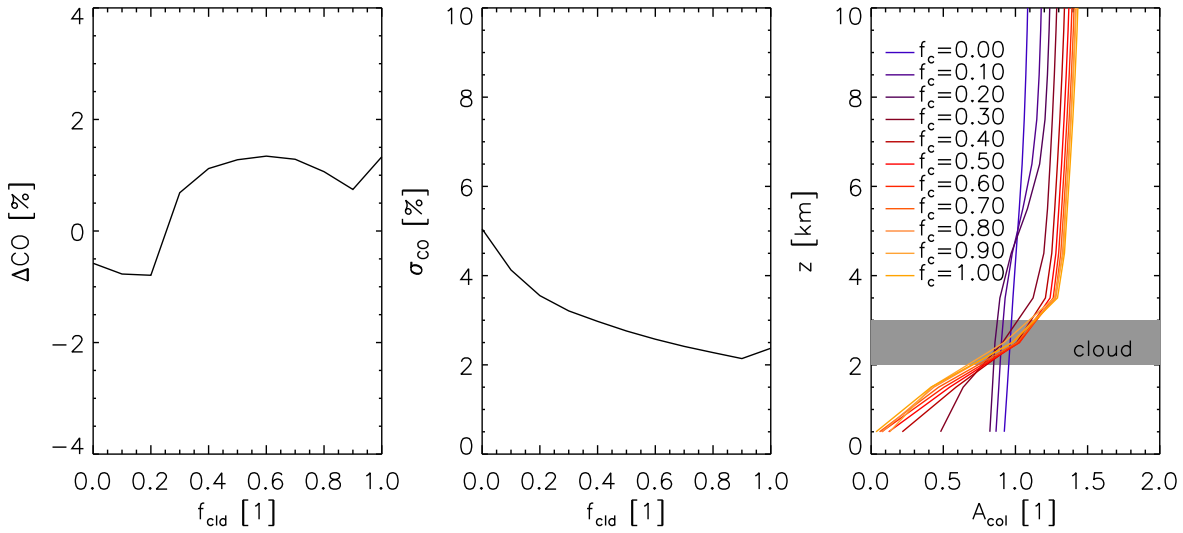

Figure 6: Example of the CO data product. The SWIR measurements are simulated for a scene partially covered by a water cloud between 2 and 3 km with optical depth $\tau_{\text{cld}} = 30$ and a surface albedo $A_s = 0.05$. Left panel: Difference Δ_{CO} between the true CO column and the retrieved CO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

5.7 Data product

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

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

Figure 6 shows an example of the CO retrieval product over land for a cloudy atmosphere and a surface albedo $A_s = 0.05$. It shows the difference between the true CO column and retrieved CO column corrected for retrieval noise (left panel), the 1-σ estimate of the corresponding retrieval noise (middle panel) and the column averaging kernel (right panel). The retrieval scales the true relative CO profile. The left panel shows only small differences ($\leq \pm 1$ %) between the true and the retrieved profile for cloud fractions $f_{\rm cld}$ = 0–1. The differences are due to the description of atmospheric scattering in the retrieval. The shielding of the atmosphere below the cloud does not reduce the retrieved CO column as a function of increased cloud fraction because the sensitivity of the measurement with respect to the CO abundance above the cloud is used to retrieve the total CO column by an appropriate scaling of the relative profile. At the same time, the retrieval noise on the CO column decreases due to the gain in SNR for increasing cloud coverage. The change of retrieval sensitivity with cloud coverage is clearly illustrated by the column averaging kernels shown in the right panel of Fig. 6. Here, the color of lines indicate the cloud fraction. When the cloud fraction is greater than zero, the column averaging kernel starts to increase above the top of the retrieved cloud height and at the same time the retrieval sensitivity decreases below the cloud such that the net effect on the retrieval CO column nearly cancels out.

6 Common aspects with other algorithms

Both, the SICOR algorithm and the operational TROPOMI CH_4 RemoTeC algorithm [RD2] will be implemented in one common processing structure as discussed in Sec. 8.1. This makes it possible to optimize synergistic use of data structures as e.g. a common lookup table for atmospheric absorption cross sections and input meteo fields.

Table 2: Microphysical properties of water and ice clouds: $n(r)$ represents the size distribution type, $r_{\rm eff}$ and $v_{\rm eff}$ are the effective radius and variance of the size distribution, $n = n_r - in_i$ is the refractive index.The ice cloud size distribution follows a power-law distribution as proposed by [RD1].

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

7 Error analysis

To assess the performance of the CO retrieval algorithm, we have generated simulated measurements for various test cases using the S-LINTRAN version 2.0 radiative transfer model [RD59]. S-Lintran is a scalar plane-parallel radiative transfer model based on the discrete ordinate method [RD60]. It fully accounts for multiple elastic scattering by particles (clouds and cirrus) and molecules, and the interaction with the reflecting Earth surface. This version of LINTRAN does not account for the polarisation properties of light, which are of minor relevance for the purpose of this study. The optical properties of water clouds are calculated using Mie theory with microphysical cloud properties given in Table 2. For ice clouds the ray tracing model of Hess et al. [RD61, RD62] is employed assuming hexagonal, columnar ice crystals randomly oriented in space. Finally, cirrus and water clouds are described by a cloud top and base height, and a cloud optical thickness. Furthermore, we assume that cirrus fully cover the observed ground scene, whereas for water clouds, partially cloud coverage is possible utilising the independent pixel approximation [RD63] for the simulation.

The radiance spectra are superimposed by measurement noise using the TROPOMI noise model of [RD64]. For this purpose, we assumed an observed ground scene of 7×7 km^2 and a telescope aperture of 6×10^{-6} m^2. The optical transmittance of the instrument is adjusted such that, for a spectral sampling of 0.1 nm, a signal to noise ratio of 120 is achieved in the continuum of the spectrum for a dark reference scene (surface albedo $A_{\rm s} = 0.05$, viewing zenith angle $\mathrm{VZA} = 0^\circ$ and $\mathrm{SZA} = 70^\circ$). The instrument noise is estimated to be 346 electrons per signal. It includes noise due to thermal background, dark current of the detector, readout noise and analog-to-digital converter noise.

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

$$\Delta\mathrm{CO} = C_{\rm CO} - \mathbf{G}_{\rm CO}\mathbf{e}_y - \mathbf{A}_{\rm CO}\,\rho_{\rm CO,true}\ , \tag{71}$$

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

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

7.1 Performance analysis for generic scenarios

For all generic cases, which we describe in more detail below, we assume the same atmospheric model. We assume the US standard atmosphere [RD65] for the profiles of dry air density, pressure, water and CO. The CH_4 profile is taken from the European background profile of the TM4 + CHIMERE chemical transport model [RD66], interpolated to the same pressure grid and converted from mixing ratios to densities using the air densities from the US standard atmosphere. We separated the water profile into individual profiles for the three isotopic components with absorption features in the TROPOMI SWIR range: $H_2{}^{16}O$, $H_2{}^{18}O$ and HDO.

In Fig. 7, we show the (normalised) atmospheric H_2O, CO, CH_4 and air density profiles, the temperature profile and the profiles for the HDO and $H_2{}^{18}O$ depletion. For all cases A-F, the measurement spectra are

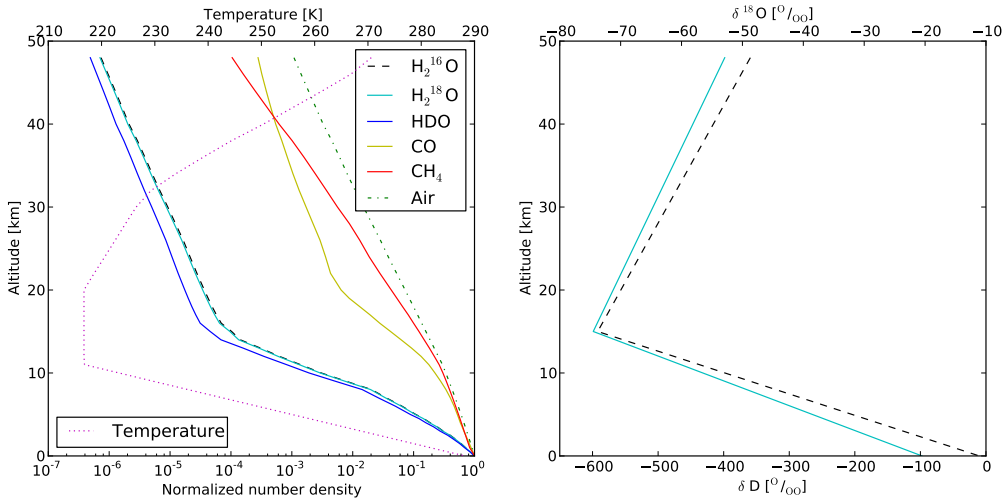

Figure 7: Left: atmospheric concentration profiles (bottom axis) and temperature profile (top axis) used as input for the model atmosphere. The concentrations are normalised to the concentration at ground level. Right: assumed profiles for the amount of HDO depletion (solid line, lower axis) and $H_2{}^{18}O$ depletion (dashed line, top axis).

Table 3: Summary of the different generic test cases A-F.

Case	Target of study	Variable X	Variable Y	Remarks
A	cloud free	surface albedo	SZA	
B	clouds	cloud top height	cloud fraction	for $\tau_{\mathrm{cld}} = 5, 10, 30, 50$
C	photon trapping	surface albedo	cloud fraction	for τ_{cld} = 2 & 5 and cloud top height = 2 & 5 km
D	aerosol	surface albedo	AOT	for a 0–2 km and 4–5 km aerosol layer
E	cirrus	surface albedo	cirrus optical depth COT	
F	multiple cloud layers	cloud fraction	COT	for A_S = 0.05, 0.1, 0.3, 0.6

simulated for a water vapor column of $4.79 \cdot 10^{22}$ molec./cm^2, a CH_4 column of $3.68 \cdot 10^{19}$ molec./cm^2, and a CO column of $2.38 \cdot 10^{18}$ molec./cm^2. In addition to this model atmosphere, we defined different cases to test the influence of clouds, aerosol, surface reflection and solar geometry. All simulations assumed a fixed sensor viewing zenith angle of 40° and an azimuthal difference between the sensor and the sun of 60°. These test cases are described below and an overview of the test cases is given in Table 3.

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

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

Case C allows us to study the effect of "photon trapping". We have varied the surface albedo between 0.03–0.6

(in 10 steps) together with the cloud fraction (between 0.0–1.0 in 20 steps) for four different water clouds: a cloud optical thickness $\tau_{\rm cld}$ of 2 and 5, combined with a cloud top height of 2 and 5 km. The geometrical cloud thickness was 1 km and the SZA was fixed at 50°.

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

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

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

7.2 Performance Analysis for generic scenarios

To test the iterative approach, the retrieval is initialised with an atmospheric state, which differs from the truth i.e. a water vapor and CO column of 4.00 10^{22} molec./cm^2 and 2.00 10^{18} molec./cm^2, respectively, a surface albedo given by the maximum Lambert-equivalent reflectivity (LER) value of the measurement defined by

$$LER(\lambda) = \frac{I^{\rm TOA}(\lambda)\pi}{\mu_o F_o(\lambda)}\,, \tag{72}$$

and a triangular scattering layer centered at 5 kilometers altitude, with a fixed geometric thickness of 2.5 km and an optical depth of 1. Figure 8 shows the CO retrieval bias ΔCO defined in Eq. (71) and the corresponding retrieval noise for case A. Overall, the retrieval bias is small with -2 % $\leq \Delta{\rm CO} \leq$ 2 %. The retrieval noise increases from values $\leq$ 1 % at high sun and high surface albedo to ≈ 11 % for low sun (SZA = 70^o) and low albedo ($A_s = 0.03$). This increase is governed by the signal strength and so by the SNR of the measurement. Both retrieval bias and retrieval noise are within the level-2 requirement specification for CO (see Sec. 4.2).

The case B retrieval performance for water clouds is shown in Fig. 9. Here, the grey areas indicate observations which are rejected by the methane cloud filter using a threshold value of Δ CH_4 of 25 %. The data filtering depends on cloud optical depth and thus on the transmission of the cloud layer. For example, measurements with a cloud top height at 3 km are rejected by the cloud filter for a cloud coverage of $f_{\rm cld} \geq 0.6$ for $\tau_{\rm cld} = 5$, of $f_{\rm cld} \geq 0.4$ for $\tau_{\rm cld} = 10$, and of $f_{\rm cld} \geq 0.3$ for $\tau_{\rm cld} = 30$ and 50. For observations accepted by the cloud filter, the retrieval bias is within $\pm 1\%$. This illustrates the capability of the retrieval to account for water clouds above dark surfaces. For all successful retrievals of case B, the retrieval noise (not shown) is better than the corresponding clear sky value because of the enhanced signal due to the presence of clouds.

In the previous case B, the effect of light path enhancement due to photon trapping, i.e. multiple reflection between the Earth surface the cloud is minor because of the low surface reflection. To estimate the algorithm performance in case of photon trapping, we show in Fig. 10 the Case C CO retrieval bias for optically thin water clouds at two different altitude levels with cloud top height $z_{\rm cld} = 2$ and 5 km as a function of surface albedo and cloud fraction. In this case, measurements are only rejected for the cloud at 5 km altitude for increasing cloud fraction and low surface albedo. The cloud filtering is stricter for the optically thicker cloud ($\tau_{\rm cld} = 5$), because the shielding of the atmosphere below the cloud becomes more efficient with increasing cloud optical thickness. Photon trapping is most efficient for optically thin clouds above bright surfaces and can compensate the light path shortening due to the reflection of light at the cloud top. Overall, in case of convergence the CO bias is between ± 2 % and only for a few case (6 out of 336 measurement simulations) the retrieval did not converge. Thus, due to the retrieval of an scattering layer the effect of photon trapping can be account for by the retrieval.

Multiple scattering by atmospheric aerosol also affects the light path of the observed light. To investigate this effect, we consider the retrieval performance for case D. Figure 11 shows the CO retrieval bias as a function

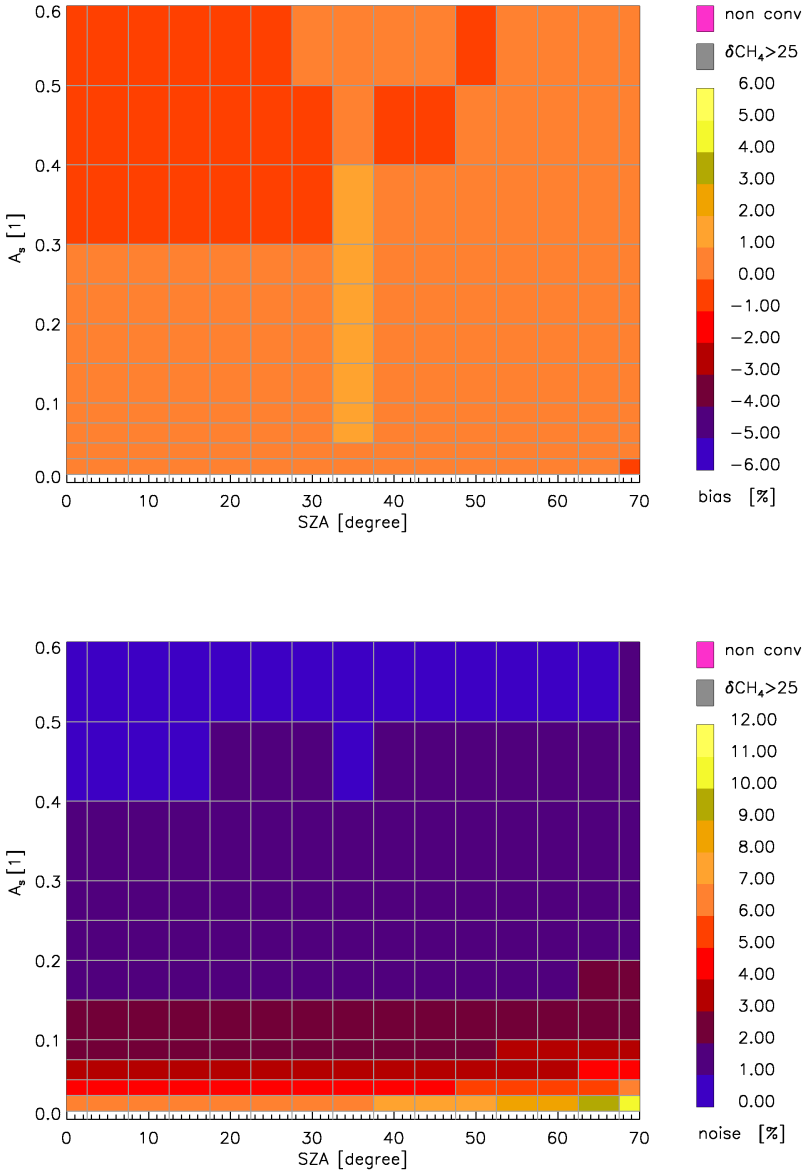

Figure 8: Retrieval bias $\Delta\mathrm{CO}$ (upper panel) and retrieval noise (lower panel) for the clear sky test case A as a function of SZA and surface albedo A_s.

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

Next, we consider Case E measurement simulations for an atmosphere with a cirrus layer between 9 and 10 km and varying optical depth. Similar to Fig. 10, the CH_4 pre-fit rejects measurements for larger optical depth and low surface albedo. The effect of photon trapping at high surface albedo and small cirrus optical depth results in cirrus biases of less than 1 % (see Fig. 12). So, this example indicates that the presence of optically thin cirrus clouds can be handled by the retrieval algorithm.

As a last test ensemble, Case F, Fig. 13 shows the CO bias for multiple cloud layers, i.e. a cirrus between 9 and 10 km and a water cloud between 2 and 3 km. For the low albedo scenes most of the data are rejected by the cloud filter, where for high albedo most measurements pass the cloud filter. Here, multiple scattering

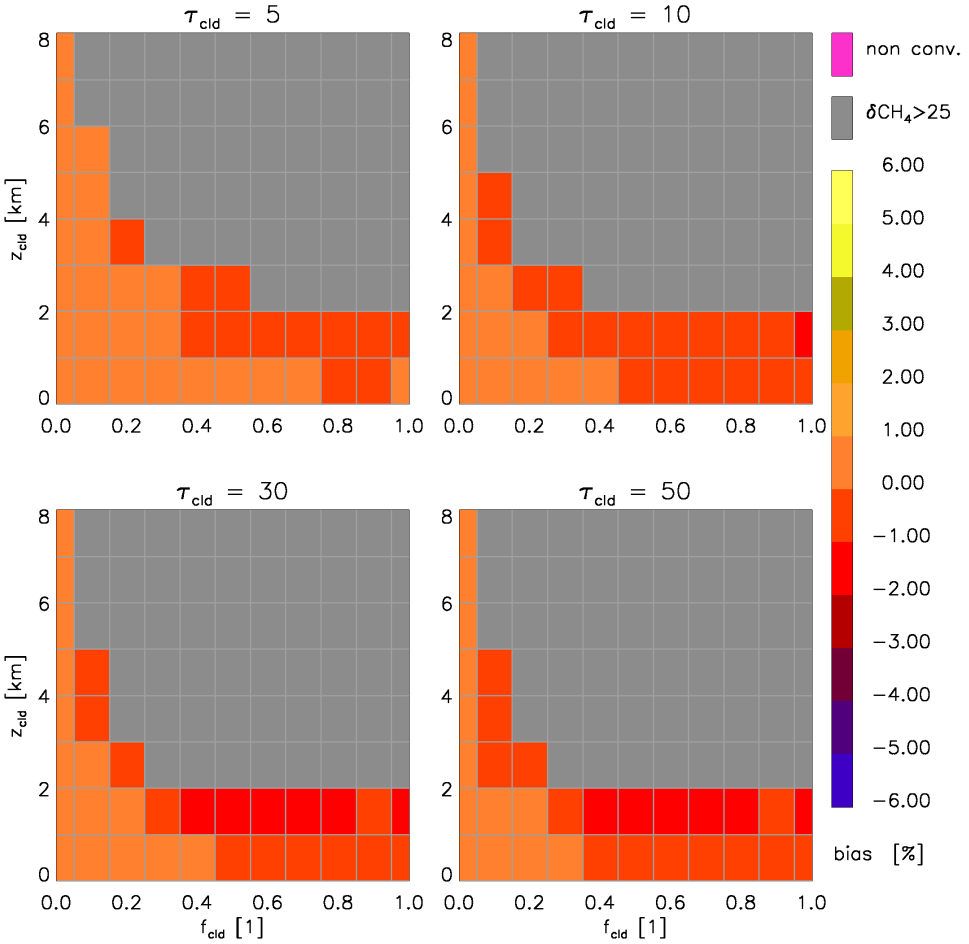

Figure 9: Retrieval bias for case B, i.e. for water clouds above a dark surface ($A_s = 0.05$) as a function of cloud top height z_{cld} and cloud fraction f_{cld} for different cloud total optical depths $\tau_{\mathrm{cld}} = 5$ (upper left panel), $\tau_{\mathrm{cld}} = 10$ (upper right panel), $\tau_{\mathrm{cld}} = 30$ (lower left panel), $\tau_{\mathrm{cld}} = 50$ (lower right panel).

between the surface, the water cloud and the cirrus takes place. The retrieval bias is in all cases below 1 %. This demonstrates nicely the strength of the algorithm to deal with multiple scattering of sunlight even for multiple scattering layers.

In summary, the retrieval performance for the measurement ensembles of case A to F are compliant with the CO level-2 requirements as described in Sec. 4.2. The bias analysis showed that the effect of water clouds, which partially cover the observed ground scene, of optically thin cirrus, of atmospheric aerosol loading and of multiple cloud layers can be treated with sufficient accuracy by a triangular scattering layer implemented in a simplified two-stream radiative transfer model.

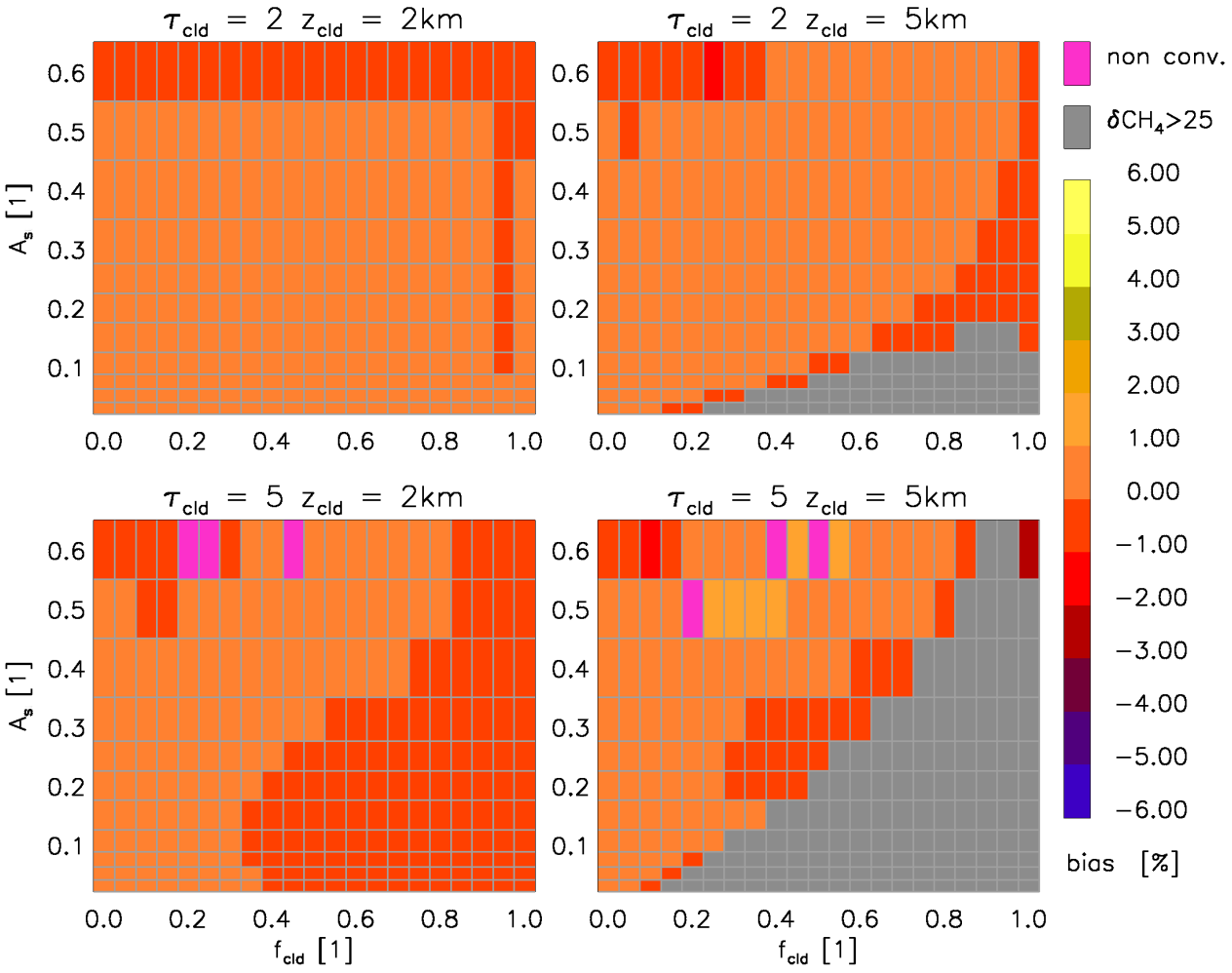

Figure 10: Retrieval bias in case of photon trapping between a water cloud and the bright surface (case C). The CO bias is shown as a function of surface albedo A_s and cloud fraction f_{cld} for a cloud with optical depth $\tau_{\mathrm{cld}} = 2$ and cloud top height $z_{\mathrm{cld}} = 2$ km (upper left), for $\tau_{\mathrm{cld}} = 2$ and $z_{\mathrm{cld}} = 5$ km (upper right), for $\tau_{\mathrm{cld}} = 5$ and $z_{\mathrm{cld}} = 2$ km (lower left) and for $\tau_{\mathrm{cld}} = 5$ and $z_{\mathrm{cld}} = 5$ km (lower right).

