BIRA·IASB aeronomie·be

S5P/TROPOMI SO_2 ATBD

sentinel-5p

document number	:	S5P-BIRA-L2-400E-ATBD
CI identification	:	CI-400E-ATBD
issue	:	2.2.0
date	:	2020 06 15
status	:	Update for version 2.1.x of the operational processor

Document approval record

	digital signature
prepared:	Nicolas Theys Digitally signed by Nicolas Theys Date: 2020.06.22 06:05:32 +02'00' Isabelle De Smedt Digitally signed by Isabelle De Smedt Date: 2020.06.22 06:05:59 +02'00' Christophe Lerot Digitally signed by Christophe Lerot Date: 2020.06.22 06:06:21 +02'00' Huan Yu Digitally signed by Huan Yu Date: 2020.06.22 06:06:50 +02'00' Michel Van Roozendael Digitally signed by Michel Van Roozendael Date: 2020.06.22 06:07:14 +02'00'
checked:	Thomas Wagner, Pascal Hedelt Dominik Brunner, Alfonso Saiz-Lopez
approved PI:	
approved PM:	

Document change record

issue	date	item	comments
0.0.1	2012-09-25	All	Initial draft version
0.0.2	2012-11-23		Changes after the first internal review
0.1.1	2013-06-04	All	CDR version for internal review
0.5.0	2013-06-21	All	CDR version
0.9.0	2013-11-28	All	Changes after the first external review
0.10.0	2014-04-15	All	Consolidated changes after the first external review
0.11.0	2014-10		Minor changes after June 2014 CDR Harmonized input/output tables between ATBDs Updated number of observations per orbit.
0.12.0	2014-12	All	Update according to L2CDR RIDs. Verification of symbol definition consistency, equation number consistency, algorithm description. Correction of several typo's layout issues.
0.13.0	2015-09-15	All	Minor changes. Update of altitude-resolved air mass factors LUT table. Update for limited release to S5P validation team
0.14.0	2015-12-11	All	Minor changes. Update of document for the instrumental binning scheme.
1.0.0	2016-01-30	All	Update after second internal review
1.1.0	2018-10-05	All	Description of algorithm updates for operational release (phase E2). Various improvements of the text.

2.0.0	219-07-31		Update for version 2.0.0 of the operational processor: • New pixel size (5.6x3.6 km²) from 06 August 2019 and on • A filter based on SO2 slant columns is now applied for the selection of the radiances used as DOAS reference. • Use of the TM5 tropopause layer index for the polluted air mass factor.
2.1.0	2020-02-28		• Reference to new pixel size of 5.5x3.5 km² • Update of reference documents numbering • Validation section: link to MPC validation. • Slight adaption of conclusions section.
2.2.0	2020-06-15		• Update for issue 2.2.0 and reference documents numbering • Cleaning of document sections

Contents

1 Introduction

1.1 Identification

This document is the Algorithm Theoretical Basis Document (ATBD) for TROPOMI sulphur dioxide (SO_2) vertical column data products, measured in the ultraviolet part of the electromagnetic spectrum.

1.2 Purpose and objective

This document outlines the theoretical basis of the retrieval of SO_2 vertical columns products. Its aim is to provide the reader with information on what can be expected from the TROPOMI instrument, how the actual retrieval algorithm functions and performs and what external information is required in order to obtain the products.

1.3 Document overview

A general introduction to SO_2 column retrieval is given in Chapter 4. The retrieval algorithm is described in Chapter 5. The feasibility of the foreseen product retrieval is discussed in Chapter 6, addressing issues as computational effort and the requirements on auxiliary data. An error analysis of the retrieval method is presented in Chapter 7. A section on validation possibilities for the retrieved SO_2 data can be found in Chapter 8.

2 Applicable and reference documents

2.1 Applicable documents

[AD01] GMES Sentinel-5 Precursor – S5P System Requirement Document (SRD); **source:** ESA/ESTEC; **ref:** S5P-RS-ESA-SY-0002; **issue:** 4.1; **date:** 2011-04-29

[AD02] Sentinel-5P Level 2 Processor Development – Statement of Work -; **source:** ESA; **ref:** S5P-SW-ESA-GS-053; **issue:** 1.1; **date:** 2012-05-21

[AD03] S5P Level 2 Processor Development – Level 2 Processor Requirements Specifications; **source:** ESA; **ref:** S5P-SW-ESA-GS-054; **issue:** 1.2 draft**; date:** 2014-09-15

[AD04] S5P/TROPOMI Level 2 Product Development Plan, **source:** KNMI**, ref:** S5P-KNMI-L2CO-0010-PL **issue:** 1.1.0 **date:** 2014-06-02

[AD05] S5P – Tailoring of ECSS Standards for the Level 2 Processor Development; **source:** ESA; **ref:** SP-RS-ESA-GS-055; **issue:** 1.1dr**; date:** 2012-10-31

[AD06] Sentinel-5P Level 2 Processor Development: Coordination Tasks; **source:** ESA; **ref:** S5P-SW-ESA-GS-081; **issue:** 1; **date:** 2012-06-27

2.2 Standard documents

There are no standard documents

2.3 Reference documents

[RD01] Terms, definitions and abbreviations for TROPOMI L01b data processor; **source**: KNMI; **ref:** S5P-KNMI-L01B-0004-LI; **issue:** 3.0.0; **date:** 2013-11-08

[RD02] Terms, and symbols in the TROPOMI Algorithm Team; **source**: KNMI; **ref**: SN-TROPOMI-KNMI-L2-049-MA; **issue**: 1.0.0; **date**: 2015-07-16

[RD03] Science Requirements Document for TROPOMI. Volume 1; **source**: KNMI & SRON; **ref:** RS-TROPOMI-KNMI-017; issue: 2.0; **date:** 2008-10-30.

[RD04] GMES Sentinels 4 and 5 Mission Requirements Document (MRD); **source:** ESA; **ref:** EO-SMA-/1507/JL; **issue:** 3; **date:** 2011-09-21

[RD05] Report Of The Review Of User Requirements For Sentinels-4/-5; **source**: ESA; **ref**: EO-SMA-/1507/JL; **issue:** 2.1; **date:** 2011-12-21

[RD06] CAPACITY: Operational Atmospheric Chemistry Monitoring Missions – Final report; **source:** KNMI; **ref:** CAPACITY; **date:** Oct. 2005.

[RD07] CAMELOT: Observation Techniques and Mission Concepts for Atmospheric Chemistry; **source:** KNMI; **ref:** RP-CAM-KNMI-050; **date:** Nov. 2009.

[RD08] TRAQ: Performance Analysis and Requirements Consolidation - Final Report; **source:** KNMI; **ref:** RP-ONTRAQ-KNMI-051; **date:** Jan. 2010.

[RD09] S5P/TROPOMI ATBD of the Aerosol data products; source: KNMI; **ref:** S5P-KNMI-L2-008-RP-TROPOMI_ATBD_AI; **issue:** 1.1.0; **date:** 2018-06-15.

[RD10] S5P/TROPOMI ATBD of the Cloud data products; source: DLR; **ref:** S5P-DLR-L2-ATBD-400I_Clouds; **issue:** 2.2.0; **date:** 2020-06-15.

[RD11] S5P/TROPOMI ATBD of the total and tropospheric NO2 data products; source: KNMI; ref: S5P-KNMI-L2-0005-RP-ATBD_NO2_data_products; **issue:** 1.4.0; **date:** 2019-02-06.

[RD12] S5P/TROPOMI ATBD of the HCHO data products; source: BIRA-IASB; r**ef:** S5P-L2-BIRA-ATBD-400F; **issue:** 2.2.0; **date:** 2020-06-15.

[RD13] S5P/TROPOMI Science Verification Plan; **source:** IUP; ref: S5P-IUP-L2-ScVR-RP-Sc_Verification_Report; **issue:** 2.1; **date:** 2015-12-18.