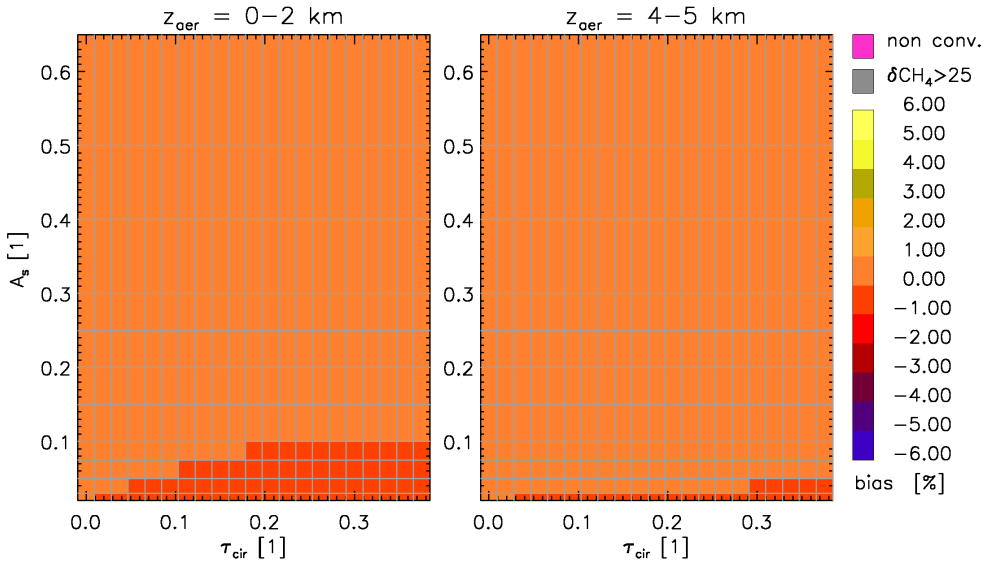

Figure 11: Retrieval bias for an aerosol loaded atmosphere of test case D. In the left panel, the CO retrieval bias is shown as a function of surface albedo and aerosol optical depth for a sulfate aerosol between the surface and 2 km altitude. The right panel shows the corresponding error analysis for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

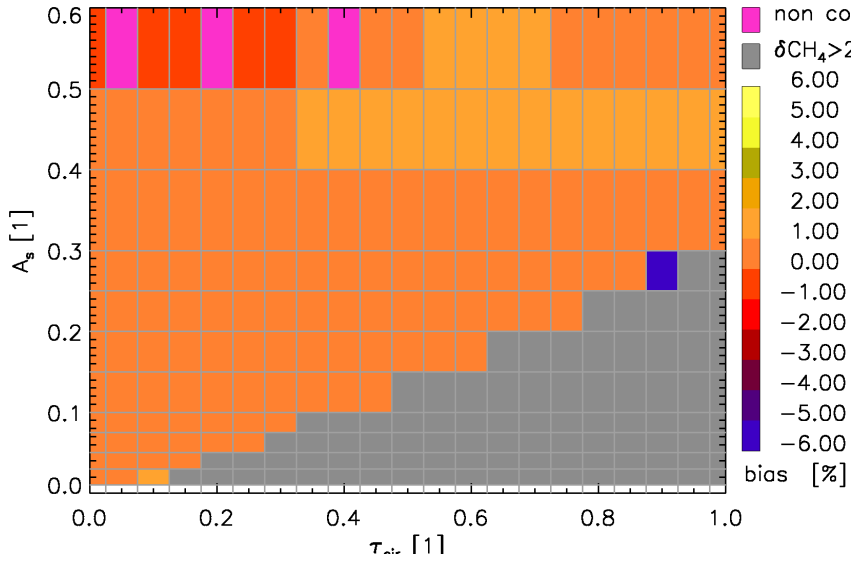

Figure 12: CO retrieval bias for measurements in presence of optically thin cirrus clouds as a function of surface albedo and cirrus optical depth (case E). The cirrus optical depth is given at 2300 nm.

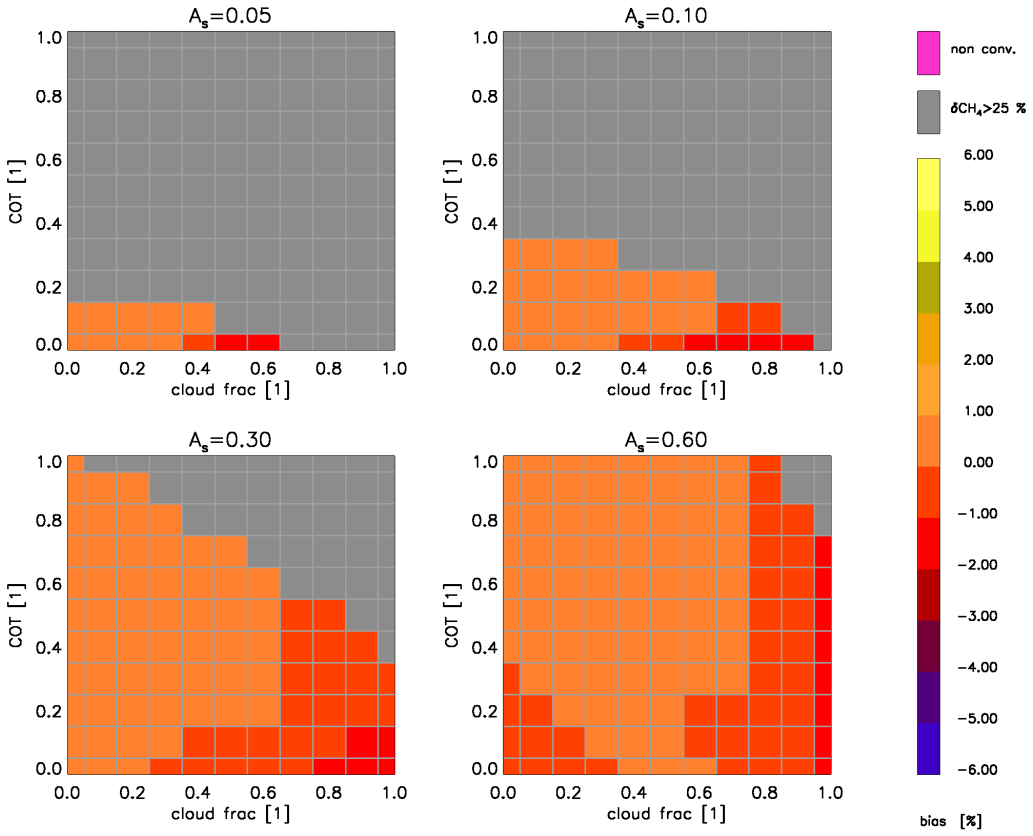

Figure 13: CO retrieval bias for measurements in presence of multiple cloud layers as a function of cirrus optical depth and cloud fraction of a water cloud for different surface albedo (case F). The cirrus optical depth is given at 2300 nm.

7.3 Geophysical test scenario for measurements over China

To test the algorithm performance for more realistic circumstances, we have simulated a measurements ensemble for China. Here, the measurement geometry and the atmospheric state of the observed scene are taken from several data sources. The viewing and solar geometry as well as the ground pixel size is adapted from a MODIS AQUA granule over China. The MODIS information is resampled on a pixel size of 10×10 km^2 at sub-satellite point, which is similar to the TROPOMI pixel size of 7×7 km^2. The pixel distortion towards the outer swath is adopted from the MODIS observation geometry. The surface albedo is taken from the MODIS MCD43C4 product for the Nadir Reflectance Band 7 at 2105–2155 nm and ground elevation is inferred from the ETOPO5 database of land and sea-floor elevations on a 5-minute latitude/longitude grid. Both surface albedo and ground elevation are interpolated to the MODIS granule grid using bilinear interpolation. The MODIS MYD06 data product with a horizontal resolution of 5×5 km^2 is used for cloud information, where we evaluate cloud top height from the MODIS cloud top pressure. Cirrus optical thicknesses are derived from the MODIS cirrus reflectance product and viewing geometry. Here, we use the algorithm by Desser et al. and Yang et al. [RD67, RD68]. The cirrus is located at 9–10 km all over the scene. One has to be aware that the cirrus data are rather uncertain due to various assumptions made in the cirrus algorithm, but also due to the MODIS data itself. Moreover, we added the MODIS aerosol product MOD04 at a spatial resolution of 10×10 km^2 at sub-satellite point to an overall rural background aerosol of an total optical depth of 0.1 at 2300 nm. ECMWF data on pressure, temperature, water vapor and air density stratification with a temporal resolution of 6 hours are interpolated to the instance of the MODIS measurement. Finally, the CHIMERE [RD69] and TM4 chemistry models [RD70] provide CO and CH_4 profiles, which are subsequently interpolated to the MODIS granule grid and to the ECMWF local pressure scales. Here, the CHIMERE chemical transport model provides information in the lower troposphere, and the TM4 fields are used to extend the CHIMERE fields to the upper troposphere and stratosphere. Examples of the ensemble fields are given in Fig. 14.

With this input and using the S-Lintran model as described above, measurements are simulated which subsequently are used as input to the SICOR retrieval algorithm. The corresponding SICOR CO bias $\Delta\mathrm{CO}$ is shown in Figure 15. About 28 % of the simulated observations are rejected by the CH_4 cloud filter causing the data gaps in the figure. For about 2 % of all converged cases, scattering had to be omitted in the inversion to avoid unphysical input to the radiative transfer model (see discussion in Sec. 5.2.5). Over the South China Sea, the retrieval converges in the presence of low water clouds. Other measurements in this region are rejected because of the methane filter or because of a too low signal. Furthermore, in the Northwest of China the CO biases exceed 5 % because of the occurrence of cirrus above a bright surface. Biases of $\leq$ 5 % occur around the city Wuhan due to the presence of tropospheric aerosols.

Figure 16 shows the bias distribution as a function of the methane filter $\Delta\mathrm{CH}_4$. Here, the CO bias distribution is not centered around zero but has a mean bias of about 1.6 % and a median bias of about 1.4 %. This overall offset can be attributed to the triangular convolution of the molecular absorption cross sections as described in Sec. (5.4). The CO bias distribution has a narrow standard deviation of 2.3 %. It is asymmetric and skewed towards positive values. However, the figure shows no significant dependence of the bias on the CH_4 pre-fit bias with a correlation coefficient $r = 0.01$. This confirms the stable performance of the retrieval for different cloud conditions.

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

Finally, Fig. 18 shows the distribution of the CO retrieval noise as a function of the LER in the continuum of the spectrum. As expected, the distribution shows a clear correlation of CO bias with the LER and the data are clustered around two maxima at LER-values of 0.1 and 0.4. The form of the CO bias distribution has its reason in the specific water cloud and cirrus distribution of our ensemble. The mean CO retrieval noise is 2.6 % and an insignificant amount of data points exceed the 6 % level. This low CO noise level may have its origin in the use of the MODIS 2.1 μm albedo to represent the surface reflection at the 2.3 μm band. The distribution of the LER values has its maximum at 0.1 and nearly no data points below 0.05. SCIAMACHY observations have indicated that the MODIS 2.1 μm albedo slightly overestimates the corresponding albedo at the 2.3 μm band by 0.03-0.06, which may cause a too low CO retrieval noise estimate for the China ensemble. However, the China test ensemble confirms the overall good retrieval performance which was already indicated by the generic test cases.

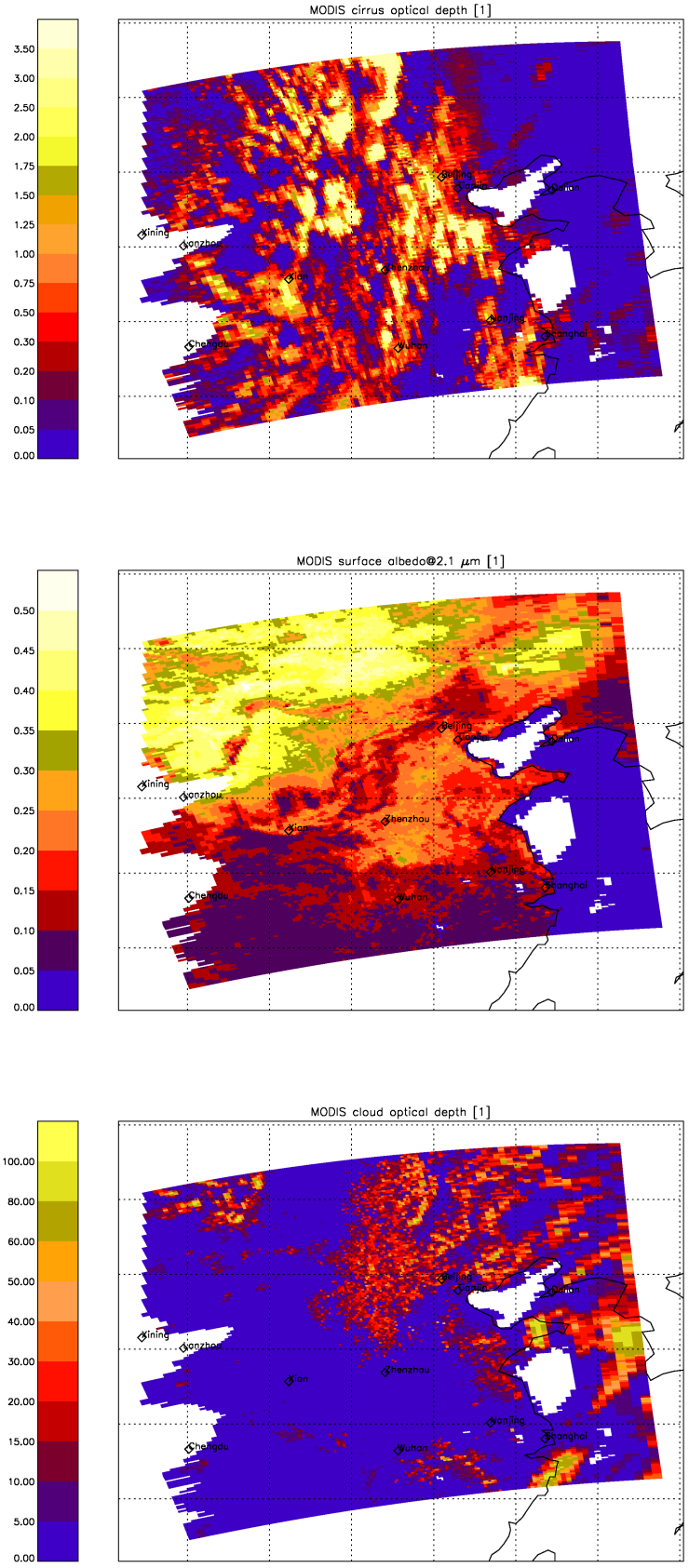

Figure 14: Test ensemble to generate TROPOMI measurements over China for a 10×10 km^2 pixel size for 10 May, 2006. The pixel distortion towards the edge of the swath is adopted from MODIS: (upper panel) MODIS cirrus optical depth, (middle) MODIS ground albedo at 2.1 μm, (lower) MODIS water cloud optical depth.

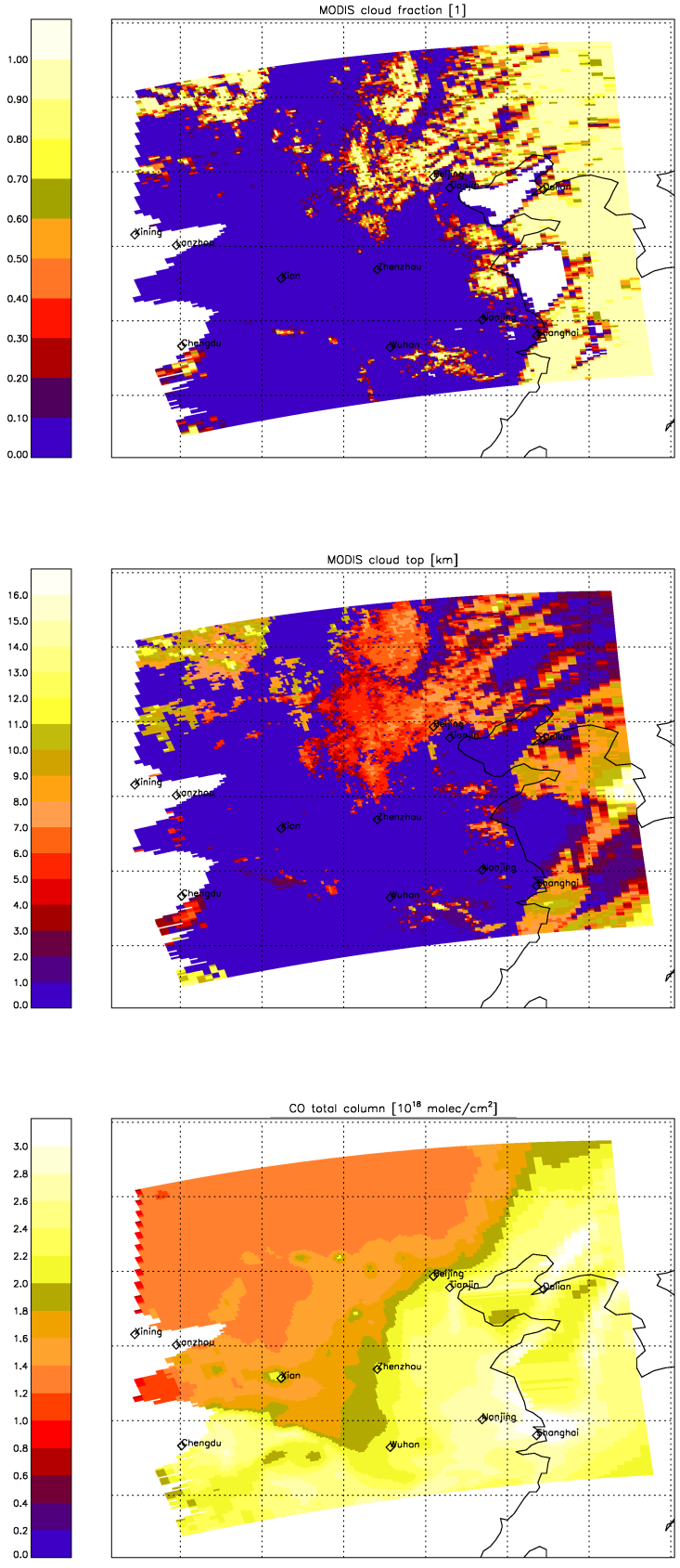

Figure 14: (Continued) (upper) MODIS cloud fraction, (middle) MODIS cloud top height, (lower) CHIMERE and TM4 CO total column.

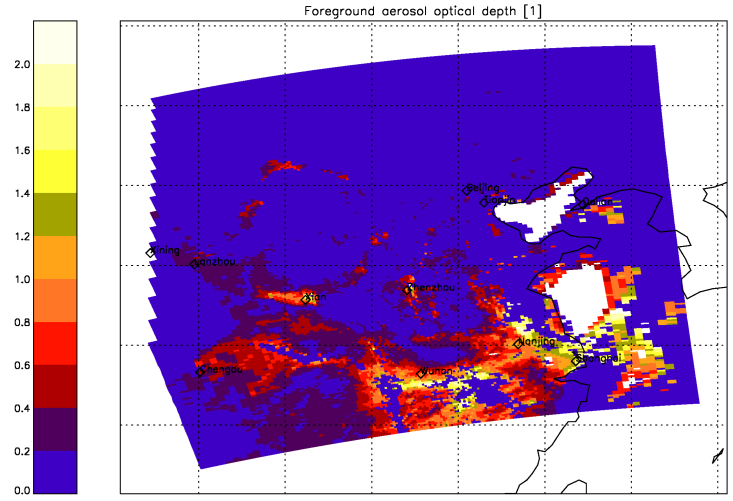

Figure 14: (Continued) MODIS aerosol optical depth at 2300 nm.

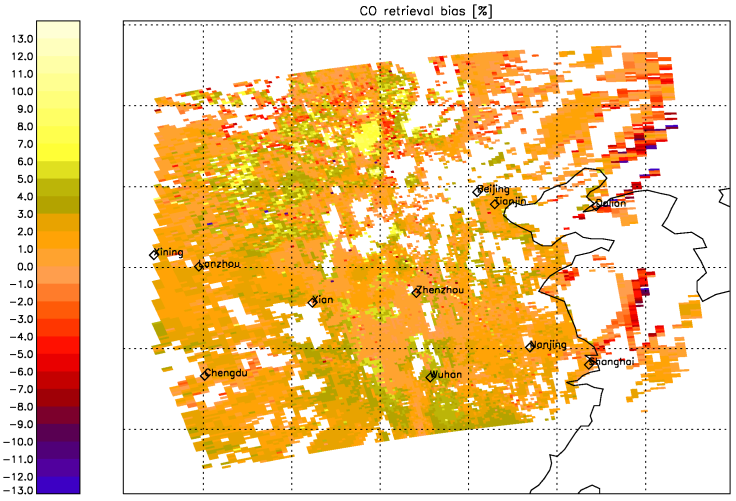

Figure 15: Bias $\Delta\mathrm{CO}$ of the retrieved CO column for the test ensemble shown in Fig. 14.

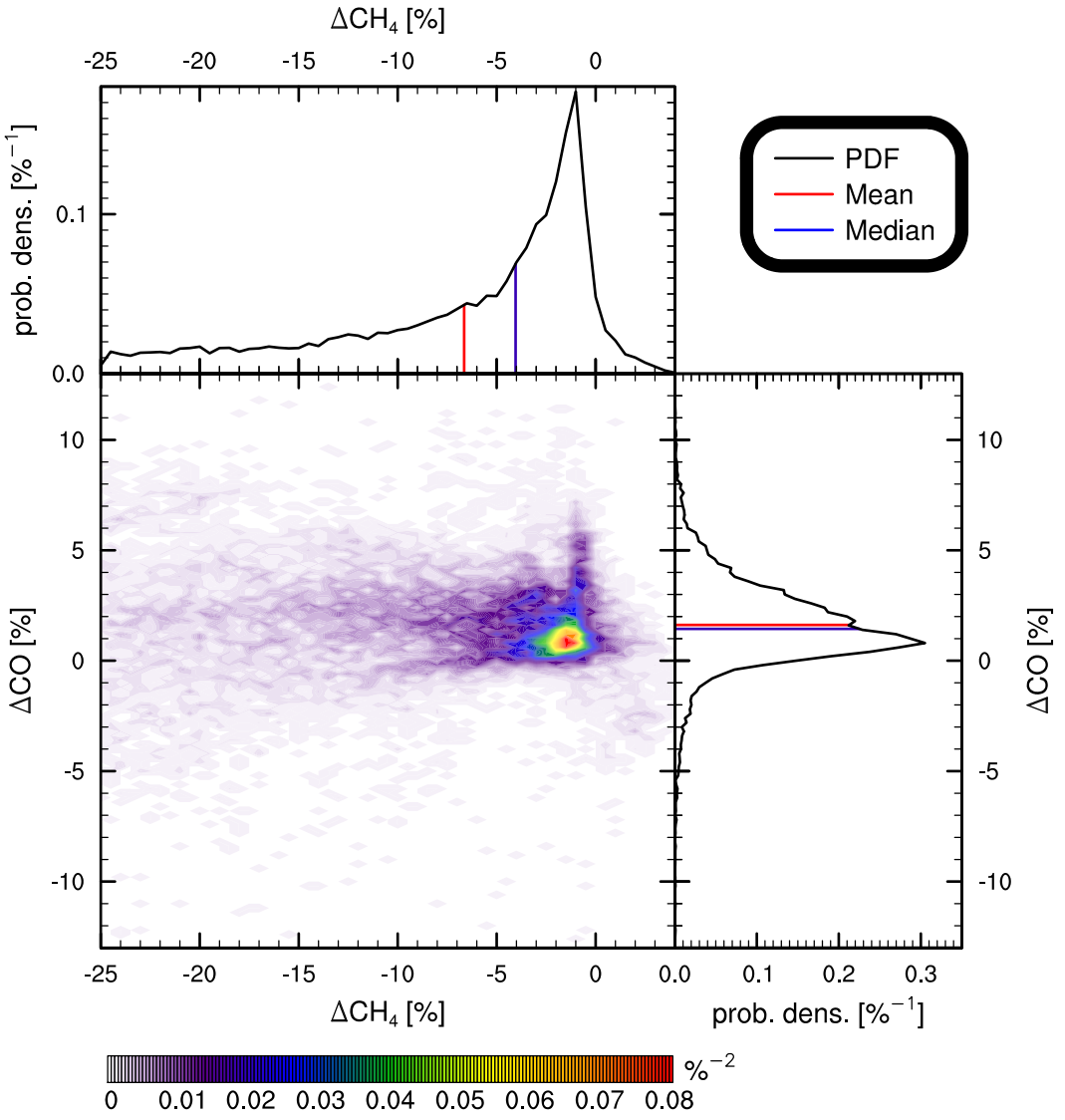

Figure 16: Bias of the retrieved CO column as a function of methane filter ΔCH_4 for the China test ensemble shown in Fig. 14. The median of the methane pre-fit bias is -4.0 %, the mean bias is -6.6 % with a standard deviation of 6.7 %. Accordingly, the median of the CO retrieval bias is $+1.4$ %, the mean bias is $+1.6$ %, with a standard deviation of 2.3 %. The correlation coefficient (r) between ΔCO and δCH_4 is $+0.01$.