[RD14] TROPOMI Instrument Performance Analyses Report; **source:** DutchSpace; **ref:** TROP-DS-0000-RP-0060**; issue:** 6.0**; date:** 2013-01-16

[RD15] S5P/TROPOMI Static input for Level 2 processors; **source:** KNMI; **ref:** S5P-KNMI-L2CO-0004-SD; **issue:** 3.0.0; **date:** 2015-02-27

[RD16] Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Sulphur Dioxide SO2; source: DLR; ref: S5P-L2-DLR-PUM-400E; issue: 2.2.0; date: 2020-06-15.

2.4 Electronic references

[URL01] http://uv-vis.aeronomie.be/software/QDOAS/

[URL02] http://sacs.aeronomie.be/

[URL03] http://www.rtslidort.com/

[URL04] http://www.spec.org/cpu2006/results/res2010q2/cpu2006-20100413-10585.html

[URL05] http://www.woudc.org/

[URL06] http://www.projects.science.uu.nl/tm5/

3 Terms, definitions and abbreviated terms

Terms, definitions and abbreviated terms that are used in development program for the TROPOMI L0-1b data processor are described in [RD01]. Terms, definitions and abbreviated terms that are used in development program for the TROPOMI L2 data processors are described in [RD02]. Terms, definitions and abbreviated terms that are specific for this document can be found below.

3.1 Terms and definitions

AK	Averaging Kernel
m	altitude-resolved air mass factor or weighting function
M	air-mass factor
N_s	slant column density
N_v	tropospheric vertical column density

3.2 Acronyms and abbreviations

AAI	Absorbing Aerosol Index
AK	Averaging Kernel
AMF	Air mass factor
AOD	Aerosol optical depth
AR	Alternative retrieval
BrO	Bromine Monoxide
CAL	Cloud As Layer
CAMELOT	Composition of the Atmospheric Mission concEpts and SentineL Observation Techniques
CAPACITY	Composition of the Atmosphere: Progress to Applications in the user CommunITY
CCD	Charged Coupled Device
CRB	Clouds as Reflecting Boundaries
CTM	Chemical Transport Model
DOAS	Differential optical absorption spectroscopy
DU	Dobson Unit (1 DU = 2.6867×10^{16} molecules cm^{-2})
ECMWF	European Centre for Medium Range Weather Forecast
ESA	European Space Agency
FT	Free-troposphere
FWHM	Full Width Half Maximum
GMES	Global Monitoring for Environment and Security
GOME-2	Global Ozone Monitoring Experiment–2
HCHO	Formaldehyde
IPA	Independent Pixel Approximation
IR	Infrared
L2WG	Level-2 Working Group
LER	Lambertian Equivalent Reflector
LIDORT	LInearized Discrete Ordinate Radiative Transfer
LOS	Line-of-sight angle
LS	Lower stratosphere
LUT	Look-up table
MAX-DOAS	Multi-axis DOAS
MR	Medium Retrieval

NO_2	Nitrogen Dioxide
NOVAC	Network for Observation of Volcanic and Atmospheric Change
NRT	Near-real time
OCRA	Optical Cloud Recognition Algorithm
O_3	Ozone
OMI	Ozone Monitoring Instrument
OMPS	Ozone Mapping Profiler Suite
PA	Prototype Algorithm
(P)BL	Planetary Boundary Layer
PCA	Principal Component Analysis
ROCINN	Retrieval Of Cloud Information using Neural Networks
RRS	Rotational Raman Scattering
RTM	Radiative transfer model
RAA	Relative azimuth angle
S-5P	Sentinel-5 Precursor
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric ChartograpHY
SCD	Slant column density
SCDE	Slant column density error
SNR	Signal-to-noise ratio
SO_2	Sulfur dioxide
SR	Standard retrieval
SWIR	Short-wave infrared
SZA	Solar zenith angle
TOMS	Total Ozone Mapping Spectrometer
TROPOMI	Tropospheric Monitoring Instrument
UPAS	Universal Processor for UV/VIS Atmospheric Spectrometers
UV	Ultraviolet
UVN	Ultraviolet/Visible/Near-infrared
VA	Verification Algorithm
VC(D)	Vertical column density
WF	Weighting Function

4 Introduction to the TROPOMI SO_2 data products.

Sulfur dioxide enters the Earth's atmosphere via both natural and anthropogenic processes. Through the formation of sulfate aerosols and sulfuric acid, it plays an important role on the chemistry at local and global scales and its impact ranges from short-term pollution to climate forcing. While about one-third of the global sulfur emissions originate from natural sources (volcanoes and biogenic dimethyl sulfide), the main contributor to the total budget is from anthropogenic emissions mainly from the combustion of fossil fuels (coal and oil) and from smelting.

4.1 Heritage

Over the last decades, a host of satellite-based UV-visible instruments have been used for the monitoring of anthropogenic and volcanic SO_2 emissions (see e.g. Figure 4-1). Total vertical column density (VCD) of SO_2 has been retrieved with the sensors TOMS (Krueger, 1983), GOME (Eisinger and Burrows, 1998; Khokar et al., 2005), SCIAMACHY (Afe et al., 2004), OMI (Krotkov et al., 2006; Yang et al., 2007, 2010; Li et al., 2013; Theys et al., 2015), GOME-2 (Rix et al., 2012; Nowlan et al., 2011; Richter et al., 2009; Bobrowski et al., 2010, Hörmann et al., 2013) and OMPS (Yang et al., 2013). In some cases, operational SO_2 retrieval streams have also been developed aiming at the delivery of SO_2 VCD in near real-time (NRT), i.e. typically with a delay of less than 3 hours (see e.g., Support to Aviation Control Service (SACS); [URL02]). Algorithms to retrieve SO_2 columns based on few wavelength pairs have been developed and extensively applied to TOMS and OMI (e.g., Krotkov et al., 2006; Yang et al., 2007 and references therein). Current algorithms exploit back-scattered radiances measurements in a wide spectral range using a direct fitting approach (Yang et al., 2010; Nowlan et al., 2011), a Principal Component Analysis method (Li et al., 2013) or (some form of) Differential Optical Absorption Spectroscopy (DOAS; Platt and Stutz, 2008), see e.g. Richter et al. (2009), Hörmann et al. (2013), Theys et al. (2015).

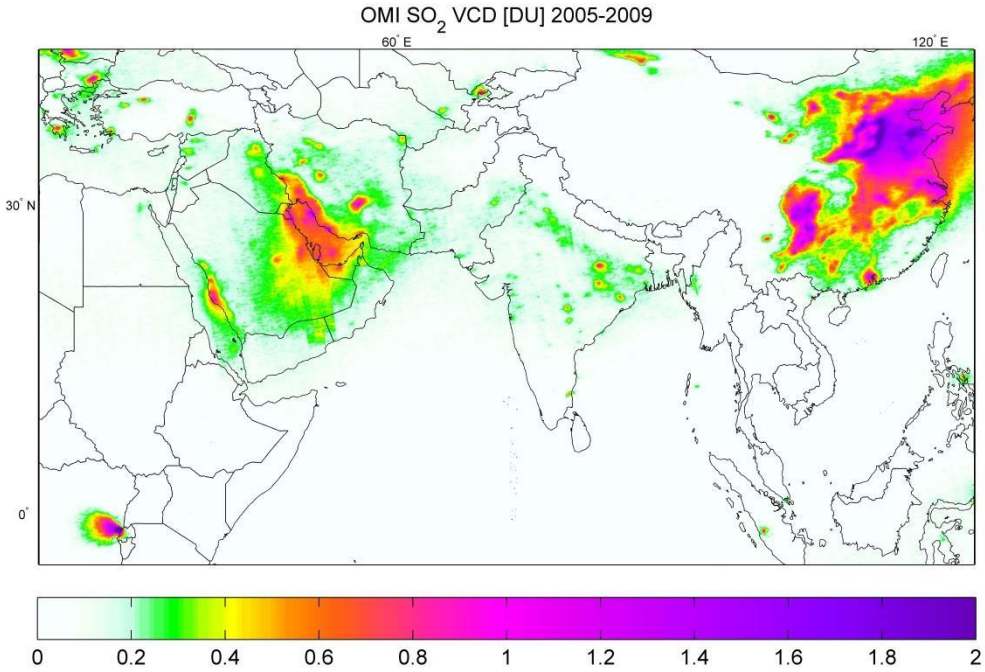

Figure 4-1 *Map of average SO_2 columns measured by OMI (Theys et al., 2015) for the 2005-2009 period, showing anthropogenic emission hotspots (China, Eastern Europe, India and the Middle East) and signals from volcanic activity (e.g. from the volcanoes in North Kivu).*

4.2 Overview of the retrieval algorithm

The retrieval algorithm is based on the DOAS technique. It is fully described in this document as well as in Theys et al. (2017) where additional verification results as part of the algorithm development are also presented. In brief, the log-ratio of the observed UV-visible spectrum, of radiation backscattered from the atmosphere, and an observed reference spectrum (solar or earthshine spectrum) is used to derive a slant column density (SCD), which represents the gas concentration integrated along the mean light path through the atmosphere. This is done by fitting absorption cross-sections of the relevant gases to the measured reflectance in a given spectral interval. In a second step, slant columns are corrected for possible biases. Finally, the slant columns are converted into vertical columns by means of air mass factors (AMF) obtained from radiative transfer calculations, accounting for the viewing geometry, clouds, surface properties and SO_2 vertical profile shapes.