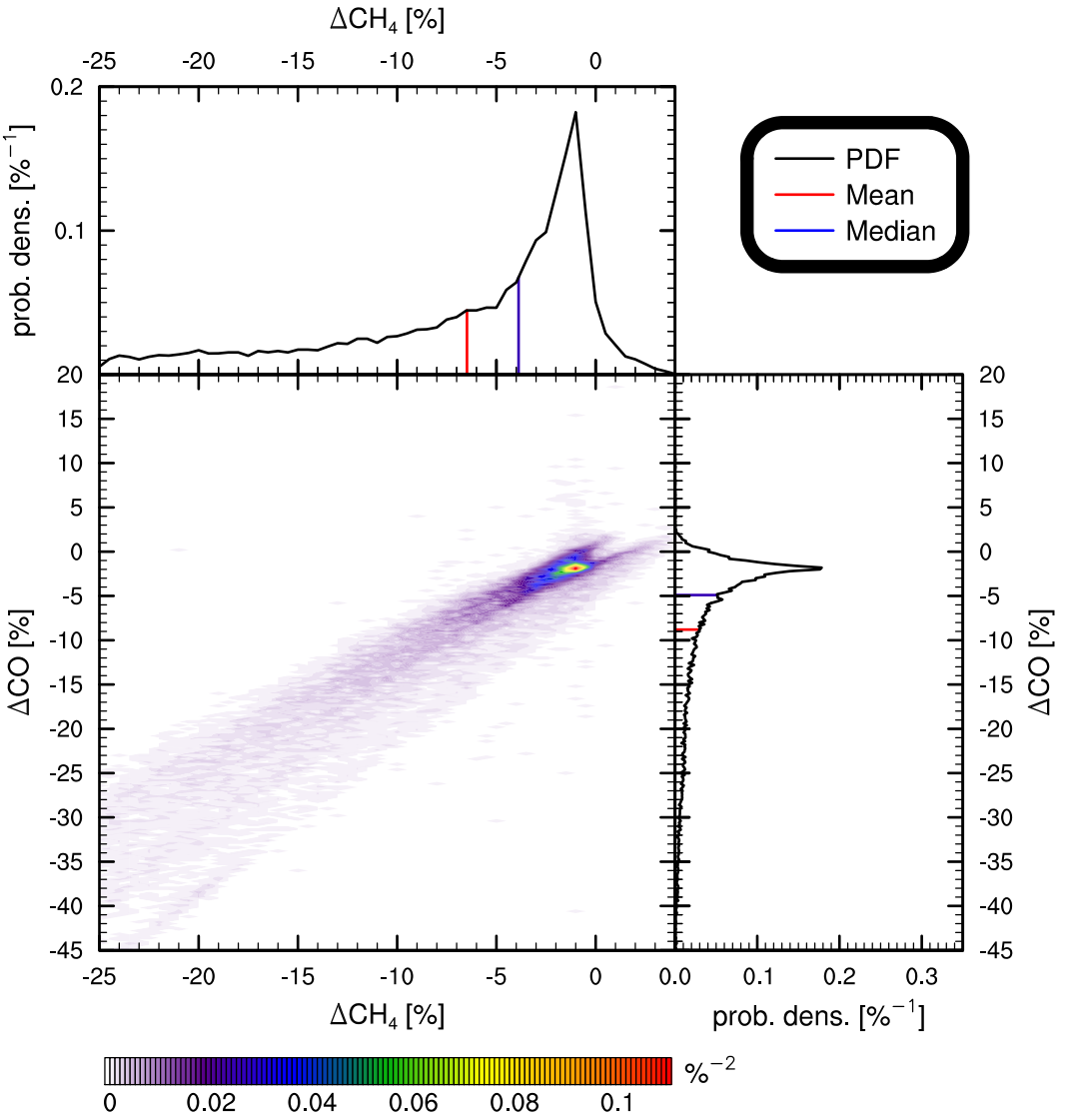

Figure 17: Same as Fig. 16 but for a non-scattering CO retrieval. The median of the CO retrieval bias is -4.90 %, the mean bias is -8.80 %. The correlation coefficient (r) between ΔCO and ΔCH_4 is 0.95.

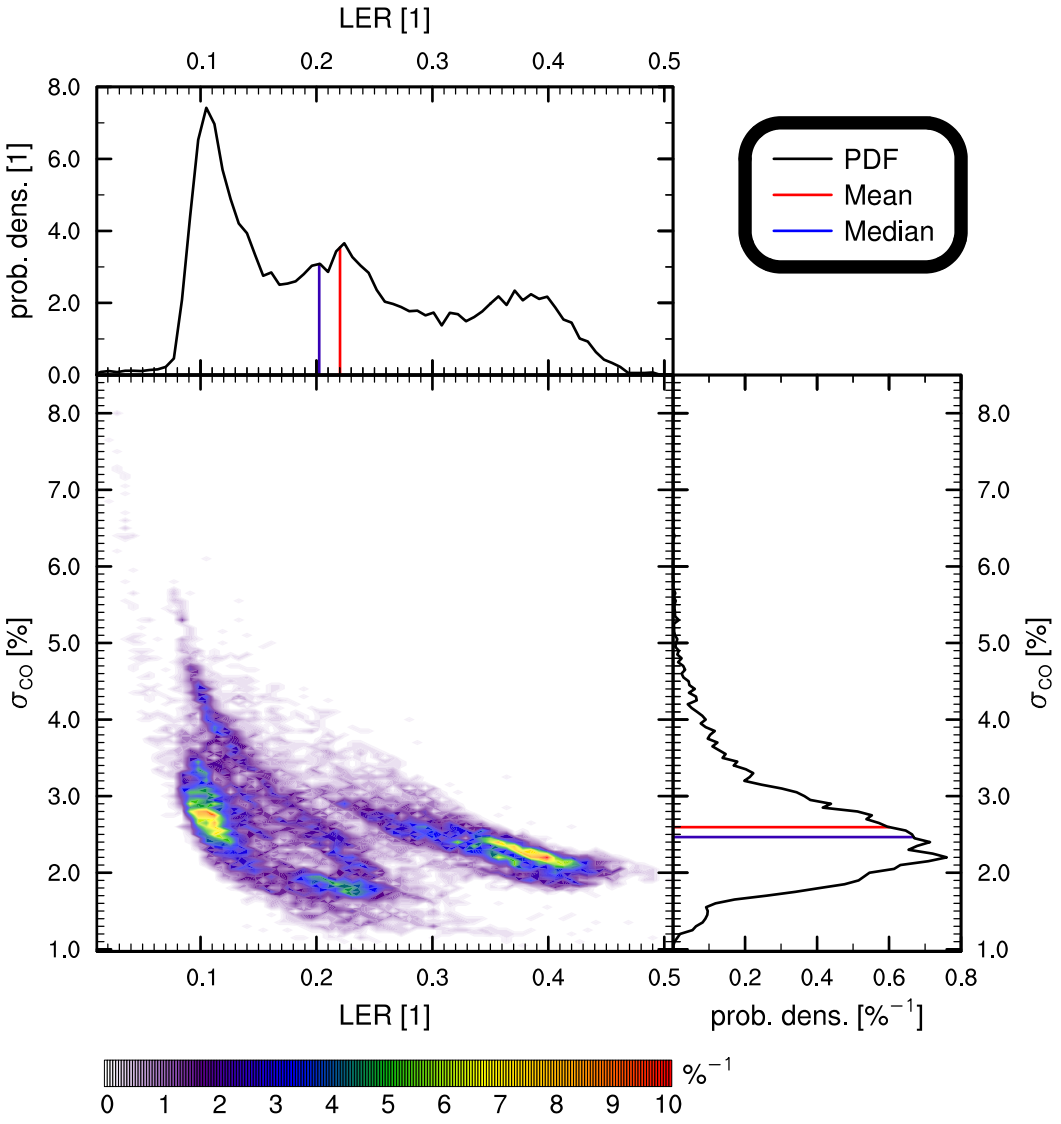

Figure 18: CO column retrieval noise as a function of methane filter ΔCH_4 for the China test ensemble shown in Fig. 14. The LER distribution as a median of 0.2, a mean value of 0.2 and a standard deviation of 0.1, whereas the CO retrieval noise distribution has a median of 2.5 % a mean value of 2.6 %, and a standard deviation of 0.8 %. The correlation coefficient (r) between CO noise and LER is -0.5.

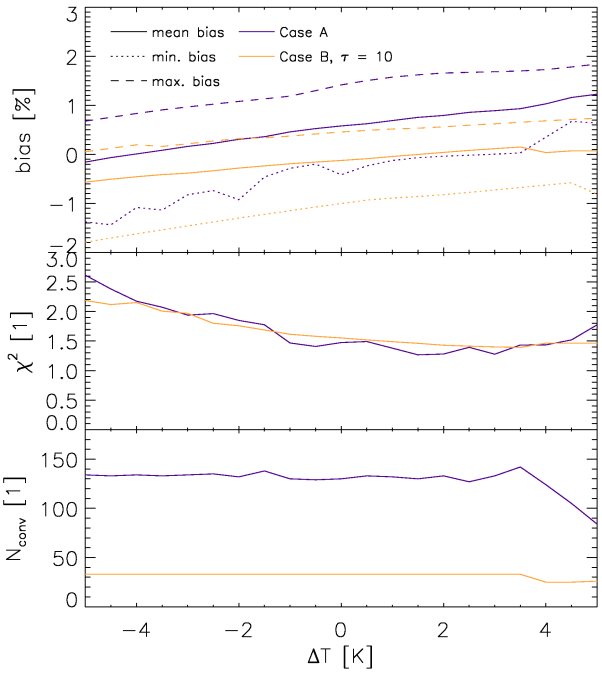

Figure 19: Upper panel: CO retrieval bias as a function of a shift ΔT of the atmospheric temperature profile. Measurement ensembles of the clear sky generic test cases A (purple lines) and the cloudy sky test case B for a cloud with an optical depth of 10 (orange lines) are considered. Here, the maximum (dashed lines), mean (solid lines) and minimum (dotted lines) bias is reported for the different temperature shifts. Middle panel: mean spectral χ^2 of the retrieval as a function of temperature shift. Lower panel: number of converged retrievals as a function of temperature shift.

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

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

- A shift of the assumed temperature profile (see Fig. 19.)
- Errors in the ground pressure and therefore a scaling in the pressure profile (see Fig. 20.)
- Wrong a priori knowledge on the total amount of methane (see Fig. 21.)
- Uncertainties in the relative profile of CH_4 and H_2O (see Figs. 22-24.)

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

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

Next, we consider the estimated surface pressure which must be representative for the observed ground scene. Errors in the surface pressure result in relative errors of the pressure profile. Similar to temperature uncertainty, this input error affects the retrieval due to wrong molecular absorption cross sections. Figure 20 shows the CO retrieval bias due to this error source for a scaling of the pressure profile between $\pm 3\%$. Within a scaling of $\pm 1\%$ the scaling affects only little the CO performance with a CO bias trend of 1.15 % for 1 % scaling

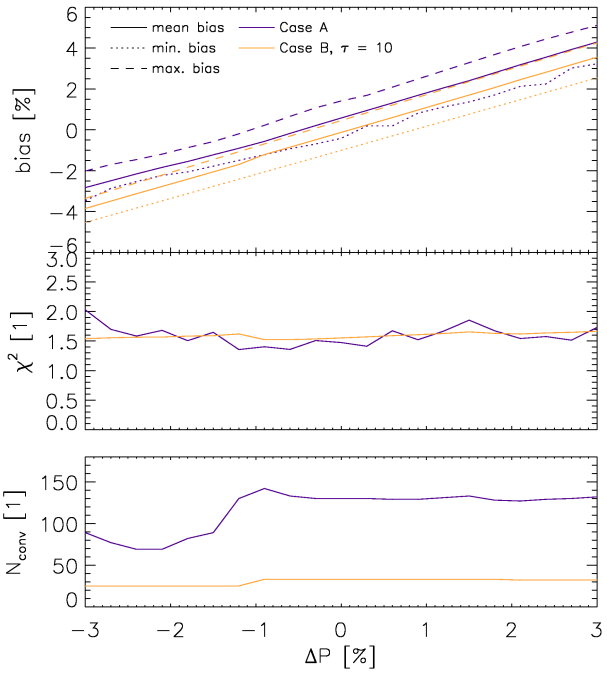

Figure 20: Same as Fig. 19 but as a function of surface pressure error ΔP .

of the pressure profile. For a scaling < -1 %, the number of successful retrievals is significantly reduced for the clear sky scenarios indicating a critical loss of retrieval performance. However, a priori knowledge of the surface pressure within 1 % is reasonable and so we expect no critical performance issues due to uncertainties in the ground pressure.

Due to our retrieval concept, which uses the atmospheric CH_4 absorptions to retrieve effective scattering layer height and optical depth, the sensitivity of the retrieval with respect to uncertainties in our CH_4 a priori knowledge is of particular relevance. Figure 21 shows a linear one-to-one relationship between the CO bias and the CH_4 a priori uncertainty. So, to keep the CO bias due to this error source within ±3% the vertically integrated amount of CH_4 must be known within 3 % as well. This puts a stringent requirement on the methane a priori knowledge accuracy. In Sec. 7.6, we discuss the current accuracy of CH_4 model prediction.

Finally, we consider the CO retrieval sensitivity to errors in the assumed relative profile of relevant atmospheric absorbers. For CO, the difference between the relative CO profile and its truth causes a null-space contribution of the retrieval which is characterised by the column averaging kernel (see discussion in Sec. 5.2.2 and [RD54]). Hence, in the following we consider only the CO retrieval biases due to deficit knowledge of the relative profile of CH_4 and H_2O. For this purpose, an ensemble of relative profiles, shown in Fig. 22, is used. With increasing profile index $I_{H_2O}, I_{CH_4} = 1 \cdots 10$ the underestimation of the relative abundance in the lower troposphere changes gradually to an overestimation, which comes along with a corresponding but opposite change in the middle troposphere. Both for CH_4 and H_2O, the true relative profile lies between profile $I_{H_2O}, I_{CH_4} = 4$ and 5. In general, the change in the relative H_2O profiles in the ensemble is larger than that of the CH_4 profiles, which reflects the stronger atmospheric variability of water. However, the assumed variability is based on ad hoc assumptions and should not be interpreted as an estimate of the a priori knowledge uncertainty, which is very hard to quantify. However, we believe that the considered range exceeds the actual uncertainty, yielding a conservative error estimate.

Figure 23 and 24 present the CO bias caused by the wrongly assumed vertical distribution of CH_4 and H_2O. For the clear sky scenarios of Case A, the dependence of the CO bias on the vertical CH_4 profile is small. Also the spectral χ^2 depends little on the profile. However, in case of a water cloud, where the retrieved properties of the scattering layer depend on methane absorption, the dependence on the vertical profile becomes significant with CO biases up to 5-10 %. The scattering layer height is adjusted by the retrieval such that the absorption by CH_4 above the cloud is described properly by the forward model. So, assuming a wrong relative profile directly relates to a wrong layer height and subsequently, it affects the retrieved CO abundance and its column averaging kernel. Here, critical CO biases are obtained for the profiles $I_{CH_4} = 7, 8, 9, 10$ with enhanced CH_4 concentration at the lowest model layer of ≥ 20 %. The crossing point of

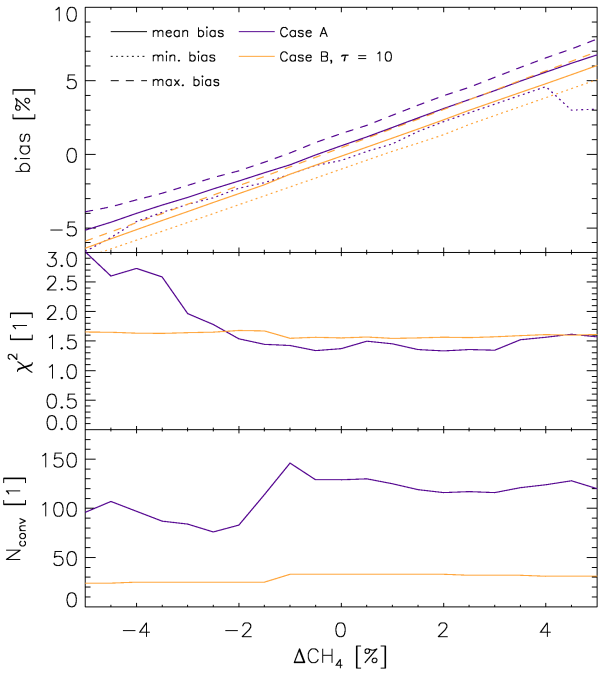

Figure 21: Same as Fig. 19 but as a function of the a priori uncertainty of the CH_4 column.

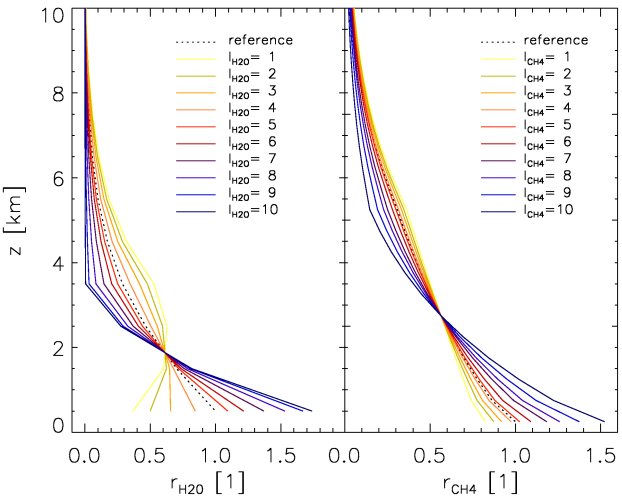

Figure 22: Relative H_2O (left panel) and CH_4 (right panel) profiles r as a function of height z. Each profile is normalised to the reference profile at 500 m, which is used for the measurement simulation.

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

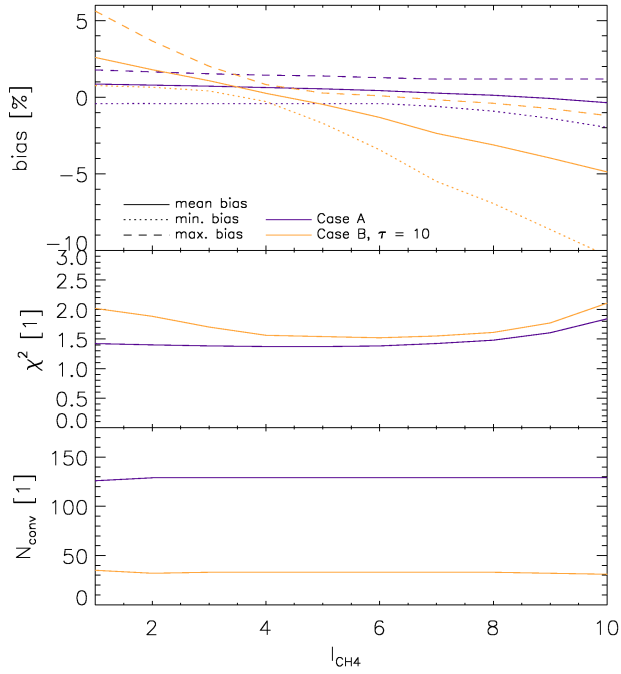

Figure 23: Same as Fig. 19 but assuming erroneous relative vertical CH_4 profile from Fig. 22 in the CO retrieval.

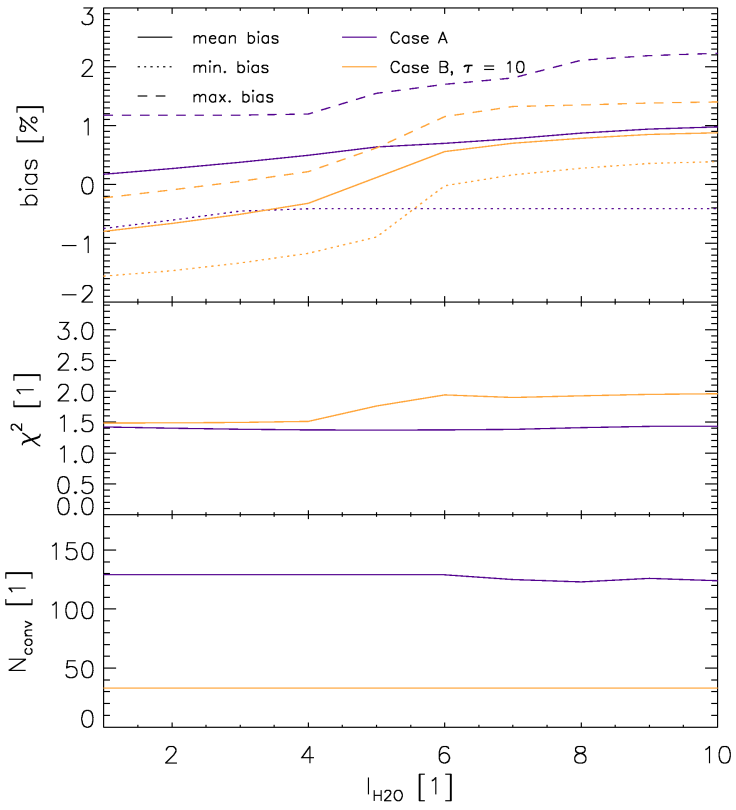

Figure 24: Same as Fig. 23 but for erroneous relative vertical H_2O profile from Fig. 22.

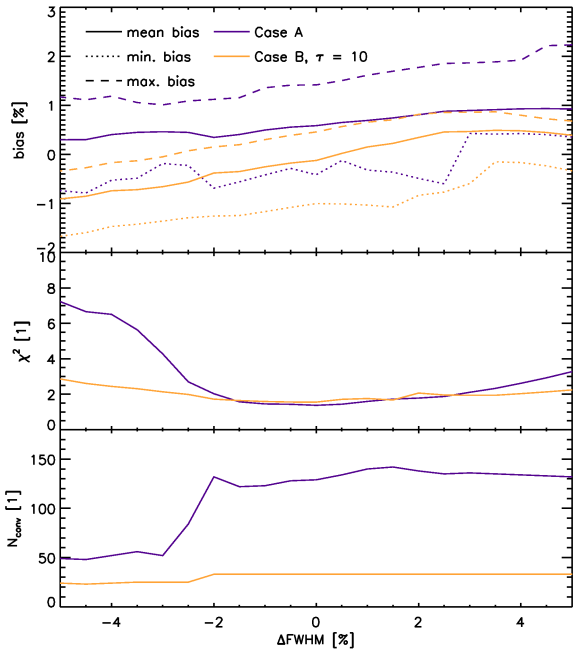

Figure 25: CO retrieval bias as a function of a FWHM error (ΔFWHM) of the ISRF. The measurement ensembles of the clear sky generic test cases A (purple lines) and the cloudy sky test case B for a cloud of optical depth 10 (orange lines) are considered. Here, the maximum (dashed lines), mean (solid lines) and minimum (dotted lines) bias is reported for the different FWHM errors. Middle panel: mean spectral χ^2 of the retrieval as a function of FWHM error. Lower panel: number of converged retrievals as a function of FWHM error.