The Sentinel-5p sensor TROPOMI samples the Earth's surface with a revisit time of one day and with an unprecedented spatial resolution of 7 x3.5 km^2 at best (5.5x3.5 km^2 from 6 August 2019). This allows the resolution of fine details and S5P is arguably be a valuable tool to better study anthropogenic SO_2 emissions but also volcanic emissions, from degassing to eruptive processes. Nevertheless, it poses additional constraints on the retrieval code, essentially for two reasons:

1) computational speed: the Level 1b data flow delivers spectral measurements for band 3 with a size of about 4 gigabytes per orbit (15 orbits daily); the SO_2 retrieval algorithm needs to be adapted for the NRT processing to reach the performance requirements (AD01);
2) treatment of non-linear effects is exacerbated by the excellent spatial resolution of the instrument (strong sources are better resolved); the SO_2 retrieval has to cope with these non-linear effects to reach the product accuracy requirements (see next section).

For very large SO_2 columns, such as occurring during explosive volcanic eruptions, the relation between the SO_2 signal and the VCD becomes strongly non-linear in the sensitive range to SO_2 absorption (310-325 nm), and DOAS typically underestimates the SO_2 VCD. Recently, alternative retrieval schemes to cope with these non-linear effects have been developed (Yang et al., 2007, 2010; Nowlan et al., 2011; Richter et al., 2009; Hörmann et al., 2013; Bobrowski et al., 2010). One option is to use a direct-fitting scheme, in which a full treatment of the radiative transfer is made for all wavelengths in the fitting window used and where simulated spectra are adjusted to the spectral observations. In recent studies, such methods have been used in the simultaneous retrieval of total column and effective altitude of volcanic SO_2 plumes (Yang et al., 2010; Nowlan et al., 2011;). The main disadvantage of direct fitting algorithms with respect to DOAS, is that they are computationally expensive and are out of reach for TROPOMI operational near-real-time processing. Even when limiting this technique to large volcanic events, a large eruption, covering 1% of the ground pixels with high SO2 content, the processing would take some several times longer than processing all pixels with a DOAS technique. Another possibility (that is adopted here) is to apply DOAS in different fitting windows (in the 310-390 nm spectral range) that are still sensitive enough to SO_2 but less affected by non-linear effects (Bobrowski et al., 2010; Hörmann et al., 2013).

4.3 Product requirements

While UV measurements are highly sensitive to SO_2 at high altitudes (upper troposphere-lower stratosphere), the sensitivity to SO_2 concentration in the boundary layer is intrinsically limited from space due to the combined effect of scattering (on molecules, aerosol and cloud particles) and ozone absorption that hamper the penetration of solar radiation into the lowest atmospheric layers. Furthermore the SO_2 absorption signature suffers from the interference with the ozone absorption spectrum. The retrieval precision (or random uncertainty) is driven by the signal to noise ratio of the recorded spectra and by the retrieval interval used, the accuracy (or systematic uncertainty) is limited by the knowledge on the auxiliary parameters needed in the different retrieval steps. Among these are the treatment of other chemical interfering species, clouds and aerosol, the representation of vertical profiles (gas, temperature, pressure), and uncertainties on data from external sources.

Requirements on the accuracy and precision for the data products derived from the TROPOMI measurements are specified in the GMES Sentinels 4 and 5 and 5p Mission Requirements Document [RD04], Report of The Review Of User Requirements for Sentinels-4/5 [RD05] and the Science Requirements Document for TROPOMI [RD03]. These requirements derive from the CAPACITY study [RD06] and have been fine-tuned by the CAMELOT [RD07] and ONTRAQ [RD08] studies. The CAPACITY study has defined three main themes: The ozone layer (A), air quality (B), and climate (C) with further division into sub themes. Requirements for SO_2 have been specified for a number of these sub themes. In the following sections, we discuss these requirements and the expected performances of the SO_2 retrieval algorithm (summary in Table 4-1).

Table 4-1 *Requirements on SO_2 vertical column products as derived from the MRTD. Numbers denote accuracy / precision, respectively.*

	Horizontal resolution [km]	Required uncertainty	Achievable uncertainty	Theme (Table in MRTD)
Enhanced stratospheric column	50-200	30% for VCD>0.5 DU	Met for VCD > 0.5DU	A3
Tropospheric column	5-20	30-60% or 1.3 x 10^{15} molecules cm^{-2} (least stringent)	50% / 3-6 x 10^{16} molec. cm^{-2}	B1, B2, B3
Total column	5-20	30-60% or 1.3 x 10^{15} molecules cm^{-2} (least stringent)	50% / 3-6 x 10^{16} molec. cm^{-2}	B1, B2, B3

4.3.1 Theme A3 Ozone layer assessment

This theme addresses the importance of measurements in the case of enhanced SO_2 concentrations in the stratosphere due to severe volcanic events. Long-term presence (up to several weeks) of SO_2 in the stratosphere contributes to the stratospheric aerosol loading and hence affect the climate and the stratospheric ozone budget. For such scenarios, the requirements state that the stratospheric vertical column should be monitored with a total uncertainty of 30%. Although such powerful volcanic events generally produce large amounts of SO_2, monitoring such a plume over extended periods of time requires the detection of the plume also after it has diluted during the weeks after the eruption.

From an error analysis of the proposed SO_2 algorithm (chapter 7), we have assessed the major sources of uncertainty in the retrieved SO_2 column. One of the main contributors to the total uncertainty is due to instrumental noise. This source of error alone limits the precision to vertical columns above about 0.25 DU. For SO_2 in the stratosphere, the summing up of the various uncertainties (Chapter 7) is believed to be around the required uncertainty of 30% for diluted SO_2 plumes, provided that the vertical column is larger than 0.5 DU. Explosive volcanic eruptions capable of ejecting SO_2 into the stratosphere regularly show stratospheric SO_2 concentrations of a few DU to several hundreds of DU or more, as was the case, for example, for the eruptions of Mt. Kasatochi (Yang et al., 2010) and Sarychev Peak (Carn et al., 2011). For very large SO_2 concentrations, the dynamical use of different fitting windows enables to reach 30 % uncertainty level.

4.3.2 Theme B – Air quality

This theme includes three sub themes:

B1 -Protocol monitoring. This involves the monitoring of abundances and concentrations of atmospheric constituents, driven by several agreements, such as the Gothenburg protocol, National Emission Ceilings, and EU Air Quality regulations.

B2 -Near-real time data requirements. This comprises the relatively fast (~30 minutes) prediction and determination of surface concentrations in relation to health and safety warnings.

B3 - Assessment. This sub theme aims at answering several air quality related science questions, such as the effect on air quality of special and temporal variations in oxidising capacity and long-range transport of atmospheric constituents.

A more detailed description of the air quality sub themes can be found in [RD04].

The user requirements on SO_2 products are equal for all three sub themes. For the total vertical column and the tropospheric vertical column of SO_2, the user requirements state an absolute maximum uncertainty of 1.3×10^{15} molecules cm^{-2} or 0.05 DU. This number derives from the ESA CAPACITY study, where the number was expressed as 0.4 ppbv for a 1.5 km thick boundary layer reaching up to 850 hPa. From the error budget (Chapter 7), the uncertainties due to instrument noise already indicate that the 0.05 DU requirement cannot be met on a single-measurement basis. This limitation was already found in the ESA CAMELOT study [RD07].

4.3.2.1 Anthropogenic pollution

In general, SO_2 emitted into the atmosphere due to anthropogenic activity is confined to the planetary boundary layer (PBL). The MRD requirement for the uncertainty on PBL columns is 1.3×10^{15} molecules cm^{-2} or 0.05 DU. However, calculations performed within the CAMELOT study showed that the smallest vertical column that can be detected in the PBL is of about 1-3 DU (for a signal-to-noise ratio (SNR) of 1000). Although pollution hotspots can be better identified by spatial or temporal averaging, several uncertainties (due to e.g. surface albedo or SO_2 vertical profile shape) are not averaging out and directly limit the product accuracy to about 50% or more.

Although the difference between the MRD requirements and the expected TROPOMI performance is rather large, one could argue that the required threshold should not be a strict criterion in all circumstances. The user requirement of 0.05 DU represents the maximum uncertainty required to distinguish (anthropogenic) pollution sources from background concentrations.

An ESA document [RD05] that reviews the MRD user requirements proposes a motivated relaxation of certain user requirements for specific conditions. For measurements in the PBL, the document proposes a relative requirement of 30-60% in order to discriminate between enhanced (> 1.5 ppbv/0.6 DU), moderate (0.5-1.5 ppbv/0.2-0.6 DU), and above background concentrations (>0.5 ppbv/0.2 DU). TROPOMI is able to discriminate these three levels with spatially-temporally averaged data.

4.3.2.2 Volcanic SO_2 in the free troposphere.

A better measurement sensitivity is expected for TROPOMI measurements for volcanic plumes in the free-troposphere. The expected precision is about 0.5 DU on the vertical column. The accuracy on the SO_2 vertical column is strongly limited by the SO_2 plume height and the cloud conditions. As these parameters are highly variable in practice, it is difficult to ascertain the product accuracy for these conditions.

5 Algorithm description

5.1 Overview of the retrieval scheme

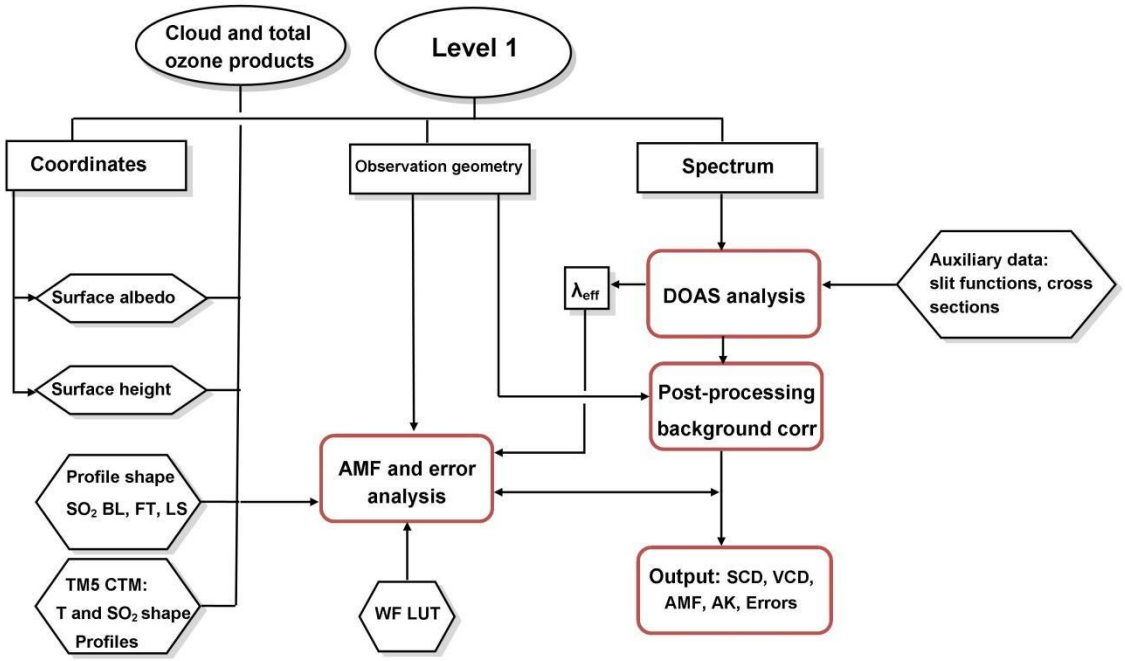

Figure 5-1 *Flow Diagram of the TROPOMI DOAS retrieval algorithm for SO_2.*

Figure 5-1 shows the full flow diagram of the SO_2 retrieval algorithm including the dependencies with auxiliary data and other L2 products. The algorithm and its application to OMI data is also described in Theys et al. (2015), although there are differences in some settings. The baseline operation flow of the scheme is based on a DOAS retrieval algorithm and is identical to that implemented in the retrieval algorithm for HCHO (also developed by BIRA-IASB, see S5P HCHO ATBD [RD12]). The main output of the algorithm are SO_2 vertical column density, slant column density, air mass factor, Averaging Kernels (AK) and error estimates. Here, we will first briefly discuss the principle of the DOAS VCD retrieval before discussing the separate steps of the process in more details.

First, the radiance and irradiance data are read from a S5P L1b file, along with geolocation data such as pixel coordinates and observation geometry (sun and viewing angles). At this stage also cloud cover information is obtained from the S5P cloud L2 data, as required for the calculation of the AMF, later in the scheme. Then relevant absorption cross section data (SO_2), as well as characteristics of the instrument (e.g., slit functions) are used as input for the SO_2 slant column density determination. As a baseline, the slant column fit is done in a sensitive window from 312 to 326 nm. For pixels with a strong SO_2 signal, results from alternative windows, where the SO_2 absorption is weaker (section 5.2.1.1), can be used instead. An empirical offset correction (dependent on the fitting window used) is then applied to the SCD. The latter correction accounts for systematic biases in the SCDs. Following the SCD determination, the AMF is estimated. For computational efficiency, the algorithm makes no 'on the fly' calculation, but uses a pre-calculated box air mass factor look-up table (LUT). This look-up-table is generated using the LIDORT radiative transfer code and has several entries (see sections 5.2.3 and 7): cloud cover data, topographic information, observation geometry, surface albedo, effective wavelength (representative of the fitting window used), total ozone column and the shape of the vertical SO_2 profile.The algorithm also includes an error calculation and retrieval characterization module (section 7) that computes the so-called DOAS-type averaging kernels (Eskes & Boersma, 2003), which characterize the vertical sensitivity of the measurement and which are required for comparison with other types of data (Veefkind et al., 2012).

The SO_2 vertical column is obtained by:

$$N_v = \frac{N_S - N_S^{back}}{M} \qquad 5\text{-}1$$

where the main quantities are the vertical column (N_v), the slant column density (N_s) and the values used for the background correction (N_S^{back}). M is the air mass factor.

Finally, the algorithm also includes a post-processing step to identify which pixels contain elevated amounts of SO_2 and tentatively attribute an emission source type (anthropogenic or volcanic). This is achieved in the form of a detection flag and the corresponding algorithm module is described in section 5.3.

5.2 Algorithm components

5.2.1 Slant column retrieval

The backscattered radiance spectrum recorded by the space instrument differs from the solar spectrum because of the interactions of the photons with the Earth's atmosphere and surface reflection. Hence the reflectance spectra contain spectral features that can be related to the various absorbing species and their amounts in the atmosphere. The DOAS method aims at the separation of the highly structured trace gas absorption spectra and broadband spectral structures. The technique relies on a number of assumptions that can be summarized as follows:

a. The spectral analysis and atmospheric radiative transfer computations are treated separately, by considering one averaged atmospheric light path of the photons travelling from the sun to the instrument.

b. The absorption cross-sections are not strongly dependent on pressure and temperature. Additionally, the averaged light path should be weakly dependent on the wavelength - for the fitting window used - which enables defining an effective absorption (slant) column density. It should be noted that strictly this is not valid for the SO_2 DOAS retrieval because of strong absorption by ozone and in some cases SO_2 itself (for large SO_2 amounts).

c. Spectrally smoothed structures due broadband absorption, scattering and reflection processes can be well reproduced by a low-order polynomial as a function of wavelength.