7.5 Robustness of the CO retrieval with respect to instrument artifacts

The accuracy of the CO retrieval product may also suffer from instrument artifacts. The acceptance level of instrument artifacts is described in detail by [RD31] [AD3]. This section is not meant to consolidate these studies but to exemplarily consider errors on five instrument-related parameters and their relevance for the CO retrieval quality. It is important to realize that the SWIR spectral range is also used to retrieve CH_4 vertical column amounts, and the required instrument performance in this spectral range is driven by both, the required CO and CH_4 retrieval uncertainty. Analogous to the previous section, the performance analysis is based on the generic ensemble of case A and case B with a water cloud of optical depth $\tau_{\mathrm{cld}} = 10$. We consider measurement artifacts due to

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

To simulate the measurement spectra of the test ensembles, line-by-line spectra are convoluted with a Gaussian ISRF with an FWHM of 0.25 nm (see Eq. 3). Errors in the ISRF may introduce spectral residuals to the spectral fit and affect the CO retrieval performance. Figure 25 shows the CO retrieval sensitivity to errors of the FWHM of the ISRF between ± 5 %, where the perturbed ISRF is applied to the Earthshine and solar spectrum. Here the retrieval bias changes only very little. However, when the FWHM is underestimated by more than 2 % the spectral χ^2 of the fit increases systematically and at the same time the number of successfully converged retrievals is reduced by a factor of 2.6. Thus, to warrant the CO retrieval quality, the FWHM of the ISRF has to be known at least with an uncertainty of 2 % within the CO spectral fit windows. This is in agreement with the Mission and System Requirement Document (see [RD31] and [AD3]), which requests the FWHM knowledge of the ISRF with an uncertainty of 1 %.

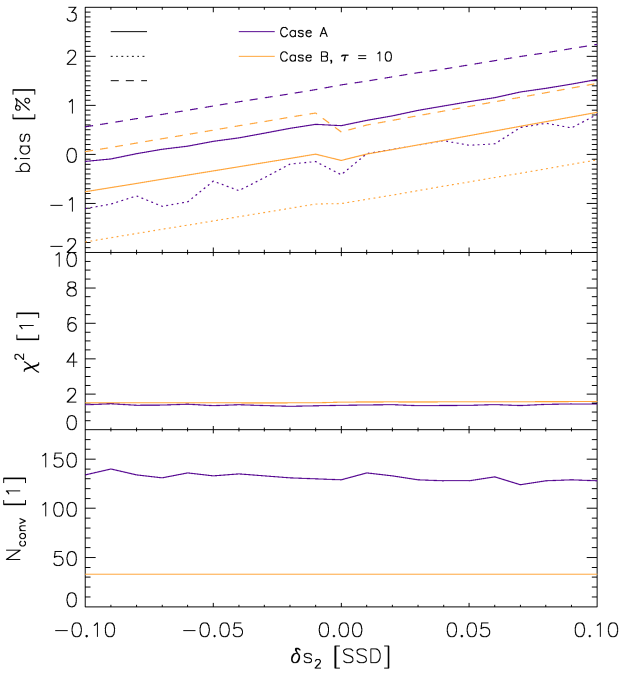

Figure 26: Same as Fig. 25 but as a function of the spectral calibration error δs_2 as described in Eq. (73).

Next, we consider an erroneous spectral calibration of the measurement. Assuming a correct instrument calibration λ_i, where spectral detector i is assigned to wavelength λ_i, we consider an erroneous calibration

$$\lambda_i' = \lambda_i + \delta s_1 + \frac{\lambda_i - \lambda_m}{\lambda_r - \lambda_m}\delta s_2 \; . \tag{73}$$

Here, $\lambda_l = 2305$ nm and $\lambda_r = 2385$ nm describe the shortwave and longwave edge of the SWIR band and $\lambda_m = 2345$ nm is its spectral center. So, δs_1 and δs_2 characterise spectral calibration errors which result in an overall spectral shift and a spectral squeeze. Both are defined over the full spectral range and the corresponding squeeze for the CO fit windows (2315-2338 nm) is about one-third of δs_2.

Because a spectral shift is a default fit parameter of the CO inversion, the retrieval is not sensitive to the calibration error δs_1. This was confirmed by retrieval simulations for the range $-0.1\text{nm} \leq \delta s_1 \leq +0.1\text{nm}$, which is the spectral range of plus-minus one spatial sampling distance (SSD) of the SWIR spectrometer (not shown). The situation differs when a calibration error δs_2 is introduced to the measurement simulation. Figure 26 shows a nearly linear dependence of the CO bias in range $-0.1 \times \text{SSD} \leq \delta s_2 \leq 0.1 \times \text{SSD}$ with a relative CO bias dependence of 1 % per 0.1 SSD. The required knowledge of the spectral position of the spectral channels is 2 pm which corresponds to 1/50 SSD [AD3]. Thus for an instrument satisfying this requirement, a calibration error conform the spectral squeeze in Eq. (73) provides a minor contribution to the CO error budget.

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

The reported error sensitivity of the CO retrieval with respect to instrument artifacts is relatively low, and the required instrument performance is mainly driven by the CH_4 column retrieval [RD2]. For CH_4, spectrally strong absorption features are needed to determined atmospheric scattering characteristics simultaneously with the abundance of water vapor and CH_4 [RD2]. Here, the radiometric accuracy in the deep absorption lines drives

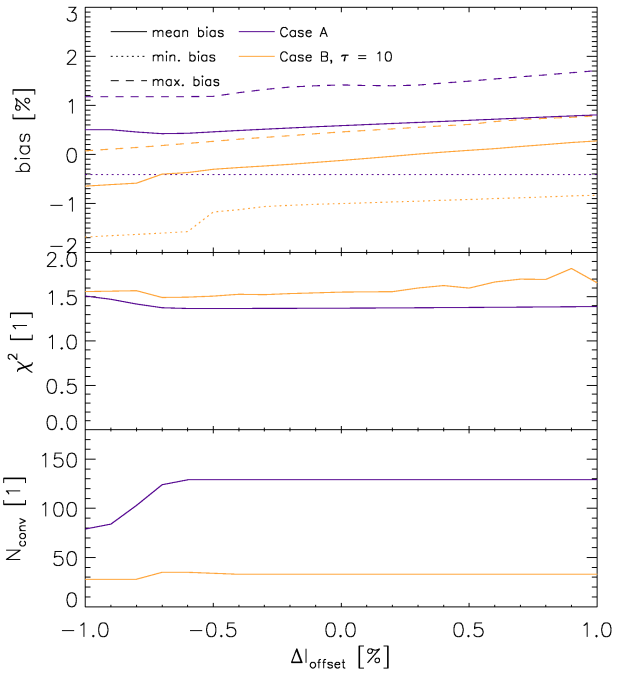

Figure 27: Same as Fig. 25 but as a function of a spectrally constant radiometric error ΔI_{offset}. Here, ΔI_{offset} is defined with respect to the continuum value of the spectrum.

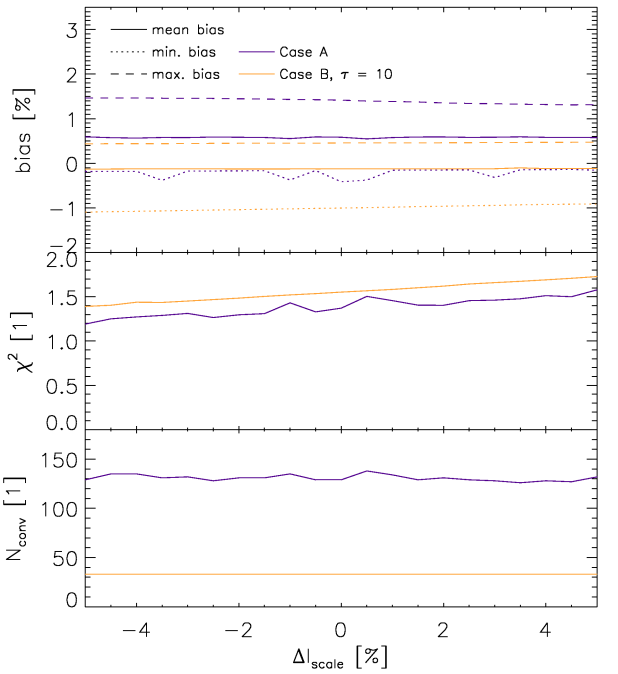

Figure 28: Same as Fig. 25 but as a function of a spectrally constant radiometric scaling error ΔI_{scale}. Here, ΔI_{scale} is defined with respect to the continuum value of the spectrum.

mainly the radiometric requirements of the instrument. For the retrieval of CO, we have selected a spectral fit window with only moderate and weak absorption of CH_4 and H_2O to reduce interference with the weak absorption lines of CO. Due to this, the sensitivity to radiometric errors is weaker for the CO retrieval. However to retrieve atmospheric scattering properties, we employ a priori knowledge about the CH_4 abundance and use the corresponding spectral features to obtain a measurement sensitivity to the height and optical depth

of a scattering layer. This puts a stringent requirement on our a priori knowledge of methane and so it shifts challenges from instrument calibration to atmospheric modeling. In the next section we discuss the accuracy of state-of-the-art chemical transport modeling to predict the CH_4 concentration.

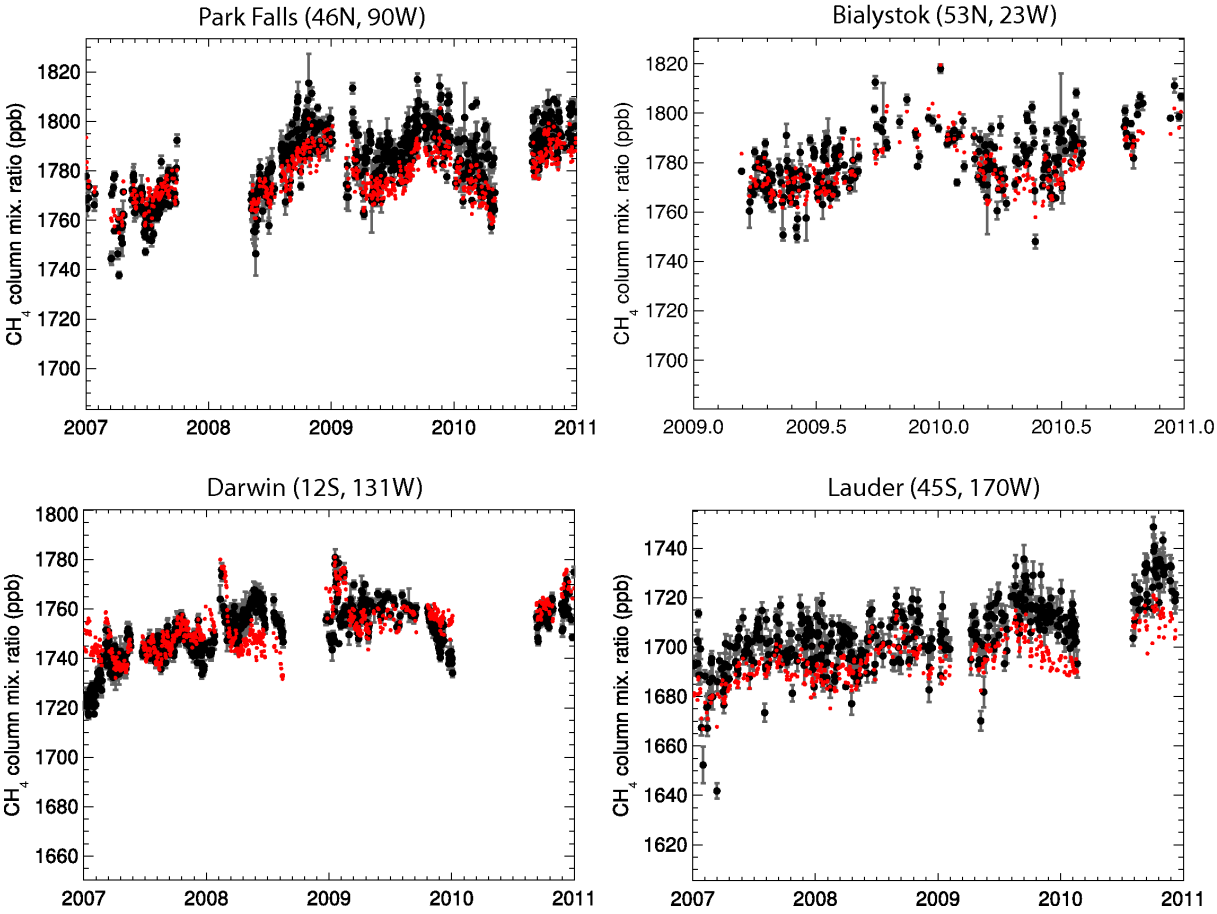

Figure 29: Comparison of TM5 simulated and in situ FTS observed total column CH_4 at selected sites of the TCCON network. Black: TCCON FTS, Red: TM5.

7.6 Quality of the model derived XCH_4

An important element of the CO retrieval approach is the use of an atmospheric transport model to estimate the methane abundance in the observed scene. Figure 21 shows the error on the retrieved CO column assuming an error of ± 3 % on the methane a priori knowledge. In both cases, the CO retrieval bias is clearly affected by this error. Overall, a ± 3 % CH_4 uncertainty causes an additional error on the retrieved CO column of about 3 %. Additionally, it affects the number of successful retrievals. So, it is essential to make a sophisticated estimate of the methane uncertainties. For this purpose, we consider the dry air mole fraction $XCH_{4,mod}$ predicted by the global transport model TM5. Moreover, we propose different options for operational delivery of $CH_{4,mod}$ in the context of the TROPOMI mission including a comparison of the expected performance.

7.6.1 Simulation of XCH_4

The TM5 model [RD33] has been used in several studies of CH_4 (e.g. [RD71, RD72, RD73]). This specific version of the model has been developed to estimate the global sources and sinks of CH_4 from atmospheric measurements using a variational inverse modeling technique. In short, the method starts with a first guess, or a priori estimate, on the basis of published emission inventories and output from land surface models. A forward run of the TM5 model yields global maps of the spatio-temporal distribution of CH_4 in the atmosphere, which are sampled according to the times and locations of available measurements. The inverse modeling technique optimizes the sources and sinks by minimising a least squares cost function, consisting of the residual differences between model and measurements and deviations from the first guess. This procedure has been used to optimize TM5 on the basis of measurements from the NOAA-ESRL global monitoring network and of CH_4 retrievals from SCIAMACHY and GOSAT.

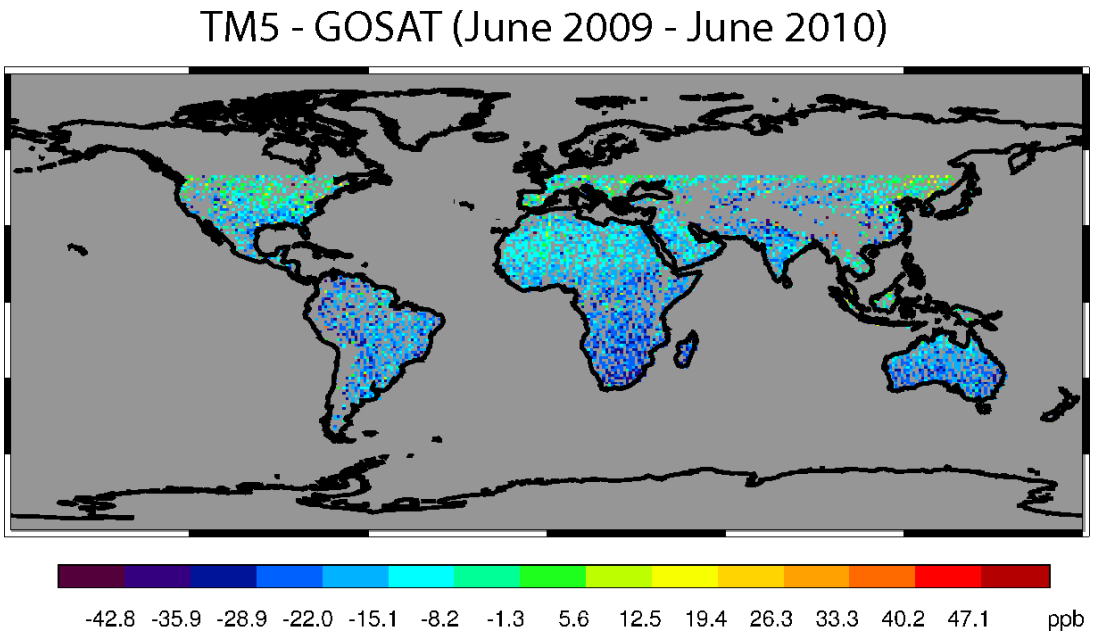

Figure 30: Comparison between TM5 simulated and GOSAT retrieved XCH_4. Differences (TM5 minus GOSAT) are shown for the period June 2009–June 2011. The TM5 results have been optimised using surface measurements.

7.6.2 Uncertainty of model-derived XCH_4

The overall uncertainty of model derived XCH_4 can be separated into two components: 1. Limitations in the state-of-the-art knowledge of XCH_4, 2. Limitations of the applied modeling procedure to reproduce the state-of-the-art. The second contribution refers to limiting conditions, mostly due to requirements on the timeliness of data delivery, which prompt for a non optimal approach.

Contribution 1 is assessed by comparing inverse modeling optimised XCH_4 to high precision measurements from the Total Carbon Column Observing Network (TCCON) of ground based Fourier Transform Spectrometers (FTS). Figure 29 shows examples of such comparisons for a selection of sites spanning a wide range in latitude. In these comparisons, the model has been optimised using measurements from 46 background sites of the NOAA-ESRL network.

The root mean square difference in diurnal mean XCH_4 between the TCCON measurements and the NOAA optimised model ranges from 8 to 22 ppb between sites, with a mean of 13 ppb for the 10 sites included in the analysis. As can be seen, the RMS differences are explained in part by a small bias of the model, which increases towards southern latitudes. Without the contribution of this systematic offset, which could in principle be corrected for, the mean RMS difference reduces to 8 ppb. However, it should be realised that the TCCON network is not designed to be well representative for typical conditions encountered over land world wide. Therefore the optimised model may on average be further from the truth than indicated by the TCCON residuals. Limited means are available for quantifying this difference. Figure 30 shows results from an attempt in this direction using retrievals from the Greenhouse Gas Observing Satellite (GOSAT). The results are largely consistent with the TCCON analysis, showing a systematic underestimation of TM5 simulated XCH_4 increasing towards southern latitudes. The RMS difference between TM5 and GOSAT amounts to 20 ppb. Part of this difference is explained by the uncertainty of GOSAT XCH_4 retrievals, which has been estimated at 15 ppb [RD74].

From Fig. 30, it can be concluded that the use of GOSAT in the optimisation of TM5 reduces the uncertainty of model simulated XCH_4. Besides measurement continuity, also timeliness needs further consideration. Commonly, inverse modeling derived estimates lag behind real-time by approximately a year. This has to do with the availability of various types of inputs that are required, including meteorological driving fields, a priori emission estimates, and measurements. The most important limitation at present is the release of new surface measurements, which cannot easily be accelerated because of standard calibration and verification procedures adopted by the laboratories involved. Alternatively, other modeling procedures can be followed to achieve near real time model estimates of XCH_4, such as:

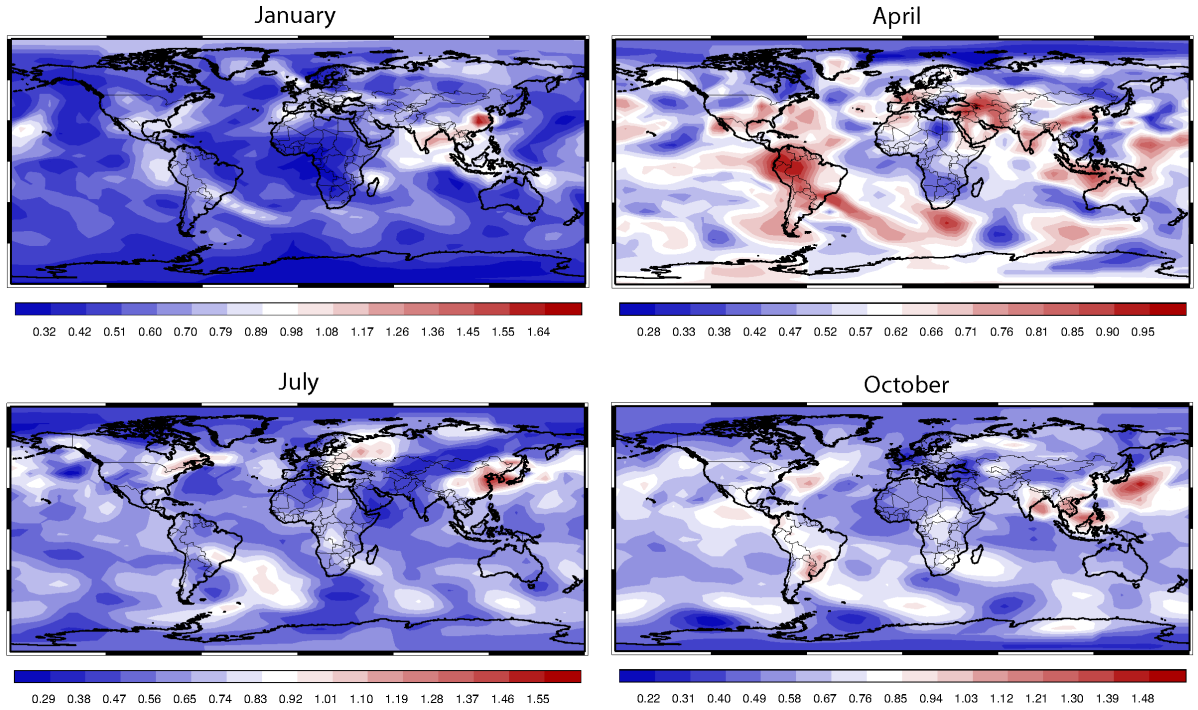

Figure 31: The standard deviation of total XCH_4 between years expressed in % of mean XCH_4. Standard deviations are calculated from TM5 XCH_4 fields, optimised using surface measurements, for the 15th day of the month in each year in the period 2003–2010.

1. Use of inversion-optimised TM5 estimates of a different year.
2. Option 1 extrapolated to the current year using a priori emissions.
3. Option 1 extrapolated to the current year using a posteriori emissions of the previous year.

It is clear that each of these options is less accurate than the NOAA optimised model discussed so far, which brings us to the estimation of uncertainty component 2 mentioned at the start of this section. For procedure 1, the uncertainty component 2 is the variability in total column XCH_4 caused by year-to-year variations in meteorology and the inter-annual variation of methane sources and sinks. The size of this error is estimated from results of a multi-year inversion, by calculating how XCH_4 on a given day of the year varied between the years. Figure 31 shows results of this procedure applied to selected days over the year.

The results highlight regions with a high variability in atmospheric transport, such as the storm tracks in the northern and southern hemisphere. The largest variations are found over South East Asia, due to large regional sources of methane. Some enhanced variability can be seen near the equator, where the north-south gradient in XCH_4 is largest. Standard deviations are on average well within 1% (18 ppb), regionally increasing up to 1.5% (27 ppb). Sporadically, standard deviations up to 3% are found, associated with biomass burning events.

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

The impact of systematic differences between a priori and a posteriori fluxes affecting procedure 1 can be reduced by extrapolating inverse modeling results using a posteriori instead of a priori fluxes. Obviously, the inconsistency that remains in this case (procedure 3) is that the fluxes are optimised for a different year. As shown in Fig. 33, this procedure avoids growing offsets in the latitudinal gradient over time. Instead, the largest variations show up in the tropics, in response, for example, to inter-annual variability of tropical sources such as wetlands and biomass burning. It should be realised that tropical sources of methane are relatively poorly constrained, both a priori and by the limited availability of surface measurements in the tropics. Compared

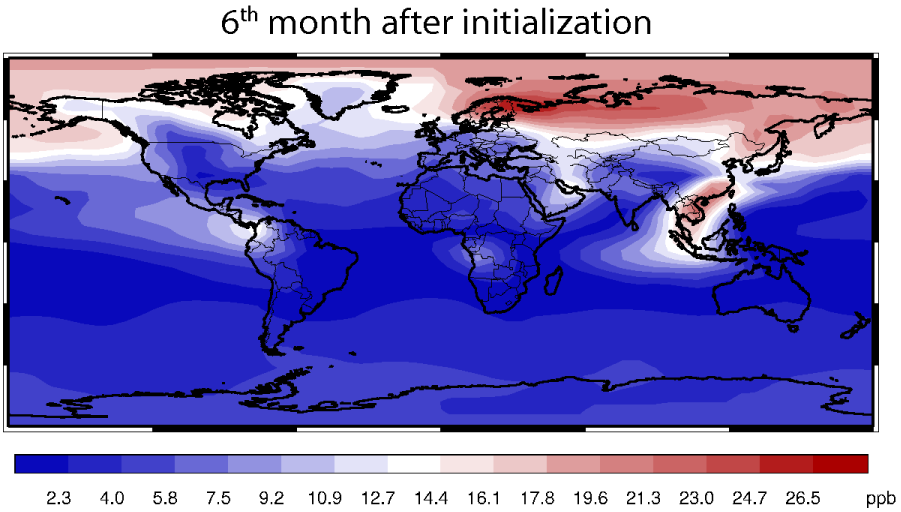

Figure 32: The RMS difference between NOAA optimised CH_4 and results of procedure 2 evaluated after 6 months. RMS values are calculated from the differences between the two simulations for all days of the 6th month after initialisation of procedure 2.