Figure 5-2 sketches the geometry of a nadir satellite radiation measurement, defining the solar and viewing zenith/line-of-sight angles (θ_0 and θ) and the solar and viewing azimuth angles (φ_0 and φ) towards the zenith in a plane parallel atmosphere.

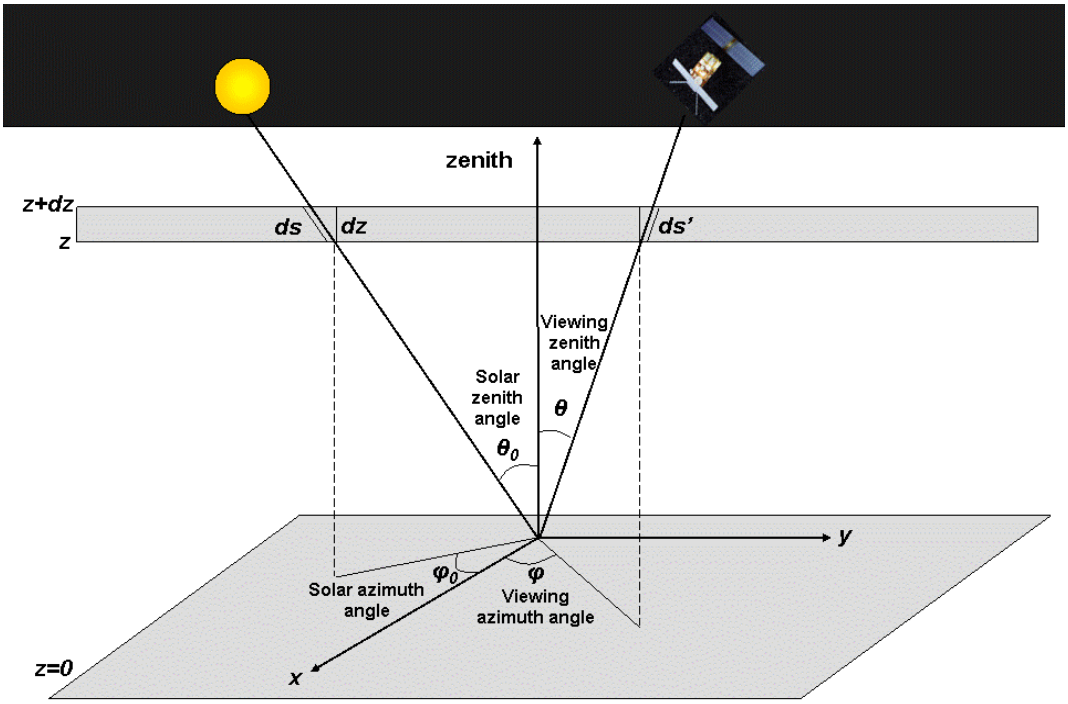

Figure 5-2 *Sketch of satellite radiation measurement and geometry in a plane parallel atmosphere.*

Photons collected by the satellite instrument may have followed very different light paths through the atmosphere depending on their scattering history. However, a single effective light path is assumed, which represents an average of the complex paths of all reflected and scattered solar photons reaching the instrument within the spectral interval used for the retrieval. This simplification is valid if the effective light path is reasonably constant over the considered wavelength range. The spectral analysis can be described by the following equation:

$$\ln\frac{\pi I(\lambda)}{E_0(\lambda)} = -\sum_j \sigma_j'(\lambda)\, Ns_j + \sum_p c_p \lambda^p \qquad 5\text{-}2$$

Here, I(λ) is the observed backscattered Earthshine radiance [W $m^{-2}nm^{-1}sr^{-1}$], E_0 is the solar irradiance [W $m^{-2}nm^{-1}$]. The first term on the right hand side indicates all relevant absorbing species with differential absorption cross-sections σ_j' [cm^2 molec.$^{-1}$]. Integration of the number densities of these species along the effective light path gives the slant column density Ns_j. [molec.cm^{-2}] Eq. 5.2 can be solved by least-squares fitting techniques (Platt and Stutz, 2008) for the slant column values. The final term in Eq.5-2 is the polynomial representing broad band absorption and (Rayleigh and Mie) scattering structures in the observed spectrum and also account for possible errors such as e.g. uncorrected instrument degradation effects, uncertainties in the radiometric calibration or possible residual (smooth) polarisation response effects not accounted for in the level 0-1 processing.

Apart from the cross-sections for the trace gases of interest, additional fit parameters need to be introduced to account for the effect of several physical phenomena on the fit result. For SO_2 fitting these are the filling-in of Fraunhofer lines (Ring effect) and the need for an intensity offset-correction. In the above, we have assumed that for the ensemble of observed photons a single effective light path can be assumed over the adopted wavelength fitting interval. For the observation of (generally small) SO_2 concentrations at large solar zenith angles this is not necessarily the case. For those long light paths, the large contribution of O_3 absorption may lead to negative SO_2 retrievals. This may be mitigated by taking the wavelength dependence of the O_3 SCD over the fitting window into account, as will be described in the next section.

The different parts of the DOAS retrieval are detailed in the next subsections and Table 5-1 gives a summary of settings used to invert SO_2 slant columns, Note that in Eq 5-2, the daily solar irradiance is used. As a better option, it is generally preferred to use daily averaged radiances, selected for each across-track position, in the equatorial Pacific. It allows for better handling of instrumental artifacts and degradation of the recorded spectra for each detector, and this is the baseline approach for TROPOMI SO_2 retrieval. In the NRT algorithm, the last 5 days are used to derive the reference average spectra, while in the offline version of the algorithm, 5 days around the current day are used. It allows e.g. for better handling of instrumental artifacts and degradation of the recorded spectra for each detector. From version 2.0.0 of the operational processor, a filter is applied on the radiances based on the retrieved SO_2 concentrations of the previous days, in order to avoid contamination of the reference radiance with volcanic signal. This is performed using a merge of two sectors over the Pacific Ocean and Indian Ocean, to increase the chances to find uncontaminated spectra.

Table 5-1 DOAS settings used to retrieved SO_2 slant columns

Fitting intervals 1 and 2	312-326 nm (w1), 325-335 nm (w2)
Cross-sections	SO_2 203K [*Bogumil et al.*, 2003] O_3 228K and 243K with *Io* correction [*Brion et al.*, 1998] Pseudo O_3 cross sections ($\lambda\sigma_{O3}$, σ_{O3}^2) [*Puķīte et al.*, 2010] Ring effect: 2 eigenvectors [*Vountas et al.*, 1998] generated for 20° and 87° solar zenith angles using LIDORT-RRS [*Spurr*, 2008]
Polynomial	5th order
Fitting interval 3	360-390 nm (w3)
Cross-sections	SO_2 *Hermans et al.* [2009] extrapolated at 203K NO_2 220K [*Vandaele et al.*, 1998] O_2-O_2 [*Greenblatt et al.*, 1990] Ring effect: single spectrum [*Chance and Spurr*, 1997]
Polynomial	4th order
Intensity offset correction	Linear offset
Spectrum shift and stretch	Fitted
Spectral spikes removal procedure	*Richter et al.* [2011]
Reference spectrum	Averaged earthshine spectrum in Pacific region (10°S-10°N, 160°E-120°W) and Indian Ocean (6°S-6°N, 50°E-100°E); separate spectrum for each detector row. NRT: averaged spectra of the last 5 available days, Off-line: averaged spectra of the 5 days around the current day. Radiances presenting

large SO_2 absorption are filtered out based on previous L2 SO_2 results.