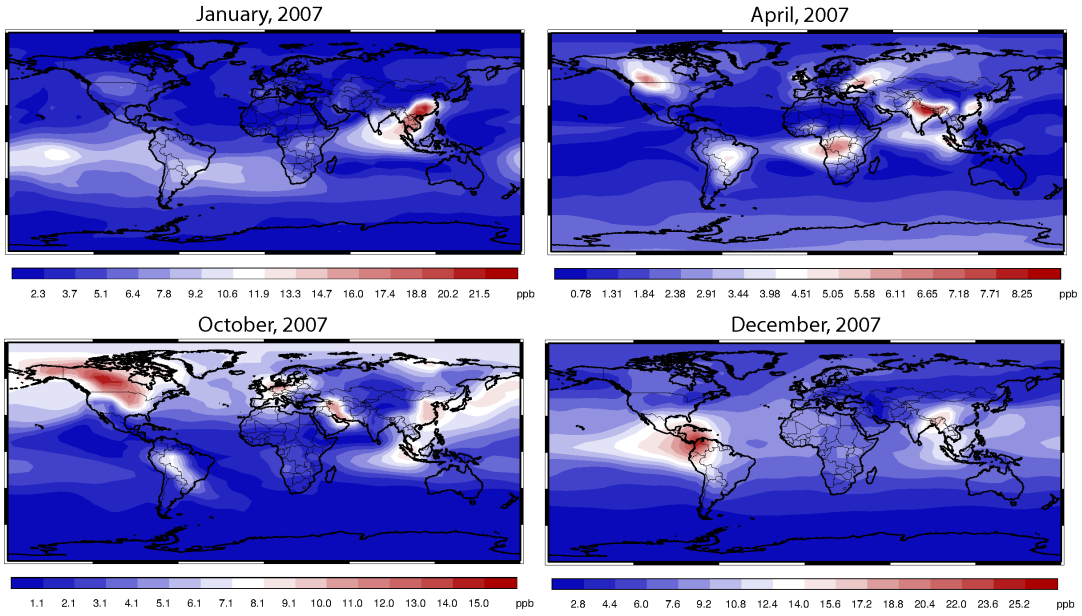

Figure 33: As Fig. 32 for procedure 3, evaluated in different months.

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

7.6.3 Discussion and conclusions

As demonstrated in the previous section, the accuracy of model-derived XCH_4 estimates depend on the optimisation approach. Options for model optimisation, in turn, depend on timeliness requirements on the delivery of data products. Table 4 lists uncertainties for each of the proposed options, calculated from the available statistics for three regions: 1. Global land, 2. South East Asia, 3. Tropical America. The most accurate approach to estimating $XCH_{4,\mathrm{mod}}$ is to optimize the model using both surface and satellite measurements of

Table 4: Estimated uncertainty in XCH_4 comparing the proposed modeling approaches.

Method	Land	S.E. Asia	Trop. America
	2σ (ppb (%))	2σ (ppb (%))	2σ (ppb (%))
Optimised TM5	26 (1.5)	-	-
Procedure 1	32 (1.9)	38 (2.2)	35 (2.0)
Procedure 2	30 (1.7)	32 (1.9)	30 (1.7)
Procedure 3	28 (1.6)	30 (1.7)	28 (1.6)

CH_4, which limits the uncertainty of $XCH_{4,\mathrm{mod}}$ to within 1% (18 ppb). When using only surface measurements, the uncertainty is estimated at 26 ppb (at 95% confidence), based on comparisons with TCCON data. This should be considered a lower limit, because of the limited representativity of the TCCON network. On the other hand, part of the error is systematic and can to first order be corrected on the basis of available TCCON and GOSAT measurements. The overall uncertainty is difficult to quantify but should remain within 2% (35 ppb).

Without the possibility to resort to optimised model estimates, additional errors are introduced, depending on the adopted procedure. The least demanding approach, with respect to timeliness requirements, is to use results from the most recent optimised year. In this case, the additional error is on average within 20 ppb (at 95% confidence), increasing to 28 ppb (1.6%) over intense source regions of South East Asia. Combining these uncertainties with the uncertainty of the optimisation approach (taken as 26 ppb), yields an overall uncertainty of 32–38 ppb (1.9–2.2%) assuming no interdependence between the errors. The transport model uncertainty can be further reduced by using actual meteorology, which limits the timeliness to about half a year behind real-time. This approach works best using optimised instead of a priori fluxes for the years that were analysed, which yields estimated uncertainties in the range of 28–30 ppb (1.6–1.7%). Our assessment of the 2^{nd} uncertainty component for procedure 1–3 may be somewhat optimistic, because of the limited resolution of the transport model and potential underestimation of inter-annual variability of CH_4 emissions using surface data. Acknowledging these limitations in our approach, an uncertainty of 2% seems a reasonable save margin that should be achievable for most conditions encountered throughout the global domain.

8 Algorithm input and output

8.1 High level processing scheme

The operational CO and CH_4 algorithms use a similar processing scheme, which for CO is summarised in Fig. 34. Red boxes indicate modules that are described in this ATBD, light green boxes are modules that are developed in the context of the overall processing framework (see e.g. [AD4]). The algorithm input is divided in static input and dynamic input. Static input is read once, during algorithm initialisation, whereas dynamic input is different for each pixel.

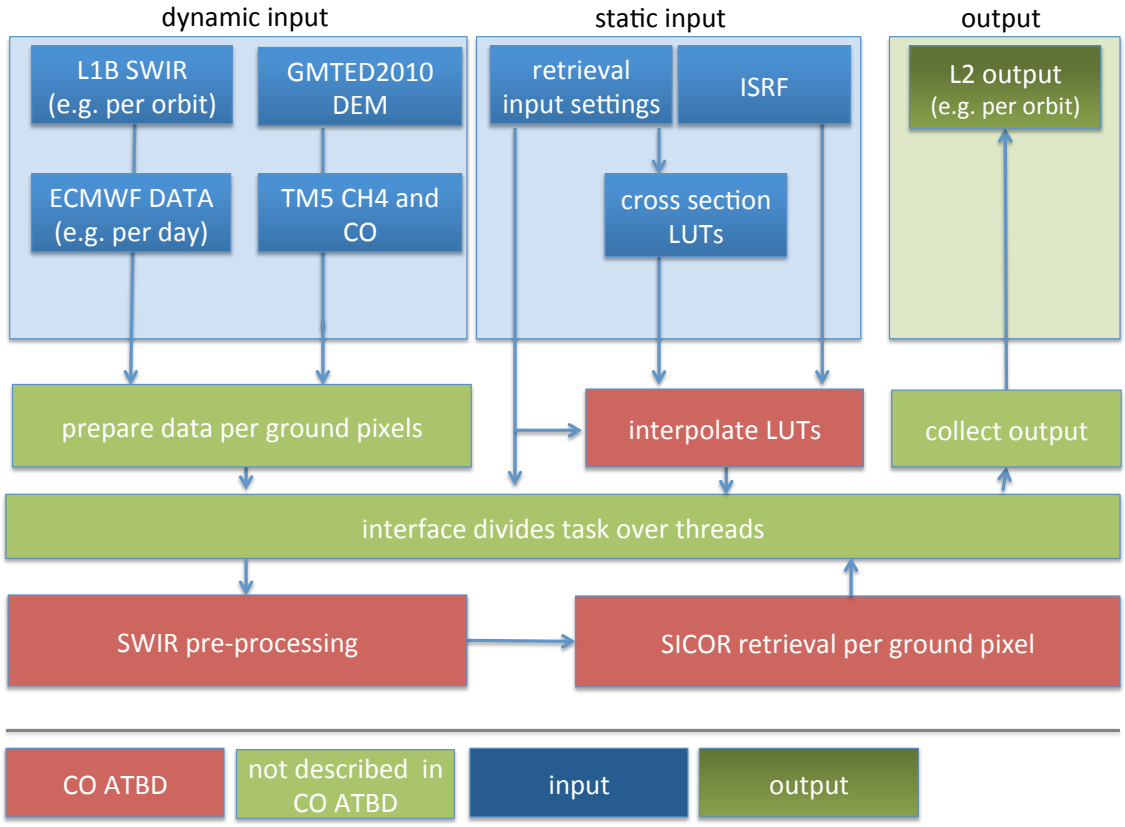

Figure 34: High level processing scheme for operational S5P CO data reduction. Modules that are described in this ATBD, are indicated by the red boxes.

8.2 Static input

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

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

8.3 Dynamic input

The dynamic input is listed in table 6. These data are specific for each measurement pixel. Dynamic data coming from different sources should be co-located to the horizontal position and observation time of the Level 1b product. Nearest neighbor approximation is also possible. Vertical domains and sampling may differ for the different input fields. The original ECMWF and TM5 data should be given on their original vertical hybrid grid.

Table 5: SICOR Static input. Calibration key data and irradiance L1b-product are semi-static, because they are provided once per processor run.

Name/Data	Symbol	Unit	Source	Pre-process needs	Comments
Calculation settings	–	–	SRON	–	
Absorption cross sections[1]	σ	cm^2	[RD57, RD75]	–	
Reference irradiance spectrum	I_{ref}	mol s^{-1} m^{-2} nm^{-1} sr^{-1}	KNMI	–	
ISRF	S_R	nm^{-1}	Calibration key data	–	
Solar irradiance SWIR	I	mol s^{-1} m^{-2} nm^{-1}	S5P L1b product	–	
Noise solar irradiance SWIR	ϵ_I	mol s^{-1} m^{-2} nm^{-1}	S5P L1b product	–	
Irradiance pixel mask	l_I	–	S5P L1b product	–	

[1] A lookup table for CO, CH_4, H_2O and HDO on a grid of temperatures, pressures and wavenumbers, covering the SWIR band.

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

Table 6: SICOR Dynamic input.

Name/Data	Symbol	Unit	Source	Pre-process needs	Backup if not available	Comments
Earth radiance SWIR	R	mol s^{-1} m^{-2} nm^{-1} sr^{-1}	S5P L1b product	–	No retrieval	
Noise Earth radiance SWIR	ϵ_R	mol s^{-1} m^{-2} nm^{-1} sr^{-1}	S5P L1b product	–	No retrieval	
Radiance pixel mask	l_R	–	S5P L1b product	–	No retrieval	
Solar zenith angle	θ_0	degree	S5P L1b product	–	No retrieval	
Viewing zenith angle	θ_v	degree	S5P L1b product	–	No retrieval	
Relative azimuth angle[1]	φ	degree	S5P L1b product	–	No retrieval	
Longitude	lat	degree	S5P L1b product	–	No retrieval	
Latitude	lon	degree	S5P L1b product	–	No retrieval	
SWIR mean pixel elevation	z_s	m	GMTED 2010	Interpolate	Same as z_{s0}	
Temperature profile	T	K	ECMWF	Interpolate	Previous day	
Specific humidity profile	q	kg kg^{-1}	ECMWF	Interpolate	Previous day	
Surface pressure	p_0	Pa	ECMWF	Interpolate	Previous day	
ECMWF surface elevation	z_{s0}	m	ECMWF	Interpolate	Previous day	
CO mixing ratio profile	m_{CO}	1	TM5	Interpolate	Previous day	
CH_4 mixing ratio profile	m_{CH_4}	1	TM5	Interpolate	Previous day	

[1] Zero degrees means that the instrument is looking towards the sun.

8.4 Algorithm output

For each spatial sounding, the level-2 output is generated and the corresponding SICOR output data structure is listed in table 7. For the operational data processing, the SICOR CO product is processed as part of both the near-real-time processor and the offline processor. Here, a subset of output fields will not be calculated in the near-real-time processor and will be filled with NetCDF fill values. The total output product is 476 bytes per pixel. More details about the SICOR CO data product is given in the CO product user manual [RD76].

Table 7: SICOR output fields. N_z is the number of layers in the model atmosphere, and is set to 50 by default.

Name/Data	Symbol	Unit	Description	Data type	Number of values per pixel	Comments
CO column	vco	mol m^{-2}		Float	1	Main product
Noise CO column	eco	mol m^{-2}		Float	1	Main product
CO column averaging kernel	avg	m		Float	N_z	Main product
Layer interface pressure	press	Pa		Float	N_z	Auxiliary product
H_2O column	vh2o	mol m^{-2}		Float	1	Auxiliary product
Noise H_2O column	eh2o	mol m^{-2}		Float	1	Auxiliary product
HDO column	vhdo	mol m^{-2}		Float	1	Auxiliary product
Noise HDO column	ehdo	mol m^{-2}		Float	1	Auxiliary product
Prefit CH_4 column	vch4_p	mol m^{-2}		Float	1	Cloud filter product
Weak-band H_2O column	vh2o_w	mol m^{-2}		Float	1	Cloud filter product
Strong-band H_2O column	vh2o_s	mol m^{-2}		Float	1	Cloud filter product
Weak-band CH_4 column	vch4_w	mol m^{-2}		Float	1	Cloud filter product
Strong-band CH_4 column	vch4_s	mol m^{-2}		Float	1	Cloud filter product
Scatterer optical depth 2330 nm	tausat	1		Float	1	Auxiliary product
Scatterer centre height	zscat	m		Float	1	Auxiliary product
Albedo 2325 nm	alb_1	1		Float	1	Auxiliary product
Albedo 2335 nm	alb_2	1		Float	1	Auxiliary product
Spectral shift	spec_off	nm		Float	1	Auxiliary product
Residual χ^2	chi2	1		Float	1	Quality product
Number of spectral bins N_y	ny	1		Integer	1	Auxiliary product
Degree of freedom for signal DFS	dfs	1		Float	1	Quality product
Iterations	niter	1		Integer	1	Quality product
Processing quality flag	PQF	–		Integer	1	Quality product

Table 8: Calculation settings and computation time for China ensemble.

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

9 Feasibility

9.1 Estimated computational effort of the SICOR module

The CO data product is processed both in the near-real-time and offline data stream of the S5P level-1 to level-2 data processor. To evaluate the computation cost of the algorithm, we assume a continuous data flow of 260 spectra/sec. during 50 minutes of the 100 minute orbit period and a continuous Level-2 data processing. About half of the measurements are performed during the eclipse period of the satellite orbit, which effectively means a mean processing request for 130 spectra/sec. To estimate the computation time of the retrievals, we consider the test ensemble over China as described in Sec. 7.3, consisting of 25862 spectra and these spectra were calculated on one local computer without any parallelisation. The local computer is a HP dc7900 SFF with Intel® Core™2 Duo CPU E8400 at 3.00 GHz and 4 GB RAM with a SPEC (Standard Performance Evaluation Corporation) floating point rate of 237, which is a computer benchmark specification for CPU's floating point processing power. To evaluate numerical performance one processor core is used. Executables were generated with the Intel FORTRAN compiler with optimisation flag O2. Table 8 provides also the computation time using the GNU gfortran compiler with optimisation flag O2. Overall, the mean computation time of the methane filter as part of the SWIR pre-processor is 0.03147 seconds and the mean computation time of the SICOR CO scattering retrieval is 0.16875 seconds (excluding file I/O).

From these 25862 spectra, 7236 (28 %) were filtered out by the methane cloud filter with a threshold of $\delta\mathrm{CH}_4 < 25\%$ and so 72 % of the data requires full processing. In contrast, for one year of GOSAT observation, which is already considered in Fig. 4, 80.6 % of the non-scattering retrievals passes the cloud filter (79.7 % for ocean pixels and 82.4 % for land pixels). Figure 35 shows the corresponding global distribution of the fraction of data to be processed. This discrepancy might be explained by the fact that the China ensemble represents an atmospheric snapshot and so cannot be used for an overall performance estimate for TROPOMI. Thus to estimate of the computational effort, we assume that a scattering retrieval has to be performed for 80.6 % of all data. This means a mean computational time of $0.03147 + 0.806 \cdot 0.16875 = 0.1675$ seconds per spectrum and so, to handle the data stream by a continuous data processing, in total 22 cores of the described hardware are needed. Concerning the software implementation, we expect no significant improvement of the algorithm runtime performance in the near future.

9.2 Spatial data selection approach

For CO, we aim to process data over land and ocean, where over ocean only cloudy scenes are considered. Here, measurements are selected with a Lambert-equivalent reflectivity (LER) > 0.03 at the continuum of the spectrum (2315 nm). The LER is provided as a standard output of the SWIR pre-processor (see Appendix B). For land pixels without inland waters, the LER value rarely goes below this threshold as indicated by the global distribution of the surface albedo, shown in Fig. 36. However, the LER value may fall below the threshold value for land pixels which are covered partly by inland waters or clear sky ocean pixels. In the SWIR spectral range, water surfaces reflect too little light to obtain useful measurements for these scenes. This LER filter reduces the number of calculations which depends mainly on the cloud coverage of ocean pixels. This is not considered in the analysis of the computational effort in Sec. 9.1.

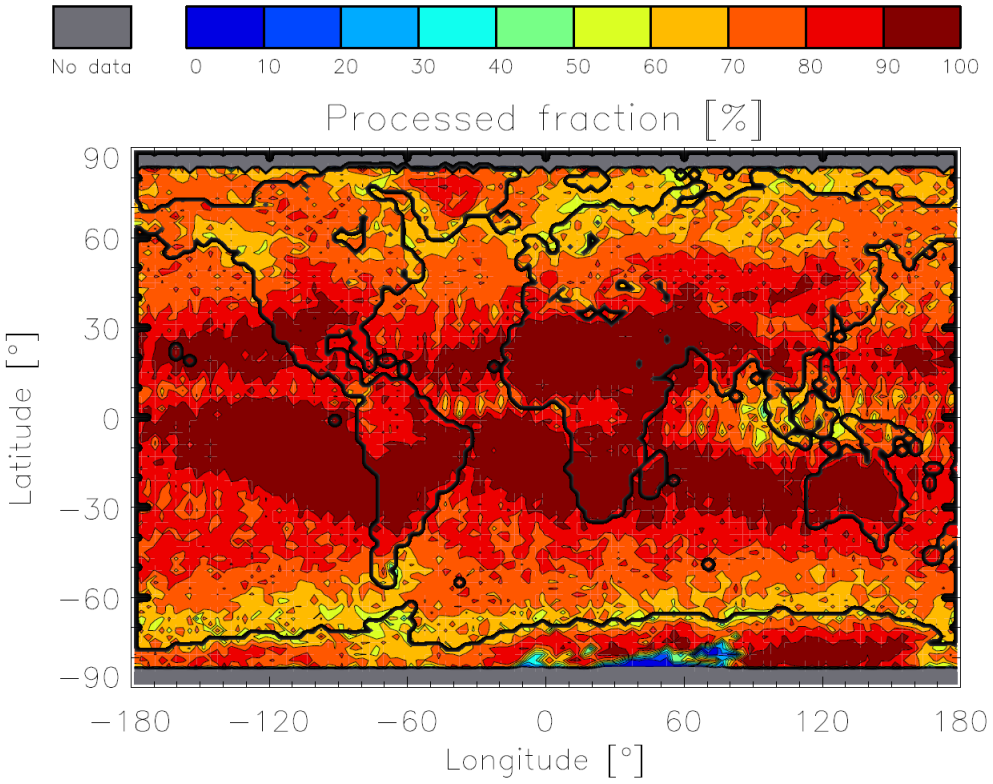

Figure 35: Fraction of GOSAT non-scattering retrievals which are accepted by the cloud filter $\Delta CH_4 < 25$ %. The analysis is based on one year (2010) of GOSAT non-scattering retrievals (RemoTeC V2.1).

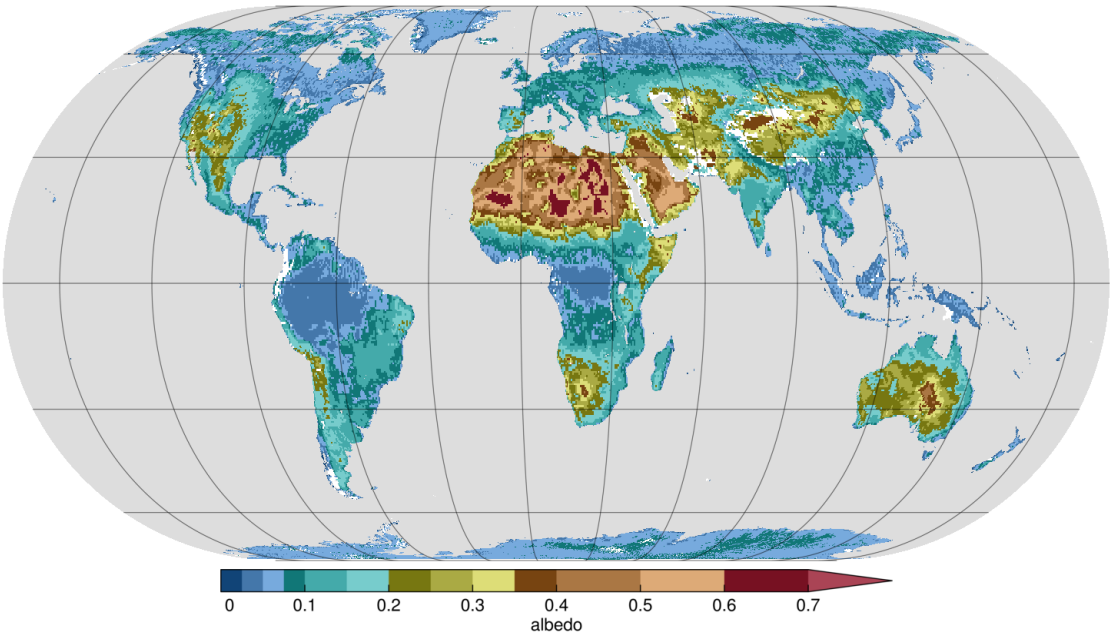

Figure 36: Average surface albedo over five years of SCIAMACHY cloud-free land observations at 2300 nm at a resolution of 0.5° (2003–2007).

10 Validation

The validation of the CO column product depends on whether or not information on the vertical profile of CO is available. In case of availability, we can refer to Eq. (53), where the independent measurement of the CO profile $\mathbf{x}_{\rm est}$ is considered as an estimate of the true profile $\mathbf{x}_{\rm true}$. Thus, the retrieved column $\hat{c}$ can be verified considering the difference

$$\delta c_{\rm val} = \hat{c} - \mathbf{A}_c \mathbf{x}_{\rm est} \; . \tag{74}$$

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

$$\delta c_{\rm val} = \hat{c} + (\mathbf{I} - \mathbf{A}_c)\mathbf{x}_{\rm ctm} - c_{\rm est} \; . \tag{75}$$

Generally, to evaluate $\delta c_{\rm val}$ in Eqs. (74) and (75), the statistical error, pseudo-statistical errors and systematical errors on predefined spatial and temporal scales have to be considered. This requires sufficient validation measurements and hence the ground based networks are considered to provide the primary data sets for validation.

In the following paragraphs, we describe shortly the different data sets that, to our knowledge, are presently available for the CO validation.

TCCON

The Total Carbon Column Observing Network (TCCON, https://tccon-wiki.caltech.edu) is a network of ground-based Fourier Transform Spectrometers recording direct solar spectra in the near-infrared spectral region in the spectral range between 0.7 and 2,5 μm. From these spectra, accurate and precise column-averaged abundance of CO_2, CH_4, N_2O, HF, CO, H_2O, and HDO are retrieved [RD77]. By default, CO is determined from a spectral window around 2.3 μm with a precision of 0.5 ppb and an accuracy of 3 ppb. The TCCON began in 2004 with the installation of the first instrument in Park Falls, Wisconsin, USA, and has since grown to 19 operational instruments worldwide.

IRGW

The Infrared Working Group (IRWG, http://www.acd.ucar.edu/irwg/) represents a similar network of infrared solar Fourier-transform spectrometers that is part of the Network for the Detection of Atmospheric Composition Change (NDACC). It is a multi-national collection of over twenty high-resolution spectrometers that regularly record the atmospheric absorption spectrum from sites distributed from pole to pole. Solar absorption spectra are used to retrieve concentrations of a number of the gaseous atmospheric components, including: O_3, HNO_3, HCl, HF, CO, N_2O, CH_4, HCN, C_2H_6, H_2O, HDO and $ClONO_2$. In contrast to the TCCON network, direct solar measurements are performed at longer wavelengths of the solar spectrum, e.g. for the retrieval of CO, three standard fit windows are employed between 4.8-4.9 μm [RD78, RD79]. The accuracy of the CO total column of this data source is about 5 % (tbc). It should be noted that in terms of geo-location the NDACC and TCCON network have a large overlap.