5.2.1.1 Wavelength fitting windows

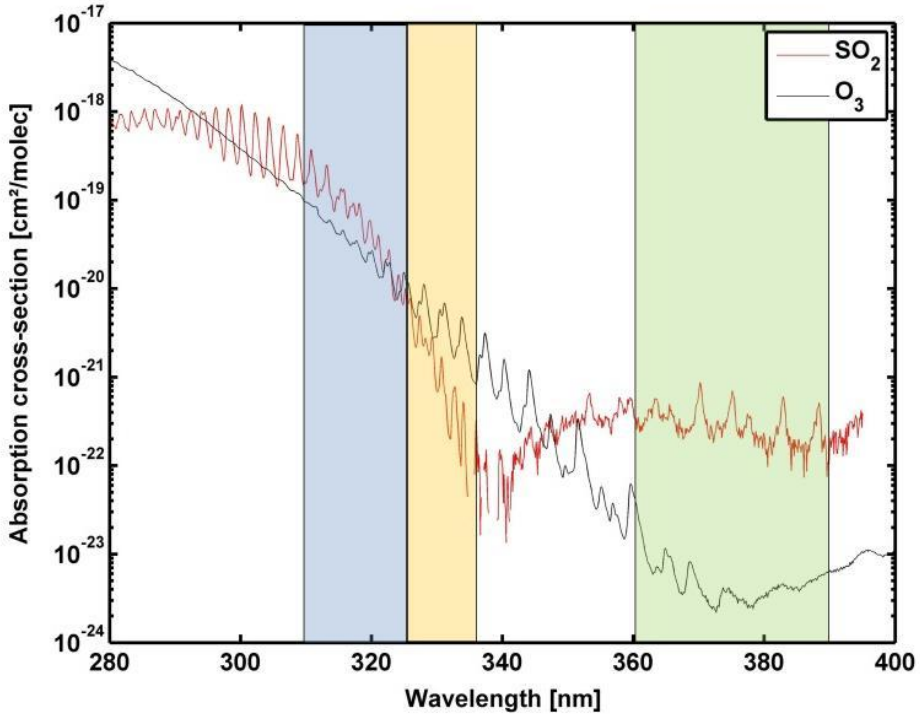

Figure 5-3 *Absorption cross-sections of SO_2 and O_3. The blue, yellow and green boxes delimit the three envisaged SO_2 fitting windows 312-326 nm, 325-335 nm and 360-390 nm, respectively.*

DOAS measurements are in principle applicable to all gases having suitable narrow absorption bands in the UV, visible, or near IR regions. However, the generally low concentrations of these compounds in the atmosphere, and the limited signal-to-noise ratio of the spectrometers, restrict the number of trace gases that can be detected. Many spectral regions contain several interfering absorbers and correlations between absorber cross-sections can sometimes lead to systematic biases in the retrieved slant columns. In general, the correlation between cross-sections decrease if the wavelength interval is extended, but the assumption of a single effective light path defined for the entire wavelength interval may not be fully satisfied, leading to systematic misfit effects that may also introduce biases in the retrieved slant columns (e.g., Pukīţe et al., 2010). To optimize DOAS retrieval settings, a trade-off has to be found between these effects. In the UV-visible spectral region, the cross-section spectrum of SO_2 has its strongest bands in the 280-320 nm range (Figure 5-3). For the short wavelengths in this range, the SO_2 signal however suffers from a strong increase in Rayleigh scattering and ozone absorption. In practice, this leads to a very small SO_2 signal in the satellite spectra compared to ozone absorption, especially for tropospheric SO_2. Consequently, SO_2 is traditionally retrieved using sensitive windows in the 310-326 nm range (GOME, SCIAMACHY, GOME-2, OMI). Note that even in this range the SO_2 absorption can be three orders of magnitude lower than that of ozone.

The TROPOMI SO_2 algorithm is using a multiple windows approach:

- 312-326 nm: classical fitting window, ideal for small columns. This window is used as baseline. If non-linear effects due to high SO_2 amounts are encountered, one of the two following windows will be used instead.
- 325-335 nm: in this window, differential SO_2 spectral features are one order of magnitude smaller than in the classical window. It allows the retrieval of moderate SO_2 columns, an approach similar to the one described by Hörmann et al. (2013).
- 360-390 nm: SO_2 absorption bands are 2-3 orders of magnitude weaker than in the classical window and are best suited for the retrieval of extremely high SO_2 columns (Bobrowski et al., 2010)

Note that in the 325-335 nm and 360-390 nm windows the Rayleigh scattering and ozone absorption are less important than in the baseline 312-326 nm window (see also Figure 5-3).

Specifically, in the first two intervals, absorption cross-sections of O_3 at 228K and 243K are included in the fit. Moreover to better cope with the strong (non-linear) ozone absorption at short wavelengths the retrieval includes the fitting of two pseudo cross-sections following the approach of Puķīte et al. (2010): $\lambda\sigma_{O3}$ and σ_{O3}^2 using the O_3 cross-section at 228K. The correction for the Ring effect is based on the technique outlined by Vountas et al. (1998). This technique involves a Principal Component Analysis of a set of Ring spectra, calculated for a range of solar zenith angles. The first two of the resulting eigenvectors appear to accurately describe the Ring spectra, with the first eigenvector representing the filling-in of Fraunhofer lines and the second mostly representing the filling-in of gas absorption features. In the retrieval algorithm, these vectors are determined by orthogonalizing two Ring spectra, calculated by LIDORT-RRS for a low SZA (20°) and a high SZA (87°), respectively.

5.2.1.2 Wavelength calibration and convolution to TROPOMI resolution

The quality of a DOAS fit critically depends on the accuracy of the alignment between the earthshine radiance spectrum, the reference (solar irradiance) spectrum and the cross sections. Although the Level 1b contains a spectral assignment, an additional spectral calibration is part of the SO_2 algorithm and is based on positioning of solar Fraunhofer lines (read below). Moreover, the DOAS spectral analysis includes also the fit of shift and stretch of radiance spectra because the TROPOMI spectral registration differs from one ground-pixel to another e.g. due to thermal variations over the orbit as well as due to inhomogeneous filling of the slit in the flight direction, etc

The wavelength registration of the reference spectrum can be fine-tuned by means of a calibration procedure making use of the solar Fraunhofer lines. To this end, a reference solar atlas E_s accurate in absolute vacuum wavelength to better than 0.001 nm (Chance and Kurucz, 2010) is degraded at the resolution of the instrument, through convolution by the TROPOMI instrumental slit function.

Using a non-linear least-squares approach, the shift (Δ_i) between the reference solar atlas and the TROPOMI irradiance is determined in a set of equally spaced sub-intervals covering a spectral range large enough to encompass all relevant fitting intervals. The shift is derived according to the following equation:

$$E_0(\lambda) = E_s(\lambda - \Delta_i) \qquad 5\text{-}3$$

where E_s is the solar spectrum convolved at the resolution of the instrument and Δ_i is the shift in sub-interval *i*. A polynomial is then fitted through the individual points in order to reconstruct an accurate wavelength calibration $\Delta(\lambda)$ for the complete analysis interval. Note that this approach allows to compensate for stretch and shift errors in the original wavelength assignment.

In the case of TROPOMI, the procedure is complicated by the fact that such calibrations must be performed (and stored) for each separate spectral field on the CCD detector array. Indeed due to the imperfect characteristics of the imaging optics, each row of the TROPOMI instrument must be considered as a separate detector for analysis purposes.

In a subsequent step of the processing, the absorption cross sections of the different trace gases must be convolved with the instrumental slit function. The baseline approach was to use slit functions determined as part of the TROPOMI key data. Slit functions are delivered for each binned spectrum and as a function of the wavelength. Note that an additional feature of the (prototype) algorithm allows to dynamically fit for an effective slit function of known line shape. This has been used for verification and monitoring purpose during commissioning . In brief, the wavelength calibration procedure allows for stretching of slit function left and right wings and the fitted stretch parameters lead top optimized slit functions, which have then be used for all subsequent spectral convolutions. After some iterations, it was shown during the commissioning phase that the optimized slit functions were very close to the latest released TROPOMI instrumental slit function data set.