IAGOS MOZAIC

As part of the European Research Infrastructure IAGOS-ERI (http://www.iagos.fr/web/), the MOZAIC program provides airborne in-situ measurements for O_3, H_2O, CO, and total nitrogen oxides (NO_y) since August 1994. Measurements are geo-localised (latitude, longitude and pressure) and come with meteorological observations (wind direction and force, temperature). Data acquisition is automatically performed during

round-trip international flights (ascent, descent and cruise phases) from Europe to America, Africa, Middle East, and Asia. During ascent and decent, trace gas height information is recorded which is of particular interest to validate the S5P CO column product. However, care must be taken in case of strong spatial gradients of CO as the TROPOMI and MOZAIC spatial sampling will not be exactly coalligned.

AIRCORE

The AirCore ([RD80] and http://www.esrl.noaa.gov/gmd/ccgg/aircraft/aircore.html) is an innovative atmospheric sampling system that consists of a long coil of stainless steel tubing. It is a recent development with great potential for the validation of satellite observations. AirCore ascends on a helium balloon and fills with surrounding atmosphere during a parachute-controlled descent, collecting a sample from balloon burst (up to 30 km) down to ground level. An AirCore sample can be analysed in the laboratory for concentrations of trace atmospheric gases. The length of the tubing and short time to analysis minimises mixing inside the tubing, so that each AirCore sample provides up to 100 measurements of CO_2, CH_4, and CO from top altitude to ground level. These measurements are calibrated on the World Meteorological Organization scales (expected within 0.05% for CO_2 and CH_4, 5% for CO) and has the potential to provide a ground-truth standard for comparison with total column measurements from either ground-based Fourier Transfer Spectrometers or satellites.

MOPITT

MOPITT (Measurements of Pollution in the Troposphere) is a payload scientific instrument launched into Earth orbit by NASA on board the Terra satellite in 1999. It is designed to monitor changes in pollution patterns. MOPITT measurements yield atmospheric profiles of CO volume mixing ratio and CO total column values using near-infrared radiation at 2.3 μm and thermal-infrared radiation at 4.7 μm. The mission was designed for a five year lifetime but the instrument is still functional. For the validation of the TROPOMI CO data product, an inter-comparison with the SWIR MOPITT data product is very useful and has to be seen complementary to the validation using ground based measurements.

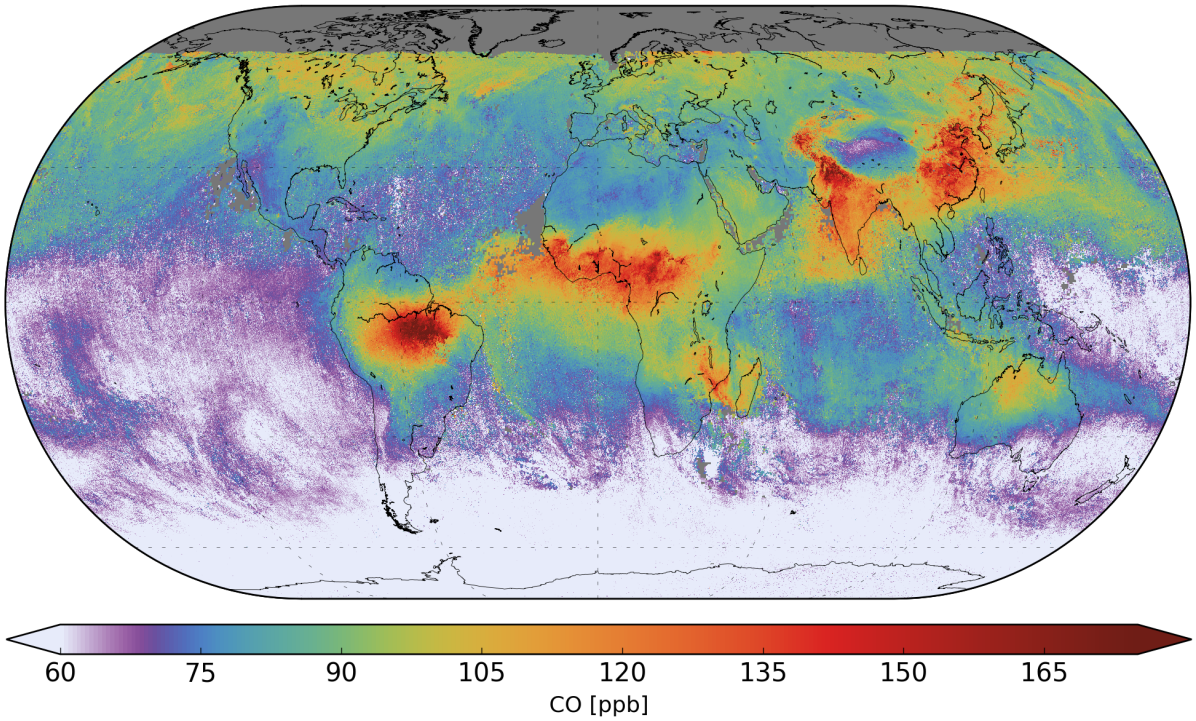

Figure 37: CO total column mixing ratios of TROPOMI averaged from November 13th to 19th, 2017 from (from Borsdorff et al. 2018)

11 Examples of TROPOMI CO 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 CO total column product of the offline data stream with remarkable accuracy. This section shows first examples of the TROPOMI CO product and summarizes preliminary data validation as presented by Borsdorff et al., 2018 [RD81, RD82].

Figure 37 shows first results of global CO observations of TROPOMI for the six subsequent days November,13th-17th, and 19th. On November 18th the instrument performed inorbit calibration measurements and so no radiance measurements are available for this day. For the data analysis, Borsdorff et al. selected only observations with a solar zenith angle (SZA) < 80^o and did not consider the two most westward pixels of the swath, because of a not yet resolved performance issue. Moreover, clear-sky and cloudy sky observations with a cloud height z_{cld} < 5000 m were selected. The figure clearly reflects CO enhancements by strong sources like wild fires (e.g. Brazil, Africa, Madagascar, and Australia) as well as anthropogenic air pollution in India and China.

The good signal-to-noise ratio of the measurements in combination with the high spatial resolution also permits to detect enhancements by weak regional sources. For example, atmospheric CO enhancements by individual wild fires can be detected by daily overpasses of TROPOMI. Beginning of May, 2018, Siberia was primed by dry and warm weather causing wild fires to lit up in East-Russia around Komsomolsk-on-Amur. The Copernicus satellite Sentinel-2 captured active fires and smoke plumes in this area, depicted in Fig. 39 for May 9th, 2018. At the same time, TROPOMI measures very high CO values of up to 1000 ppb, as shown in Figure 40 with a clear CO plume due to winds from South-West. These two images nicely illustrates the potential synergies of different Copernicus missions.

Moreover, already in this early phase of the mission TROPOMI demonstrated its capability to detect air pollution above cities, urban and industrial areas on a daily basis, which belongs to the most ambitious objectives of the mission. For example, Fig. 38a shows enhanced CO values over the industrial area near to Venice as well as pollution above Turin, Milan, and Rome. For example, Figure 38b depicts an orbit overpass over Saudi Arabia and Egypt and shows distinct pollution patterns over Mecca, Jeddah and Cairo. Furthermore, enhanced CO values along the Nile indicate air pollution in this densely populated region. Figure 38c clearly

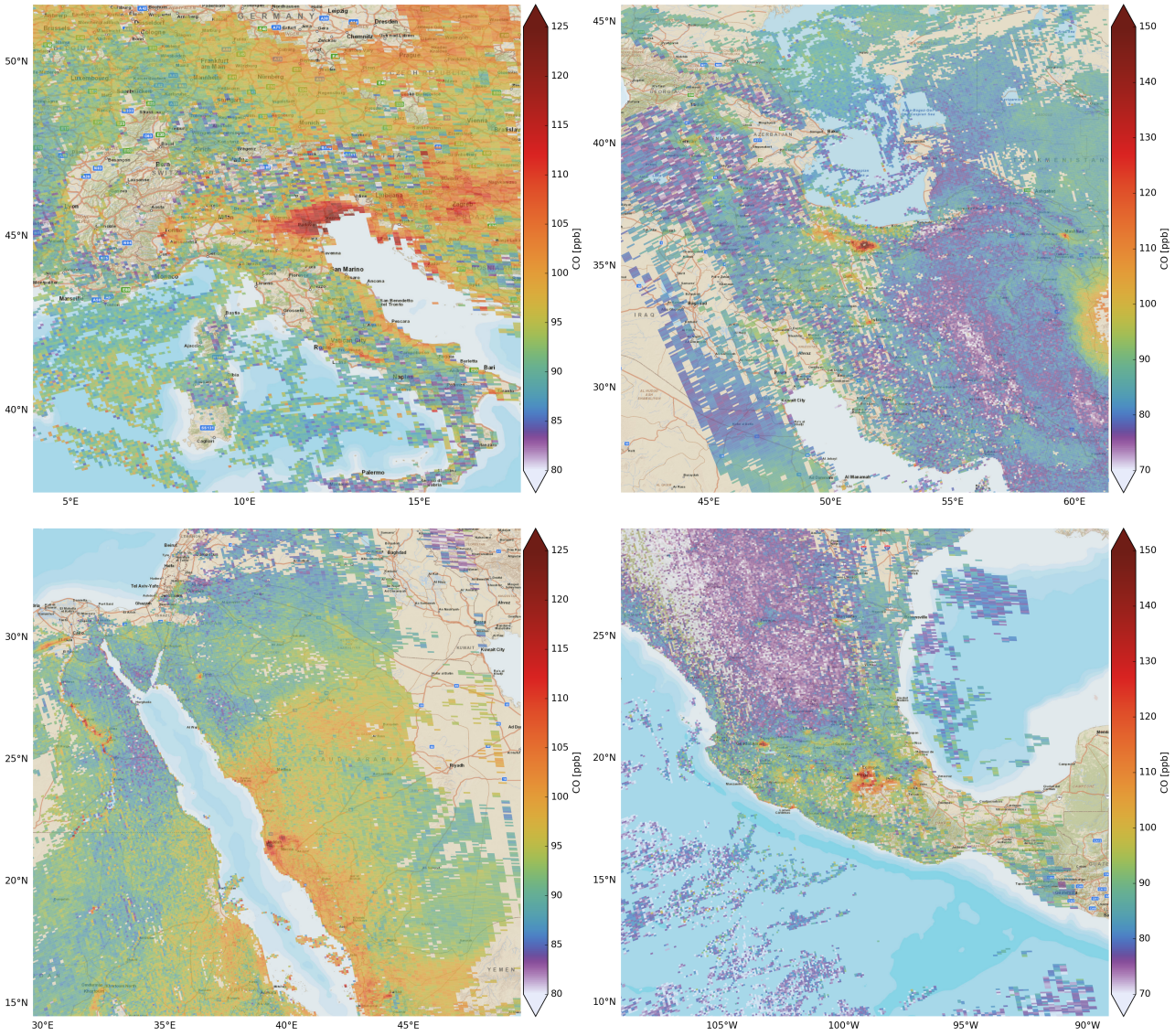

Figure 38: Total column mixing ratio (XCO) for individual TROPOMI ground pixels for (a) Italy on 25th December, (b) Saudi Arabia and Egypt on 12th November 2017, (c) Iran on 17th November 2017, and (d) Mexico on 25th November 2017.

shows the enhanced CO values above Tehran, in agreement with the urban area of the city. Also smaller cities in the region like Qom, Isfahan, and Mashhad can be distinguished from the background CO level. Finally, Fig. 38d shows strong CO enhancements above Mexico City, Guadalajara, Torreón, and Monterrey. Data gaps in the figures are caused by the filtering of measurements under clear-sky conditions over the oceans and measurements contaminated by high altitude clouds. The daily global coverage of TROPOMI and so the temporal evolution of air pollution on city scales opens up new possibilities to monitor the effect of emission regularization but also requires estimates of the absolute uncertainty of the TROPOMI CO product.

For a first data quality assessment, the TROPOMI CO data product was compared with the near-real-time data analysis of the ECMWF Integrated Forecasting System (IFS) assimilating IASI and MOPITT observations of CO 51 [RD83], which are provided by the Copernicus Atmosphere Monitoring Service (CAMS). For this purpose, Borsdorff et al. [RD81] collocated the TROPOMI CO retrieval with the 6 hourly CAMS CO fields, interpolating the CAMS data to the time and location of the individual TROPOMI measurement. Subsequently, the integration of the CAMS profile using the column averaging kernel provides the corresponding column density, which takes into account the vertical retrieval sensitivity. This approach allows a one-to-one comparison of CAMS and TROPOMI data, shown in Fig. 42. Overall, the TROPOMI and CAMS CO fields agree well. Figure 43 shows the corresponding histogram of the differences with small mean difference of +3.2 % between the TROPOMI and CAMS CO data with a standard deviation of 5.5 %. Both data sets are strongly correlated

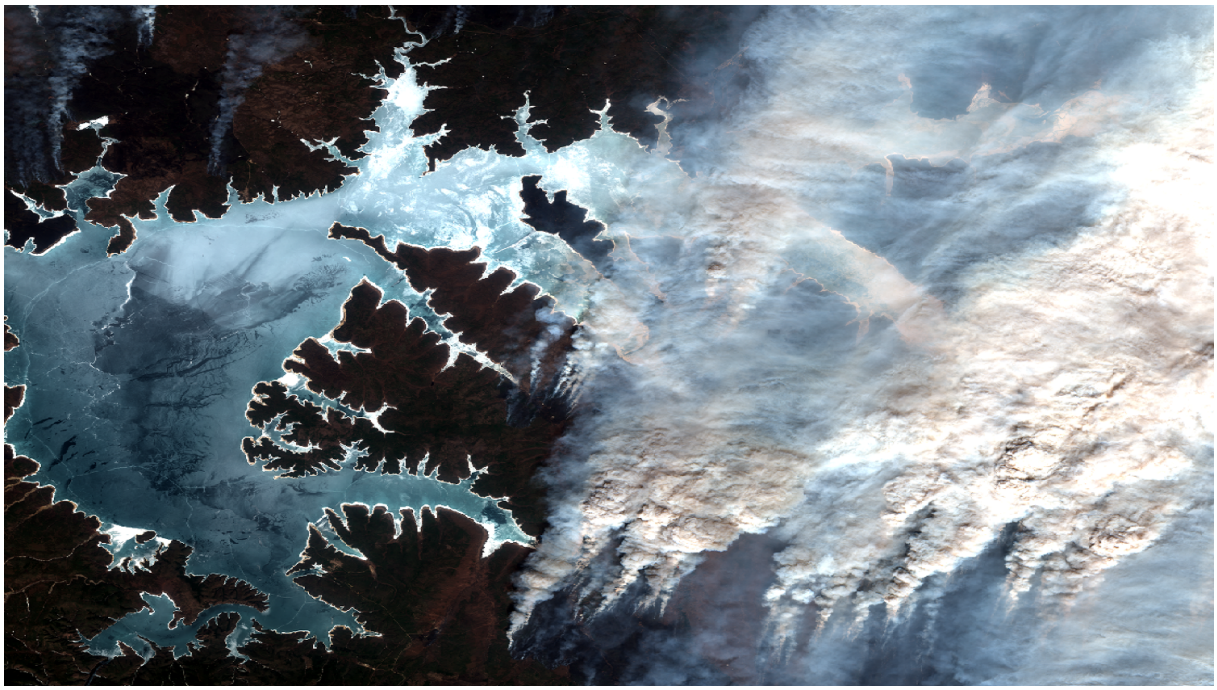

Figure 39: Copernicus Sentinel-2B image of wild fires in East Russia around Komsomolsk-on-Amur at May 9th, 2018.

with a Pearson correlation coefficient of 0.97. Also interesting is the good agreement of the two data sets over the oceans. Here, data can only be inferred from cloudy observations. In the shortwave infrared spectral range, the ocean surface is very dark (except for glint observation geometry) and so cloud-free measurements generally do not record sufficient light to achieve a meaningful retrieval. The good agreement gives confidence in valuable TROPOMI CO data product for cloudy conditions.

In addition to this overall agreement, we can also identify clear differences between CAMS and TROPOMI CO fields. Figure 41 shows the temporal evolution of the CO concentration above New Delhi (top panel), Kabul (middle panel), and Oqoltin in Uzbekistan (lower panel) during the days November 13th -19th. The TROPOMI CO columns (pink) and collocated CAMS fields (blue) are averaged in a 60km radius around the locations. The enhanced CO concentration above New Delhi shows a steady decline with time. The northeastward transport of the pollution from India strongly increases the CO concentration above Kabul and Oqoltin on November 14th and 15th. Both TROPOMI and CAMS indicate that the CO concentration above Kabul is back at normal levels on November 19. However, for Oqoltin TROPOMI shows that still some pollution is trapped between the mountains of the Ulugnor District on November 19 which is not reflected by CAMS. These differences may provide useful information to improve the CAMS forecast services as foreseen as one of the major objectives of the Sentinel 5 Precursor mission. Overall, the striking agreement with the CAMS analysis data but also the observed differences indicate new exciting application possibilities for the TROPOMI CO data product.

The quality of the TROPOMI CO data product needs to be validated in more detail using independent on-ground reference observations both for clear-sky and cloudy TROPOMI measurements. To this end, Borsdorff et al., 2018, [RD82] performed a first validation with CO observations at nine ground-based FTS stations operated by the TCCON network (see Table 9). The study selected sites in both the northern and southern hemisphere at low and high elevation on the continents and islands [RD84, RD85, RD86, RD87, RD88, RD89, RD90, RD91]. For the comparison, the authors used TROPOMI observations co-located with the TCCON sites by selecting all TROPOMI retrievals from the same day within a radius of 50 km around each station. The retrieved CO column of TROPOMI is adapted to the altitude of the station by either cutting off the scaled mixing ratios profile at the station altitude or extending it assuming a constant elongation of the mixing ratio to lower altitude. Finally, the daily averages of the dry air mixing ratios (XCO) is calculated using the adapted TROPOMI retrievals and the TCCON measurements.

Figure 44 shows an example for the TCCON sites at Ascension Island and Reunion. Data gaps in the TROPOMI time series are partly caused by disregarding observations with high clouds (z_{cld}>5000m) but also due to observation time reserved for in-orbit instrument characterization during the instrument commissioning

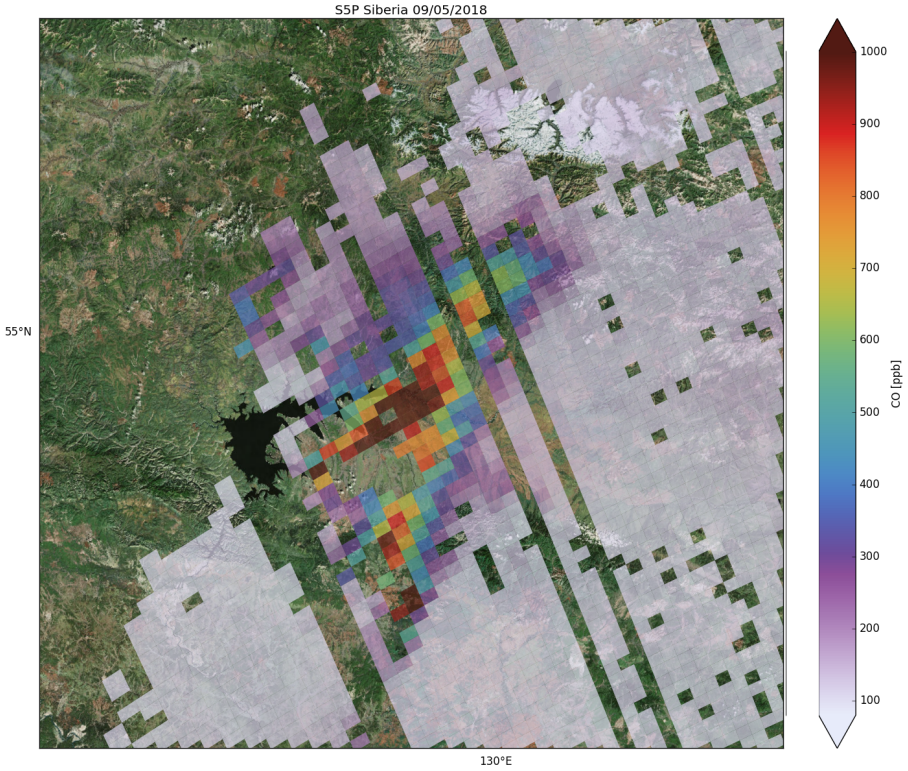

Figure 40: Sentinel 5 Precusor/TROPOMI CO total column mixing ratio for May 9th, 2018

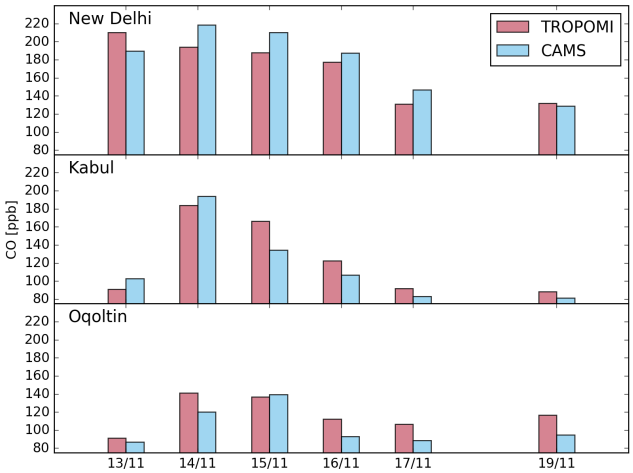

Figure 41: CO total column mixing ratios over New Delhi, Kabul, and Oqoltin. The data is averaged in a radius of 60 km around the locations. TROPOMI retrievals (pink) and CAMS data (blue) are shown for the individual days November 13th-15th and 19th.

phase. Figure 45 depicts the corresponding bias for each TCCON station for clear-sky and cloudy-sky conditions and the combination of both, as well as the standard deviation and the number of coincident daily mean values of TROPOMI and TCCON.

With the limited data available at the time of their study, Borsdorff et al. found good agreement with a small mean bias of TROPOMI CO versus TCCON of 6.0 ppb for clear-sky, 6.2 ppb for cloudy-sky TROPOMI retrievals and 5.8 ppb for the combination of both with a station-to-station deviation of 3.9 ppb for clear-sky, 2.4 ppb for cloudy-sky, and 2.9 ppb for the combination case. Furthermore, the mean standard deviation of the bias is 3.9 ppb for clear-sky, 2.4 ppb for cloud-sky, and 2.9 ppb for the combination. The good agreement

Table 9: Ground-based FTIR stations used for validation. The latitude and longitude values are given in degrees, the surface elevation in km.