More specifically, wavelength calibrations are made for each TROPOMI orbit as follows:

1. The TROPOMI irradiances (one for each row of the CCD) are calibrated in wavelength over the 311-391 nm wavelength range, using 10 sub-windows.
2. The earthshine radiances and the absorption cross-sections are interpolated (cubic spline interpolation) on the calibrated wavelength grid, prior to the analysis.
3. During spectral fitting, shift and stretch parameters are further derived to align radiance and irradiance spectra. The reference wavelength grid used in the DOAS procedure is the (optimised) grid of the TROPOMI solar irradiance.

5.2.1.3 Spike removal algorithm

A method to remove individual hot pixels or pixels affected by the South Atlantic Anomaly has been presented for NO_2 retrievals in Richter et al. (2011). Often only a few individual detector pixels are affected and in these cases, it is possible to identify and remove the noisy points from the fit. However, as the amplitude of the distortion is usually only of the order of a few percent or less, it cannot always be found in the highly structured spectra themselves. Higher sensitivity for spikes can be achieved by analysing the residual of the fit where the contribution of the Fraunhofer lines, scattering, and absorption is already removed.

When the residual for a single pixel exceeds the average residual of all pixels by a chosen threshold ratio (the tolerance factor, here fixed to 5), the pixel is excluded from the analysis, in an iterative process. This procedure is repeated until no further outliers are identified, or until the maximum number of iterations is reached (here fixed to 3). This is especially important to handle the degradation of instruments such as OMI or TROPOMI.

5.2.1.4 Fitting window selection.

The implementation of multiple fitting windows retrieval requires selection criteria for the transition from one window to another. These criteria are based on the measured SO_2 slant columns. As a baseline, the SO_2 SCD in the 312-326 nm window is retrieved for each satellite pixel. When the resulting value exceeds a certain criterion, the slant column retrieval is repeated with an alternative window. As part of the algorithm development and during the verification exercise [RD13], closed-loop retrievals have been performed and application of the algorithm to real data from the GOME-2 and OMI instruments lead to threshold values and criteria as given in Table 5-2.

Table 5-2 *Criteria for selecting alternative fitting windows*.

Window number	w1	w2	w3
Wavelength range	312 – 326 nm	325-335 nm	360-390 nm
Derived slant column	S1	S2	S3
Application	Baseline for every pixel	S1 > 15 DU and S2 > S1	S2 > 250 DU and S3 > S2

5.2.2 Offset correction

When applying the algorithm to OMI and GOME-2 data, across-track/viewing angle dependent residuals of SO_2 are found over clean areas and negative SO_2 SCDs are found at high SZA which need to be corrected. A background correction scheme was found mostly necessary for the SO_2 slant columns retrieved in the baseline fitting window 1. The adopted correction scheme depends on across-track position and measured O_3 slant column as described below.

The correction is based on a parameterization of the background values that are then subtracted from the measurements. The scheme first removes pixels with high SZA (>70°) and SCDs larger than 1.5 DU (measurements with presumably real SO_2) and then calculates the offset correction by averaging the SO_2 data on an ozone slant column grid (bins of 75 DU). This is done independently for each across-track position and hemisphere, and the correction makes use of measurements averaged over a time period of five days preceding/around the measurement of interest for the offline/near-real-time processing, respectively (this is to improve the statistics and minimize the impact of a possible extended volcanic SO_2 plume on the averaged values).

It should be noted that the O_3 slant column is dependent on the wavelength when applying the approach of Puķīte et al. (2010):

$$SCD(\lambda) = SCD_{T1} + SCD_{T2} + \lambda.SCD_{\lambda} + \sigma_s(\lambda)SCD_S \qquad 5\text{-}4$$

SCD_{T1} and SCD_{T2} are the retrieved ozone slant columns corresponding to the ozone cross-sections at two temperatures included in the fit. SCD_{λ} and SCD_s are the retrieved parameters for the two pseudo cross-sections $\lambda.\sigma_s$ and σ_s^2 (σ_s being the O_3 cross-section at T1). In order to apply the background correction, the O_3 slant column expression (Eq 5-4) is evaluated at 313 nm (read below).

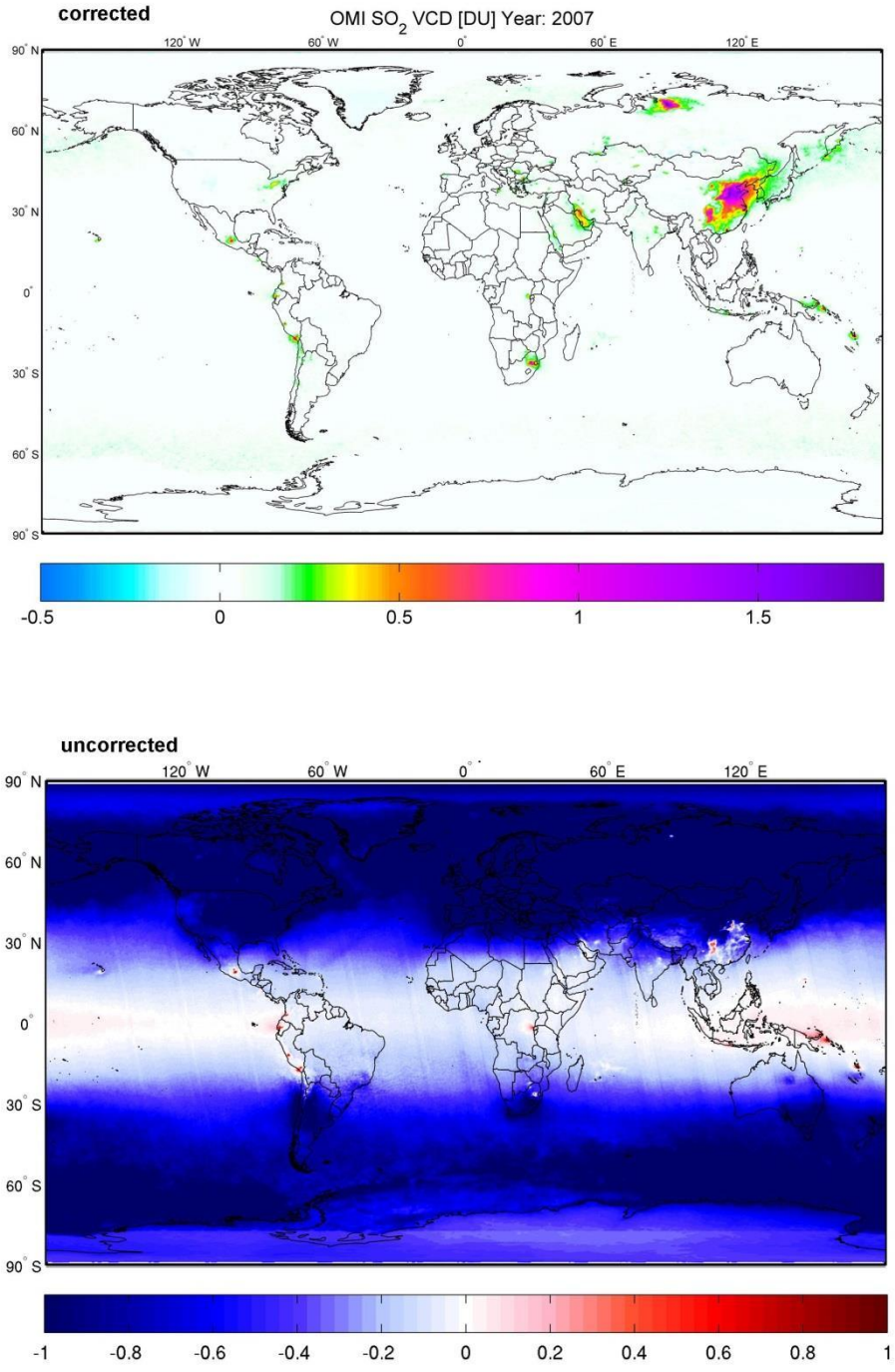

Figure 5-4 OMI *SO_2 vertical columns averaged for the year 2007 (top) with and (bottom) without background correction. Only clear sky pixels (cloud fraction lower than 30%) have been kept. AMFs calculated from SO_2 profiles from the IMAGES global model are applied to the slant columns (Theys et al., 2015).*

An example of the effect of the background correction is shown in Figure 5-4 for OMI. One can see that after correction (top panel) the retrievals shows smooth/unstriped results and values close to zero outside the polluted areas. In some regions (in particular at high latitudes), residual columns can be found but are generally lower than 0.2 DU.