Name	Latitude	Longitude	Altitude	Type
Karlsruhe	49.10	8.44	0.11	TCCON
Garmisch	47.48	11.06	0.75	TCCON
Zugspitze	47.42	10.98	2.96	TCCON
JPL	34.20	−118.18	0.39	TCCON
Caltech	34.14	−118.13	0.24	TCCON
Izaña	28.31	−16.50	2.37	TCCON
Mexico City	19.33	−99.18	2.26	Bruker Vertex 80
Ascension Island	−7.92	−14.33	0.03	TCCON
Réunion	−20.90	55.49	0.09	TCCON
Lauder	−45.04	169.68	0.37	TCCON

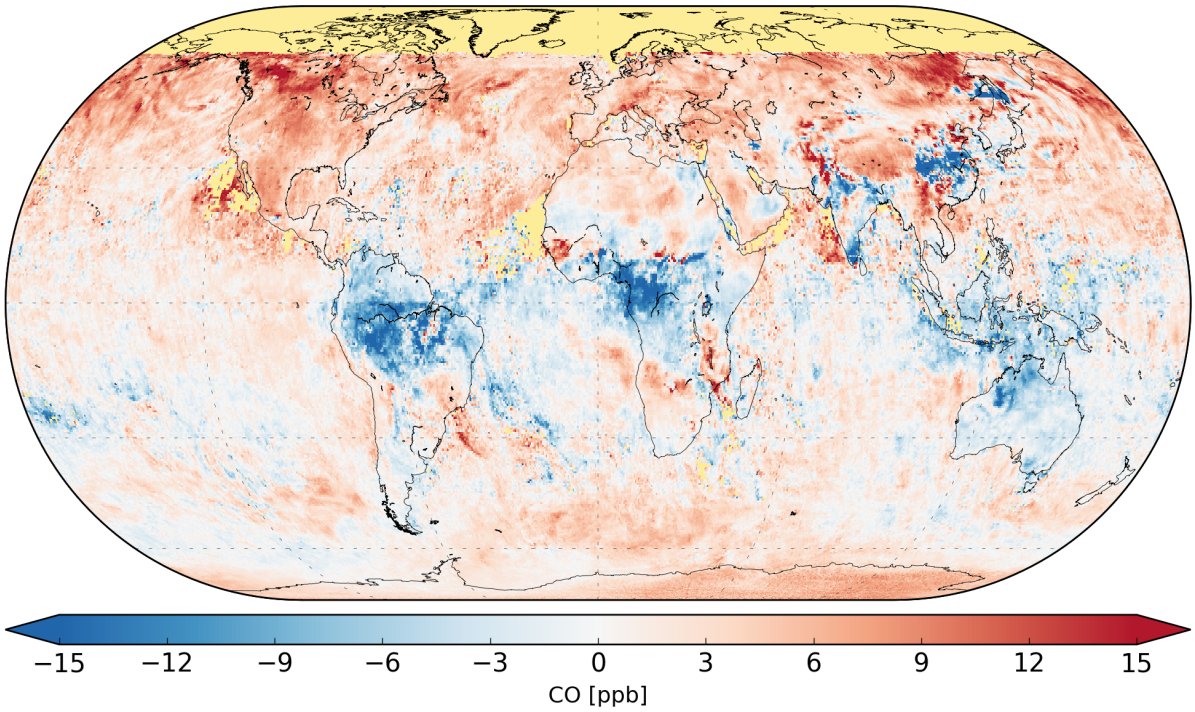

Figure 42: Differences of CO total column mixing ratios (TROPOMI - CAMS) averaged over the same time period as shown in Fig. 37 (from Borsdorff et al. 2018)

between clear-sky and cloudy-sky retrieval underlines the validity of the data retrieval for cloudy scenes, a key aspect of the SICOR algorithm to achieve the data coverage of the TROPOMI CO product.

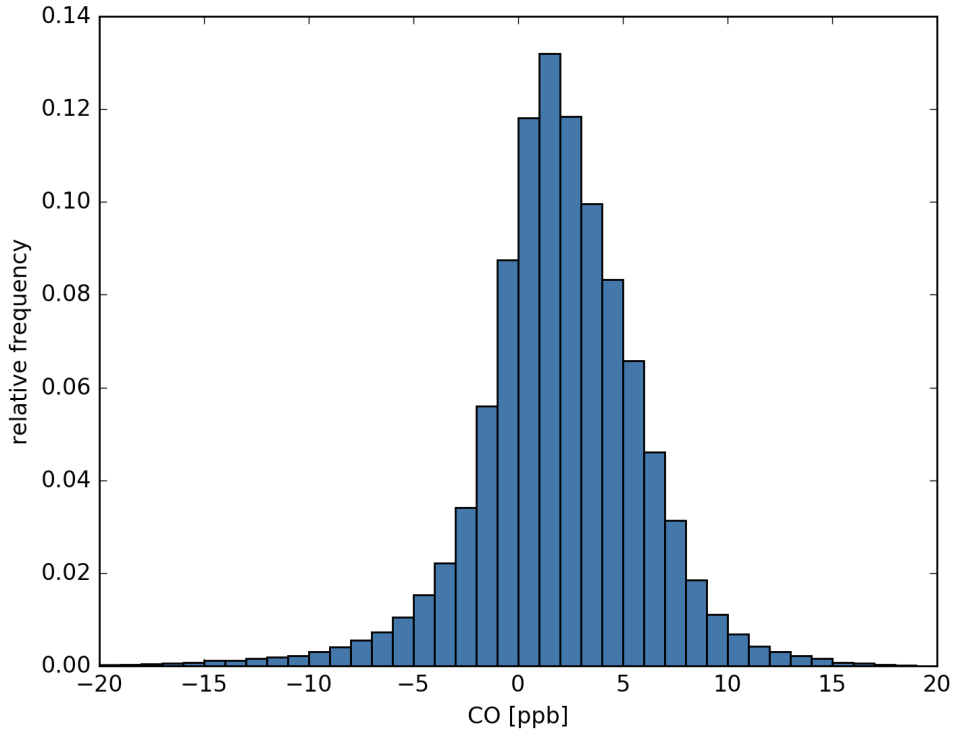

Figure 43: Histogram of the differences shown in Fig. 42 (from Borsdorff et al. 2018)

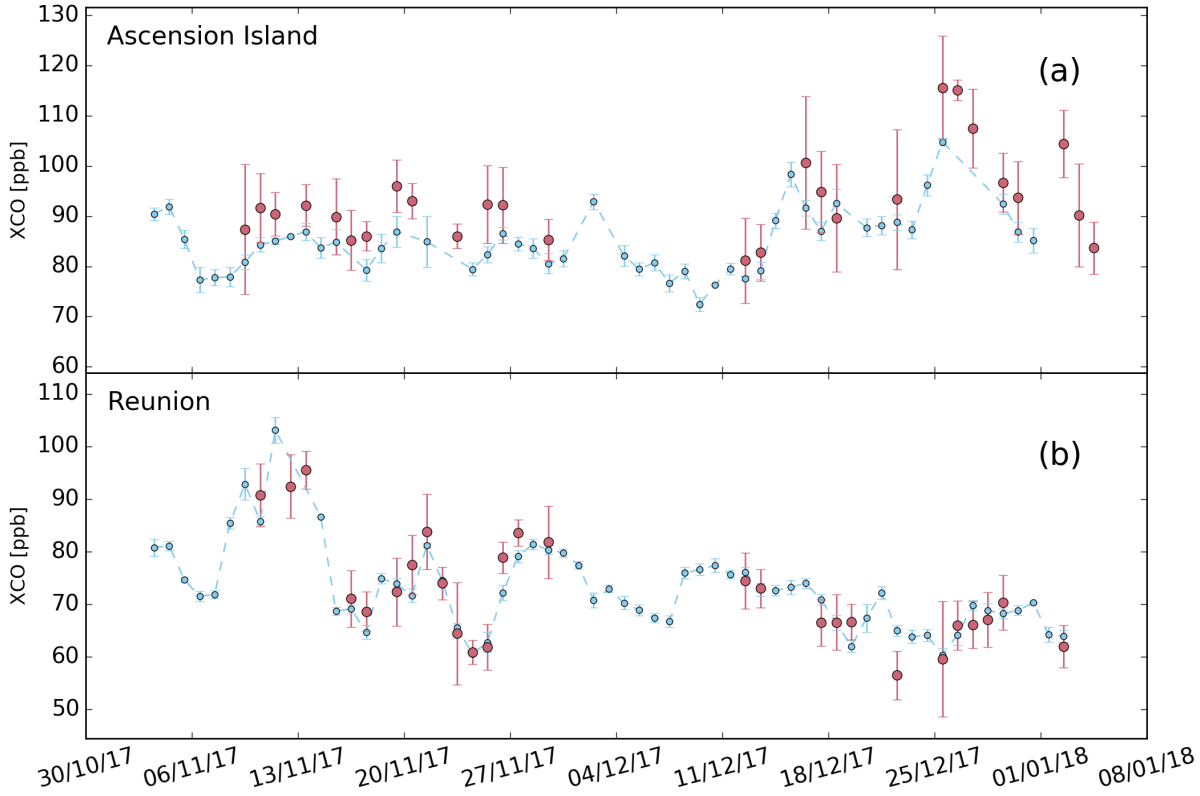

Figure 44: Daily means of dry air column mixing ratios (XCO) measured by TROPOMI (pink) and TCCON stations (blue) for (a) Ascension Island and (b) Reunion. A co-location radius of 50 km is used. The standard deviation of individual retrievals within a day is shown as error bar (from Borsdorff et al. 2018)

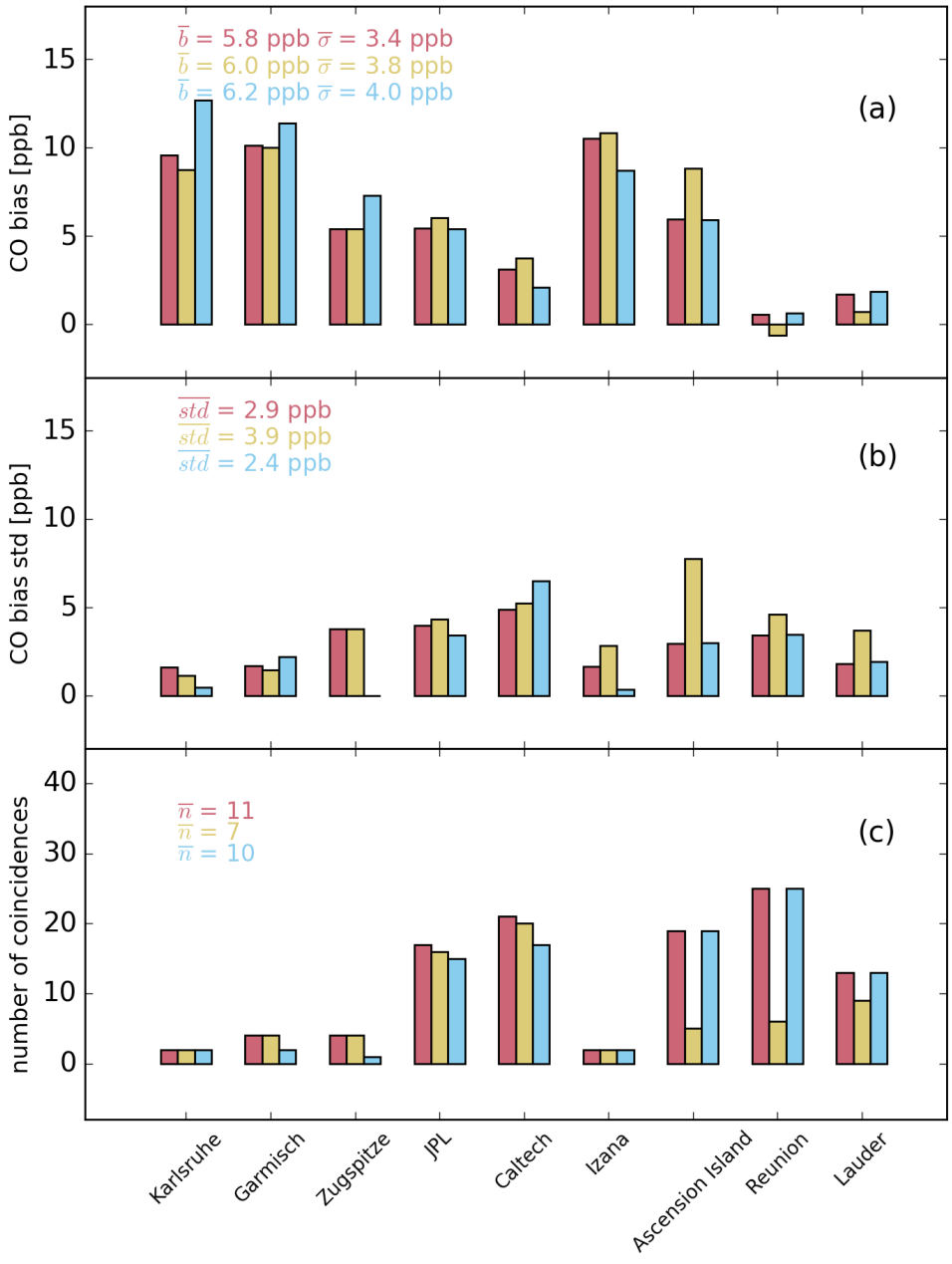

Figure 45: Mean bias (TROPOMI - TCCON) between co-located daily mean XCO values of TROPOMI and TCCON (a), the standard deviation of the bias (b), and the number of coincident daily mean pairs (c). $\bar{b}$ is the global mean bias (average of all station biases) and $\bar{\sigma}$ its station-to-station . $\bar{\text{std}}$ is the average of all standard deviations and $\bar{\text{n}}$ the average number of coincident pairs. TROPOMI retrievals under clear-sky (yellow), cloudy-sky (blue) and the combination of both (pink) are distinguished. (from Borsdorff et al. 2018)

12 Conclusion

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

The SICOR algorithm relies on a two-step retrieval from TROPOMI SWIR measurements. In the first step as part of the SWIR pre-processing, a non-scattering retrieval of the total amount of CH_4 in the spectral range 2315-2324 nm is performed to filter on the presence of high and optically thick clouds. In a second step, the CO column is inferred from SWIR measurements of an adjacent spectral window 2324-2338 nm. Here, a priori knowledge on the atmospheric methane abundance is used to retrieve effective cloud parameters from methane absorption bands simultaneously with atmospheric CO and H_2O abundances. The algorithm uses two highly efficient numerical modules. The inversion infers the trace gas column information by a least squares fitting of a scaling of a CO reference profile. From a theoretical point of view, the approach is identical to a regularized CO profile retrieval using Tikhonov regularization of first order and infinite regularization strength. This analogy provides us an analytical approach to calculate the column averaging kernel, which describes the sensitivity of the retrieved column to changes in the corresponding trace gas profile as a function of altitude. Moreover, the retrieval employs a two-stream radiative transfer model, which is linearized with respect to the parameters to be retrieved. The two-stream approximation is a simple approximation to account for multiple light scattering in atmospheric radiative transfer and its numerical implementation comes along with low computational cost.

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

An extensive performance analysis demonstrated a robust algorithm with stable numerical implementation, where we considered both erroneous atmospheric input fields and instrument malfunctions. The retrieval performance was analyzed for 6 generic measurement ensembles, where solar geometry, surface albedo and aerosol, cirrus and cloud parameters were varied systematically over a realistic parameter domain. Furthermore, we investigated the performance for an ensemble of about 26 000 simulated measurements over China, combining different data sources that describe the observed scenes as realistically as possible. For this purpose, we used MODIS cloud, aerosol and surface information, ECMWF pressure, water and temperature and CO and CH_4 profiles, simulated with the chemical transport models TM4 and CHIMERE. Measurements are simulated with the S-LINTRAN version 2.0 radiative transfer model, which accounts for multiple scattering by water clouds and cirrus. Additionally, we considered the robustness of the algorithm with respect to serval potential malfunction of the instrument, i.e. erroneous instrument spectral response, spectral calibration, a constant radiometric offset and a multiplicative radiometric error. Overall, the analysis showed a very stable numerical performance of the SICOR CO retrieval and the accuracy and precision of the CO data product is fully compliant with the level 2 user requirements. In this context, we identified the a priori knowledge of the atmospheric methane abundance as the most critical input parameter and to keep the corresponding CO retrieval bias within 3 %, the vertically integrated column density of CH_4 must be known with an accuracy of also 3 %. We showed that the forecast skills of present state-of-the-art chemical transport models are sufficient to predict methane with the required accuracy.

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

1. Relative profile of CO and H_2O.
2. Temperature-pressure profile.
3. Surface pressure on SWIR pixel level.
4. Accurate estimate of atmospheric methane.
5. Geo-potential height of the different input profiles.

6. First guess of the total amount of CO and H_2O.

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

Based on the analysis of the CO offline data product during the 6 month Sentinel 5 Precursor commissioning phase lasting till April 2018, we conclude that the mission ground segment provides a reliable high-quality CO data product, detecting the distribution of CO on global, regional and city scales with daily global coverage. Preliminary validations with the CAMS CO data product of ECMWF and ground based TCCON observations indicated a CO data quality already complaint with the mission requirements. Due to the high radiometric performance combined with the daily global coverage, the TROPOMI SWIR spectrometer opens doors for new research of natural and man-made emission of CO and air quality monitoring, one of the missions main objectives.

Acknowledgements

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

A Appendix: Flux method PIFM

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

$$\frac{dF^{\uparrow}}{d\tau} = \alpha_{1,n}F^{\uparrow} - \alpha_{2,n}F^{\downarrow} - \alpha_{3,n}\frac{S}{\mu_o} \tag{76}$$

$$\frac{dF^{\downarrow}}{d\tau} = \alpha_{2,n}F^{\uparrow} - \alpha_{1,n}F^{\downarrow} - \alpha_{4,n}\frac{S}{\mu_o} \tag{77}$$

$$\frac{dS}{d\tau} = -(1-\omega_n f_n)\frac{S}{\mu_o} \tag{78}$$

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

$$\alpha_{1,n} = U_{\uparrow}(1-\omega_n(1-\bar{\beta}_n)) \tag{79}$$

$$\alpha_{2,n} = U_{\downarrow}\bar{\beta}_n\omega_n \tag{80}$$

$$\alpha_{3,n} = (1-f_n)\omega_n\beta_n(\mu_o) \tag{81}$$

$$\alpha_{4,n} = (1-f_n)\omega_n(1-\beta_n(\mu_o)) \tag{82}$$

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

For a N-layer atmosphere, the general solution of the two-stream model of Zdunkowski et al. [RD36] can be expressed by a linear combination of the internal boundary conditions

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

(see also Eq. (6)) with coefficients

$$a_{1,n} = \exp\left[-\frac{(1-\tilde{\omega}_n f_n)\Delta\tau_n}{\mu_o}\right] \tag{83}$$

$$a_{2,n} = -a_{4,n}\gamma_{2,n} - a_{5,n}\gamma_{1,n}a_{1,n} + \gamma_{2,n}a_{1,n} \tag{84}$$

$$a_{3,n} = -a_{5,n}\gamma_{2,n} - a_{4,n}\gamma_{1,n}a_{1,n} + \gamma_{1,n} \tag{85}$$

$$a_{4,n} = E_n\frac{1-M_n^2}{1-E_n^2M_n^2} \tag{86}$$

$$a_{5,n} = M_n\frac{1-E_n^2}{1-E_n^2M_n^2} \tag{87}$$

and with

$$E_n = \exp(-\epsilon_n\Delta\tau_n);\;\; M_n = \frac{\alpha_{2,n}}{\alpha_{1,n}+\epsilon_n};\;\; \epsilon_n = \sqrt{\alpha_{1,n}^2+\alpha_{2,n}^2} \tag{88}$$

$$\gamma_{1,n} = \frac{(1-\omega_n f_n)\alpha_{3,n} - \mu_o(\alpha_{1,n}\alpha_{3,n}+\alpha_{2,n}\alpha_{4,n})}{(1-\omega_n f_n)^2 - \epsilon_n^2\mu_o^2} \tag{89}$$

$$\gamma_{2,n} = \frac{-(1-\omega_n f_n)\alpha_{4,n} - \mu_o(\alpha_{1,n}\alpha_{4,n}+\alpha_{2,n}\alpha_{3,n})}{(1-\omega_n f_n)^2 - \epsilon_n^2\mu_o^2} \tag{90}$$

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

The numerical effort to calculate the coefficients $\bar{\beta}$ and $\beta(\mu_o)$ in Eq. (16) and (15) and the fraction f scattered in forward direction can be significant, and so different approaches are suggested to estimate this qualities in a simplified manner. Zdunkowski et al. [RD36] proposed the following approximations:

$$\bar{\beta}_n = \frac{3 - p_{1,n}}{8} \tag{91}$$

$$\beta_n(\mu_o) = \frac{1}{2} - \frac{\mu_o}{4}\frac{p_{1,n} - 3f_n}{1 - f_n} \tag{92}$$

$$f_n = \frac{p_{1,n}^2}{3^2} \tag{93}$$

with the diffusivity factors

$$U_\downarrow = U_\uparrow = U = 2 \tag{94}$$

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

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

$$\frac{\partial\alpha_{1,n}}{\partial\omega_n} = -U(1 - \bar{\beta}_n) \tag{95}$$

$$\frac{\partial\alpha_{2,n}}{\partial\omega_n} = -U\bar{\beta}_n \tag{96}$$

$$\frac{\partial\alpha_{3,n}}{\partial\omega_n} = (1 - f_n)\beta_n(\mu_o) \tag{97}$$

$$\frac{\partial\alpha_{4,n}}{\partial\omega_n} = (1 - f_n)(1 - \beta_n(\mu_o)) \tag{98}$$

Furthermore,

$$\frac{\partial\epsilon_n}{\partial\omega_n} = \frac{1}{\epsilon_n}\left(\alpha'_{1,n}\alpha_{1,n} - \alpha'_{2,n}\alpha_{2,n}\right) \tag{99}$$

$$\frac{\partial E_n}{\partial\omega_n} = \epsilon'_n\tau_n E_n \tag{100}$$

$$\frac{\partial M_n}{\partial\omega_n} = \frac{\alpha'_{2,n}(\alpha_{1,n} + \epsilon_n) - \alpha_{2,n}(\alpha'_{1,n} + \epsilon'_n)}{(\alpha_{1,n} + \epsilon_n)^2} \tag{101}$$

and

$$\frac{\partial\gamma_{1,n}}{\partial\omega_n} = \frac{-f_n\alpha_{3,n} + (1 - \omega_n f_n)\alpha'_{3,n} - \mu_o(\alpha'_{1,n}\alpha_{3,n} + \alpha_{1,n}\alpha'_{3,n} + \alpha'_{2,n}\alpha_{4,n} + \alpha_{2,n}\alpha'_{4,n})((1 - \omega_n f_n)^2 - \epsilon_n^2\mu_o^2)}{((1 - \omega_n f_n)^2 - \epsilon_n^2\mu_o^2)^2} + \frac{[2f_n(1 - \omega_n f_n) + 2\epsilon_n\epsilon'_n\mu_o^2]((1 - \omega_n f_n)\alpha_{3,n} - \mu_o(\alpha_{1,n}\alpha_{3,n} + \alpha_{2,n}\alpha_{4,n}))}{((1 - \omega_n f_n)^2 - \epsilon_n^2\mu_o^2)^2} \tag{102}$$

$$\frac{\partial\gamma_{2,n}}{\partial\omega_n} = \frac{f_n\alpha_{4,n} - (1 - \omega_n f_n)\alpha'_{4,n} - \mu_o(\alpha'_{1,n}\alpha_{4,n} + \alpha_{1,n}\alpha'_{4,n} + \alpha'_{2,n}\alpha_{3,n} + \alpha_{2,n}\alpha'_{3,n})((1 - \omega_n f_n)^2 - \epsilon_n^2\mu_o^2)}{((1 - \omega_n f_n)^2 - \epsilon_n^2\mu_o^2)^2} + \frac{[2f_n(1 - \omega_n f_n) + 2\epsilon_n\epsilon'_n\mu_o^2]((1 - \omega_n f_n)\alpha_{4,n} - \mu_o(\alpha_{1,n}\alpha_{4,n} + \alpha_{2,n}\alpha_{3,n}))}{((1 - \omega_n f_n)^2 - \epsilon_n^2\mu_o^2)^2} \tag{103}$$

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

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

$$\frac{\partial \alpha_{1,n}}{\partial \Delta\tau_n} = \frac{\partial \alpha_{2,n}}{\partial \Delta\tau_n} = \frac{\partial \alpha_{3,n}}{\partial \Delta\tau_n} = \frac{\partial \alpha_{4,n}}{\partial \Delta\tau_n} = 0 \tag{104}$$

$$\frac{\partial \epsilon_n}{\partial \Delta\tau_n} = \frac{\partial M_n}{\partial \Delta\tau_n} = 0 \tag{105}$$

$$\frac{\partial E_n}{\partial \Delta\tau_n} = -\epsilon_n E_n \tag{106}$$

$$\frac{\partial \gamma_{1,n}}{\partial \Delta\tau_n} = \frac{\partial \gamma_{2,n}}{\partial \Delta\tau_n} = 0\,. \tag{107}$$

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

$$\frac{\partial \alpha_{1,n}}{\partial p_{1,n}} = \frac{\partial \alpha_{2,n}}{\partial p_{1,n}} = -\frac{U\omega_n}{8} \tag{108}$$

$$\frac{\partial \alpha_{3,n}}{\partial p_{1,n}} = \omega_n \left[\beta'_n(\mu_o)(1-f_n) - \beta_n(\mu_o)\frac{2p_{1,n}}{9} \right] \tag{109}$$

$$\frac{\partial \alpha_{4,n}}{\partial p_{1,n}} = -\omega_n \left[\beta'_n(\mu_o)(1-f_n) + (1-\beta_n(\mu_o))\frac{2p_{1,n}}{9} \right] \tag{110}$$

with

$$\beta'_n(\mu_o) = \frac{\partial \beta_n(\mu_o)}{\partial p_{1,n}} = -\frac{\mu_o}{4}\left[\frac{1-\frac{2p_{1,n}}{3}}{1-f_n} + \frac{p_{1,n}-3f_n}{(1-f_n)^2}\frac{2p_{1,n}}{9} \right] \tag{111}$$

and

$$\frac{\partial \epsilon_n}{\partial p_{1,n}} = -\frac{U\omega_n}{8\epsilon_n}(\alpha_{1,n}-\alpha_{2,n}) \tag{112}$$

$$\frac{\partial E_n}{\partial p_{1,n}} = \epsilon'_n E_n \tag{113}$$

$$\frac{\partial M_n}{\partial p_{1,n}} = \frac{\alpha'_{2,n}(\alpha_{1,n}+\epsilon_n) + \alpha_{2,n}(\alpha'_{1,n}+\epsilon'_n)}{(\alpha_{1,n}+\epsilon_n)^2} \tag{114}$$

$$\frac{\partial \gamma_1}{\partial p_{1,n}} = \frac{(1-\omega_n f_n)\alpha_{3,n} - \omega\alpha_{3,n}\frac{2}{9}p_{1,n} - \mu_o(\alpha_{1,n}\alpha'_{3,n} + \alpha'_{1,n}\alpha_{3,n} + \alpha_{2,n}\alpha'_{4,n} + \alpha'_{2,n}\alpha_{4,n})}{((1-\omega_n f_n)^2 - \epsilon_n^2\mu_o^2)^2} \tag{115}$$

$$\frac{\partial \gamma_2}{\partial p_{1,n}} = \frac{(1-\omega_n f_n)\alpha_{4,n} - \omega_n\alpha_{4,n}\frac{2}{9}p_{1,n} - \mu_o(\alpha_{1,n}\alpha'_{4,n} + \alpha'_{1,n}\alpha_{4,n} + \alpha_{2,n}\alpha'_{3,n} + \alpha'_{2,n}\alpha_{3,n})}{((1-\omega_n f_n)^2 - \epsilon_n^2\mu_o^2)^2}\,. \tag{116}$$

With Eq. 95-116, we can calculate the derivatives of the matrix elements $a_{1,n}$, $a_{2,n}$, $a_{3,n}$, $a_{4,n}$ and $a_{5,n}$ in Eq. (83)- (87):

$$\frac{\partial a_{1,n}}{\partial \omega_n} = \frac{f_n\Delta\tau_n}{\mu_o}a_{1,n} \tag{117}$$

$$\frac{\partial a_{2,n}}{\partial \omega_n} = -a'_{4,n}\gamma_{2,n} - a_{4,n}\gamma'_{2,n} - a'_{5,n}\gamma_{1,n}a_{1,n} - a_{5,n}\gamma'_{1,n}a_{1,n} - a_{5,n}\gamma_{1,n}a'_{1,n} + \gamma_{2,n}a'_{1,n} + \gamma'_{2,n}a_{1,n} \tag{118}$$

$$\frac{\partial a_{3,n}}{\partial \omega_n} = -a'_{5,n}\gamma_{2,n} - a_{5,n}\gamma'_{2,n} - a'_{4,n}\gamma_{1,n}a_{1,n} - a_{4,n}\gamma'_{1,n}a_{1,n} - a_{4,n}\gamma_{1,n}a'_{1,n} + \gamma'_{1,n} \tag{119}$$

$$\frac{\partial a_{4,n}}{\partial \omega_n} = \frac{[E'_n(1-M_n^2) - 2E_nM'_nM_n](1-E_n^2M_2^2) + 2E_n(1-M_n^2)E_nM_n[E'_nM_n + E_nM'_n]}{(1-E_n^2M_n^2)^2} \tag{120}$$

$$\frac{\partial a_{5,n}}{\partial \omega_n} = \frac{[M'_n(1-E_n^2) - 2M_nE'_nE_n](1-E_n^2M_2^2) + 2M_n(1-E_n^2)E_nM_n[E'_nM_n + E_nM'_n]}{(1-E_n^2M_n^2)^2} \tag{121}$$

$$\frac{\partial a_{1,n}}{\partial \Delta\tau_n} = \frac{1-\omega_n f_n}{\mu_o} a_{1,n} \tag{122}$$

$$\frac{\partial a_{2,n}}{\partial \Delta\tau_n} = a'_{4,n}\gamma_{2,n} - a'_{5,n}\gamma_{1,n}a_{1,n} - a_{5,n}\gamma_{1,n}a'_{1,n} + \gamma_{2,n}a'_{1,n} \tag{123}$$

$$\frac{\partial a_{3,n}}{\partial \Delta\tau_n} = -a'_{5,n}\gamma_{2,n} - a'_{4,n}\gamma_{1,n}a_{1,n} - a_{4,n}\gamma_{1,n}a'_{1,n} \tag{124}$$

$$\frac{\partial a_{4,n}}{\partial \Delta\tau_n} = \frac{(1-M_n^2)E'_2(1+E_n^2M_n^2)}{(1-E_n^2M_n^2)^2} \tag{125}$$

$$\frac{\partial a_{5,n}}{\partial \Delta\tau_n} = \frac{2E_nE'nM_n(1-M_n^2)}{(1-E_n^2M_n^2)^2} \tag{126}$$

and

$$\frac{\partial a_{1,n}}{\partial p_{1,n}} = -\frac{2\omega_n\Delta\tau_n p_{1,n}}{9\mu_o} a_{1,n} \tag{127}$$

$$\frac{\partial a_{2,n}}{\partial p_{1,n}} = -a'_{4,n}\gamma_{2,n} - a'_{4,n}\gamma_{2,n} - a'_{5,n}\gamma_{1,n}a_{1,n} - a_{5,n}\gamma'_{1,n}a_{1,n} - a_{5,n}\gamma_{1,n}a'_{1,n} + \gamma_{2,n}a'_{1,n} + \gamma'_{2,n}a_{1,n} \tag{128}$$

$$\frac{\partial a_{3,n}}{\partial p_{1,n}} = -a'_{5,n}\gamma_{2,n} - a_{5,n}\gamma'_{2,n} - a'_{4,n}\gamma_{1,n}a_{1,n} - a_{4,n}\gamma'_{1,n}a_{1,n} - a_{4,n}\gamma_{1,n}a'_{1,n} + \gamma'_{1,n} \tag{129}$$

$$\frac{\partial a_{4,n}}{\partial p_{1,n}} = \frac{[E'_n(1-M_n^2) - 2E_nM'_nM_n](1-E_n^2M_2^2) + 2E_n(1-M_n^2)E_nM_n[E'_nM_n + E_nM'_n]}{(1-E_n^2M_n^2)^2} \tag{130}$$

$$\frac{\partial a_{5,n}}{\partial p_{1,n}} = \frac{[M'_n(1-E_n^2) - 2M_nE'_nE_n](1-E_n^2M_2^2) + 2M_n(1-E_n^2)E_nM_n[E'_nM_n + E_nM'_n]}{(1-E_n^2M_n^2)^2} \tag{131}$$

Finally to calculate the Jacobian in Eq. (39), we need the corresponding derivatives of the response vector

$$\mathbf{R} = \frac{1}{4\pi}(u_1, v_1, w_1, \cdots, u_{N+1}, v_{N+1}, w_{N+1}) \tag{132}$$

with respect to the atmospheric parameters $\Delta\tau_n$, ω_n, β_n and the phase function P_n. Using Eq. (26), we obtain:

$$\begin{aligned}
\frac{\partial u_n}{\partial \omega_m} &= \delta_{nm}P_nt_n(\tilde{\mu}) \\
\frac{\partial v_n}{\partial \omega_m} &= U^{\uparrow}\left[\delta_{nm}(1-\beta_n)t_n(\mu_v) + \delta_{n-1,m}(1-\beta_{n-1})t_{n-1}(\mu_v)\right] \\
\frac{\partial w_n}{\partial \omega_m} &= U^{\downarrow}\left[\delta_{n,m}\beta_nt_n(\mu_v) + \delta_{n-1,m}\beta_{n-1}t_{n-1}(\mu_v)\right]
\end{aligned} \tag{133}$$

and

$$\begin{aligned}
\frac{\partial u_n}{\partial P_m} &= \delta_{n,m}\omega_nt_n(\tilde{\mu}) \\
\frac{\partial v_n}{\partial P_m} &= 0 \\
\frac{\partial w_n}{\partial P_m} &= 0
\end{aligned} \tag{134}$$

and

$$\begin{aligned}
\frac{\partial u_n}{\partial \beta_m} &= 0 \\
\frac{\partial v_n}{\partial \beta_m} &= -U^{\uparrow}\left[\delta_{n,m}\omega_nt_n(\mu_v) + \delta_{n-1,m}\omega_{n-1}t_{n-1}(\mu_v)\right] \\
\frac{\partial w_n}{\partial \beta_m} &= U^{\downarrow}\left[\delta_{n,m}\omega_nt_n(\mu_v) + \delta_{n-1,m}\omega_{n-1}t_{n-1}(\mu_v)\right]
\end{aligned} \tag{135}$$

and

$$\begin{aligned}
\frac{\partial u_n}{\partial \Delta\tau_m} &= \frac{1}{4\pi\mu_o}\omega_n P_n t'_{n,m}(\tilde{\mu}) \\
\frac{\partial v_n}{\partial \Delta\tau_m} &= \frac{U^{\uparrow}}{4\pi}\left[\omega_n(1-\beta_n)t'_{n,m}(\mu_v) + \omega_{n-1}(1-\beta_{n-1})t'_{n-1,m}(\mu_v)\right] \\
\frac{\partial w_n}{\partial \Delta\tau_m} &= \frac{U^{\downarrow}}{4\pi}\left[\omega_n\beta_n t'_{n,m}(\mu_v) + \omega_{n-1}\beta_{n-1}t'_{n-1,m}(\mu_v)\right]
\end{aligned} \tag{136}$$

with the derivative

$$t'_{n,m}(\mu) = \frac{\partial t_n(\mu)}{\partial \Delta\tau_m} = \begin{cases} -\frac{1}{\mu_v}t_n(\mu) & \text{for } m < n \\ \frac{1}{\mu}\exp(-\frac{\tau_{n-1}}{\mu_v})\exp(-\frac{\Delta\tau_n}{\mu})] & \text{for } m = n = 1,\cdots,N \\ 0 & \text{for } m > n \end{cases} \tag{137}$$

of the auxiliary function t_n in Eq. (23).

B Appendix: SWIR Pre-Processing

To optimise synergies between the different SWIR retrievals, i.e. the CO and HDO/H_2O retrieval described in this document and the CH_4 retrieval as described by [RD2], a SWIR pre-processing module is developed. Exploiting spectral information from the SWIR spectral ranges, the module provides the following auxiliary products:

- A cloud filter based on a non-scattering retrieval of the total amount of CH_4 and a corresponding column estimate from the TM5 forecast (CH_4 a priori cloud filter)
- A cloud filter using a non-scattering retrieval of the total amount of CH_4 in a strong and weak methane absorption band (CH_4 two-band cloud filter)
- A cloud filter using a non-scattering retrieval of the total amount of H_2O in a strong and weak water absorption band (H_2O two-band cloud filter)

Technically, the pre-processing module is implemented in the CO prototype software to reduce interfaces with the processing framework as indicated in Fig. 2. However, functionally the module does not depend on the CO retrieval. The overall algorithm structure is summarised in Fig. 46. Via the CO interface to the processing framework, the TROPOMI measurements in the SWIR spectral range are allocated in combination with the meteo fields for temperature, pressure and water vapour abundance, information on the pixel's orography and the model forecast of the CH_4 and CO profiles. The different auxiliary products are processed in a hierarchical manner to optimise the computational effort. In a first step, only measurements are accepted with a continuum LER value exceeding a certain threshold. Due to this, only measurements are processed with an acceptable signal level. For example, clear sky ocean measurements are rejected by this inquiry due to the very low ocean surface reflection in the SWIR spectral range. Subsequently, the CH_4 a priori filter is processed. A comparison

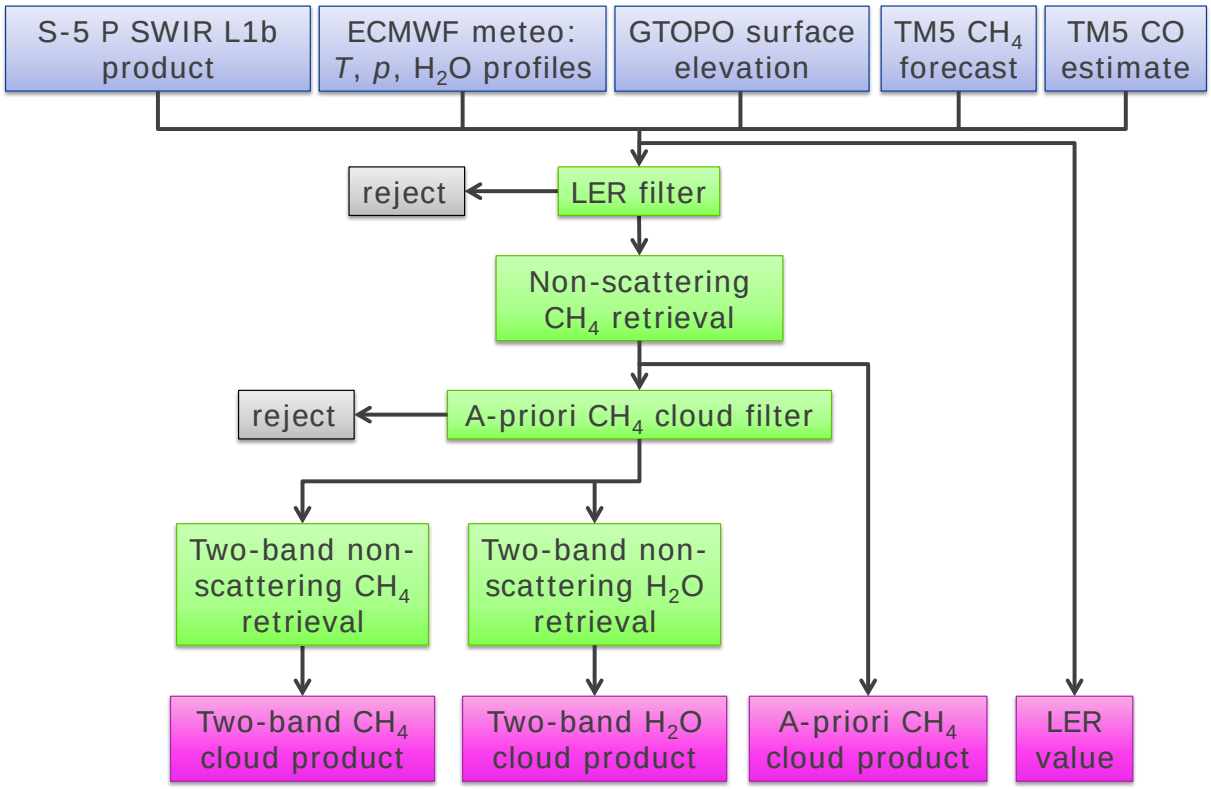

Figure 46: Overall structure of the SWIR preprocessor.

of a non-scattering methane column retrieval with a model a priori methane estimate is used to reject spectral observations contaminated by high and optically thick clouds. This cloud filter is required for the SICOR CO retrieval, which is designed to process clear sky SWIR observation and measurements of scenes with low water clouds and optically thin cirrus and aerosol loads (see Sec. 5). For the processing of CH_4 and HDO/H_2O an even stricter cloud screening is required. Therefore, the remaining auxiliary products, viz. the CH_4 and H_2O two-band cloud filters, are only processed for observations which pass the CH_4 a priori filter. The retrieval of the auxiliary products are described in more detail below.

Non-scattering total column retrieval

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

$$I^{\mathrm{TOA}}(\lambda) = A_s(\lambda)\frac{\mu_o F_o}{\pi}\exp(-\frac{1}{\tilde{\mu}}\tau_{\mathrm{tot}}(\lambda))\,, \tag{138}$$

where A_s is the surface albedo, $\mu_o = \cos(\Theta_o)$ with the solar zenith angle Θ_o. This cosine is corrected for the sphericity of the Earth according to Kasten and Young (1989) [RD92]. F_o is the solar irradiance and

$$\frac{1}{\tilde{\mu}} = \frac{\mu_o + \mu_v}{\mu_o \mu_v} \tag{139}$$

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

$$\tau_{\mathrm{tot}}(\lambda) = \sum_k \int \sigma_k(z,\lambda)\,\rho_k(z)\,dz\,, \tag{140}$$

where z indicates the altitude, index k represents the relevant absorbers CO, CH_4 H_2O and HDO, $\rho_k(z)$ is the concentration of absorber k at altitude z, $\sigma_k(z,\lambda)$ is the corresponding wavelength-dependent absorption cross sections.

In the following, we assume that the relative profile

$$\rho_k^{\mathrm{rel}} = \frac{\rho_k}{c_k} \tag{141}$$

of absorber k is constant, where

$$c_k = \int \rho_k(z)dz \tag{142}$$

is the column density of this absorber. So,

$$\frac{\partial \tau_{\mathrm{tot}}}{\partial c_k} = \frac{1}{c_k}\int \sigma_k(z,\lambda)\,\rho_k(z)\,dz \tag{143}$$

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

$$\frac{\partial I^{\mathrm{TOA}}}{\partial c_k} = -\frac{I^{\mathrm{TOA}}}{\tilde{\mu} c_k}\int \sigma_k(z)\rho_k(z)dz\,. \tag{144}$$

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

$$\frac{\partial I^{\mathrm{TOA}}}{\partial A_s} = \frac{\mu_o F_o}{\pi} \exp(-\frac{1}{\tilde{\mu}} \tau_{\mathrm{tot}}) \, . \tag{145}$$

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

CH_4 a priori cloud filter

The methane a priori cloud filter relies on a non-scattering methane column retrieval. In the presence of high and optically thick clouds, the lightpath gets effectively shortened due to reflection of light by clouds. Thus assuming a clear sky model atmosphere, the retrieved methane total column is underestimated depending on cloud occurrence in the observed scene. Figure 47 shows the induced error in the non-scattering methane column retrieved from the spectral window 2315–2324 nm for two cloudy scenes. For a water cloud with an optical depth of 5, the methane error increases with cloud coverage and cloud height. In case of a cirrus at 10 km height, the non-scattering CH_4 column can be used to identify cloud cirrus contamination with optical depth > 0.5 for low and moderate surface reflection. However for bright surfaces, the lightpath shortening due the reflection of light by the cirrus is compensated by an enhancement of the lightpath because of multiple reflection of light between the cirrus layer and the surface. In this case, the non-scattering CH_4 column is less suited for cirrus detection. Keeping in mind that the current accuracy of the methane column forecast is in the order of 2–3 % using state-of-the-art chemical transport models (see Sec. 7.6), the CH_4 a priori cloud filter is a powerful tool to screen measurement with respect to the presence of high and optically thick clouds.

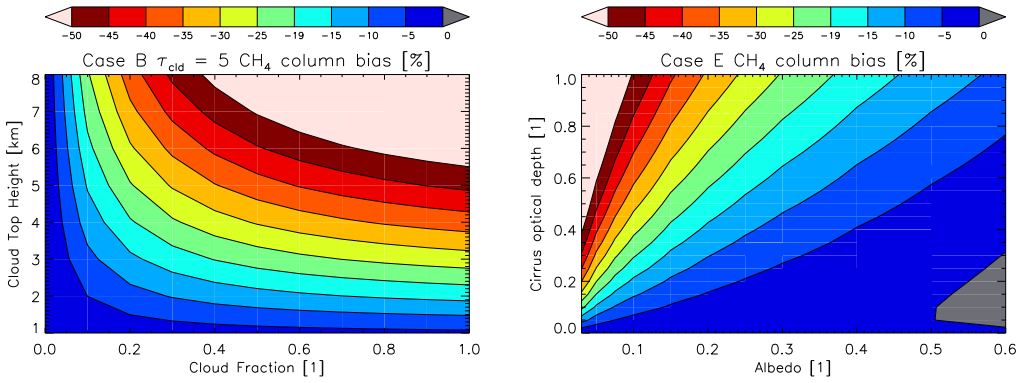

Figure 47: CH_4 error of a non-scattering retrieval from the SWIR 2315–2324 nm spectral window for a water cloud with optical thickness of 5 as function of cloud height and cloud fraction (left panel, for more details see generic scenario B in Sec. 7.1) and for a cirrus cloud at 10 km height as function of surface albedo and cirrus optical thickness (right panel, for more details see generic scenario E in Sec. 7.1).

CH_4 two-band cloud filter

Complementary to the CH_4 a priori filter, the SWIR pre-processor includes a cloud filter based on two methane non-scattering retrievals that utilise two different spectral bands with a strong and weak absorption band, respectively. This method relies on the fact that in a strong absorption band, photons along an enhanced light path contribute less to the total signal than in a weak absorption band. Thus, we expect that the non-scattering methane column, which are inferred from a weak absorption, exceeds the corresponding column retrieved from a strong absorption band. Moreover, the difference between the two columns indicates changes of the lightpath due to atmospheric scattering and so can be used as a cloud filter. In Fig. 48, this difference is depicted for the cloud scenarios of Fig. 47 using the strong CH_4 absorption at 2363-2373 nm and the weak absorption at 2310-2315 nm. The figure indicates that the methane two-band retrieval is well suited to detect cirrus cloud even above bright surfaces and complements the CH_4 a priori filter. Therefore, this cloud filter is particular valuable for cirrus screening needed by the CH_4 and HDO/H_2O data processing.

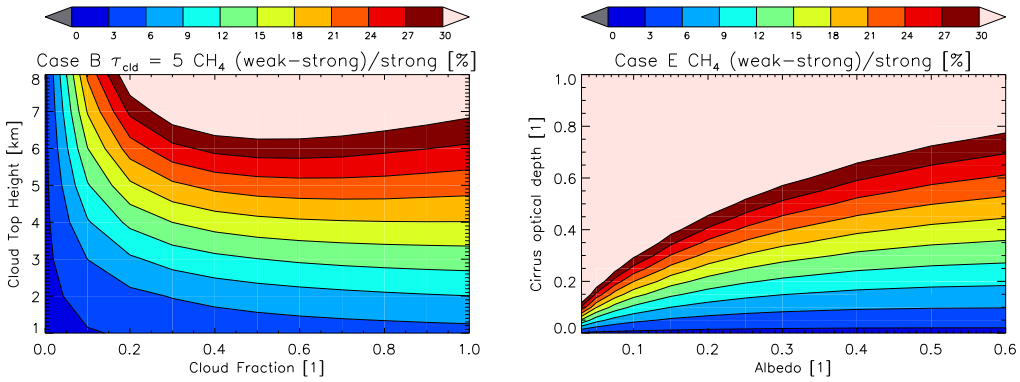

Figure 48: CH_4 two-band cloud filter for the cloud scenarios of Fig. 47. The methane cloud filter relies on non-scattering methane column retrieval from strong and weak absorption features at 2363-2373 nm and 2310-2315 nm, respectively.

H_2O two-band cloud filter

The concept of the two-band CH_4 cloud filter can be applied to H_2O as well. The two-band cloud filter does not rely on a-priori knowledge, so the poor a-priori knowledge of H_2O does not hamper the application of the two-band cloud filter for water. Compared to CH_4, H_2O absorption lines are narrower, resulting in a different sensitivity towards pressure broadening close to the surface. The H_2O two-band cloud filter is set up analogous to the CH_4 two-band cloud filter. Here, the window 2329-2334 nm contains weak H_2O absorption and the window 2367-2377 nm includes strong H_2O absorption features.

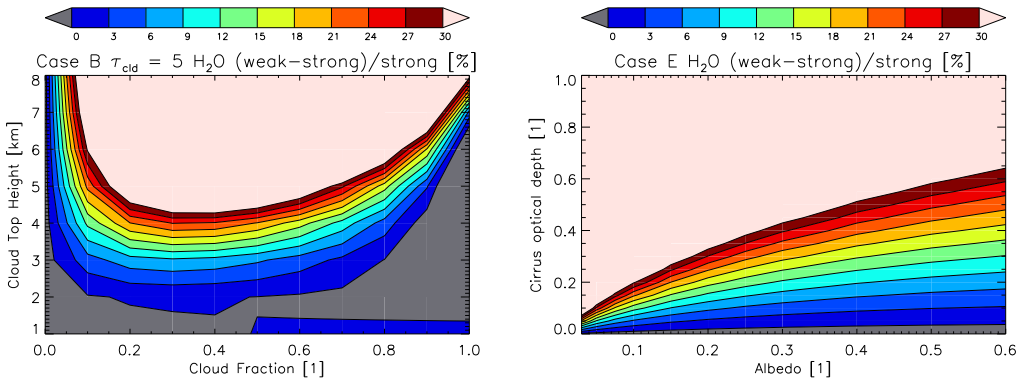

Figure 49: H_2O two-band cloud filter for the cloud scenarios of Fig. 47. The filter relies on non-scattering methane column retrieval from strong and weak absorption features at 2367-2377 nm and 2329-2334 nm, respectively.

Figure 49 shows the relative difference in the retrieved water columns using the weak and strong absorption bands. Compared to the CH_4 two-band cloud filter in Fig. 48, the H_2O two-band filter shows larger sensitivity to clouds and so can be used to compliment the other cloud filters.