For the two additional fitting windows, residual SO_2 levels are relatively small in comparison to the volcanic column amounts expected to be retrieved in these windows. However, simplified background corrections are also applied to the alternative windows: the offset corrections use parameterizations of the background slant columns based on latitude (bins of 5°), cross-track position and time (two weeks moving averages as for the baseline window). To avoid contamination by strong volcanic eruptions, only the pixels are kept with SCD less than 50DU and 250DU for the fitting windows 325-335nm and 360-390nm, respectively.

It should be noted that the background corrections do not imply to save five days of SO_2 L2 data in the memory, but only the averaged values ($\Sigma_{i=1,N}$ SCD_i/ N) over the predefined working grids (note: the numerators $\Sigma_{i=1,N}$ SCD_i and denominators N need to be stored separately).

This background correction is well suited for the case of a 2D-detector array such as TROPOMI, for which across-track striping can possibly arise, due to imperfect cross-calibration and different dead/hot pixel masks for the CCD detector regions. This instrumental effect can also be found for scanning spectrometers, but since these instruments only have one single detector, such errors do not appear as stripes, but rather as constant, unknown offsets. These different retrieval artefacts can be compensated (up to a certain extent) using background corrections which depend on the across-track position. These corrections are also meant to handle the time-dependent degradation of the instrument. Note that experience with OMI shows that the most efficient method to avoid across-track stripes in the retrievals is to use row-dependent mean radiances as reference spectrum in the DOAS fit.

5.2.3 Air mass factors

The DOAS method assumes that the retrieved slant column (after appropriate background correction) can be converted into a vertical columns using a single air mass factor *M* (representative for the fitting interval):

$$M = \frac{Ns}{Nv} \qquad 5\text{-}5$$

and which is determined by radiative transfer calculations with LIDORT version 3.3 (Spurr, 2008). The AMF calculation is based on the formulation of Palmer et al. (2001):

$$M = \int m'(p) \cdot s(p) \mathrm{d}p \qquad 5\text{-}6$$

with $m'=m(p)/C_{temp}(p)$,where m(p) is the so-called weighting function (WF) or box air mass factor, C_{temp} is a temperature correction (see section 5.2.3.7) and *s* is the SO_2 normalized a priori mixing ratio profile, as function of pressure (*p)*.

The AMF calculation assumes Lambertian reflectors for the ground and the clouds and makes use of pre-calculated box air mass factor LUTs at 313, 326 and 375 nm (depending on the fitting window used; see also section 7). Calculating the AMF at these three wavelengths was found to give the best results using closed-loop retrievals.The box air mass factor depends on observation geometry (solar zenith angle: SZA, line-of-sight angle: LOS, relative azimuth angle: RAA), total ozone column (TO3), surface albedo (alb), surface pressure (p_s), cloud top pressure (p_{cloud}) and effective cloud fraction (f_{eff}).

Examples of SO_2 box air mass factors are displayed in Figure 5-5 (as a function of height for illustration purpose) and show the typical variations of the measurement sensitivity as a function of height, wavelength and surface albedo.

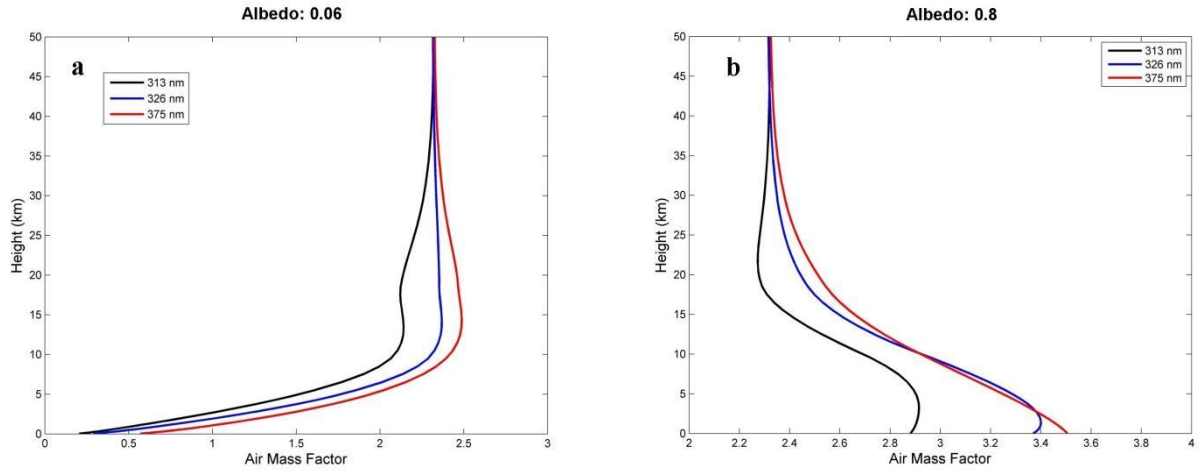

Figure 5-5 *SO_2 box air mass factors at 313, 326 and 375nm for albedo of (a) 0.06 and (b) 0.8. SZA: 40°, LOS: 10°, RAA: 0°, Surface Height: 0 km.*

The generation of the box air mass factor LUT has been done for a large range of physical parameters, listed in Table 5-3. In practice, the box air mass factor for each pixel is computed by linear interpolation of the box air mass factor LUT at the apriori profile pressure grid and using the auxiliary data sets described in the following sub-sections. Linear interpolations are performed along the cosine of solar and viewing angles, relative azimuth angle and surface albedo, while a nearest neighbour interpolation is performed in surface pressure. In particular, the grid of surface pressure is very thin near the ground, in order to minimise interpolation errors caused by the generally low albedo of ground surfaces. Furthermore, the LUT and model pressures are scaled to the respective surface pressures, in order to avoid extrapolations outside the LUT range.

Table 5-3 *Physical parameters that define the* box air mass factor *look-up table*

Parameter	Number of grid points	Grid values	Symbol
Atmospheric pressure [hPa]	64	1056.77, 1044.17,1031.72, 1019.41, 1007.26, 995.25, 983.38, 971.66, 960.07, 948.62, 937.31, 926.14, 915.09, 904.18, 887.87, 866.35, 845.39, 824.87, 804.88, 785.15, 765.68, 746.70, 728.18, 710.12, 692.31, 674.73, 657.60, 640.90, 624.63, 608.58, 592.75, 577.34, 562.32, 547.70, 522.83, 488.67, 456.36, 425.80, 396.93, 369.66, 343.94, 319.68, 296.84, 275.34, 245.99, 210.49, 179.89, 153.74, 131.40, 104.80, 76.59, 55.98, 40.98, 30.08, 18.73, 8.86, 4.31, 2.18, 1.14, 0.51, 0.14, 0.03, 0.01, 0.001	p_l
Altitude corresponding to the atmospheric pressure, using an US standard atmosphere [km]	64	-0.35, -0.25, -0.15, -0.05, 0.05, 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, 0.95, 1.10, 1.30, 1.50, 1.70, 1.90, 2.10, 2.30, 2.50, 2.70, 2.90, 3.10, 3.30, 3.50, 3.70, 3.90, 4.10, 4.30, 4.50, 4.70, 4.90, 5.25, 5.75, 6.25, 6.75, 7.25, 7.75, 8.25, 8.75, 9.25, 9.75, 10.50, 11.50, 12.50, 13.50, 14.50, 16.00, 18.00, 20.00, 22.00, 24.00, 27.50, 32.50, 37.50, 42.50, 47.50, 55.00, 65.00, 75.00, 85.00, 95.00	z_l
Solar zenith angle [°]	17	0, 10, 20, 30, 40, 45, 50, 55, 60, 65, 70, 72, 74, 76, 78, 80, 85	θ_0
Line of sight angle [°]	10	0, 10, 20, 30, 40, 50, 60, 65, 70, 75	θ
Relative azimuth angle [°]	5	0, 45, 90, 135, 180	φ
Total ozone column [DU]	4	205, 295, 385, 505	TO3
Surface albedo	14	0, 0.01, 0.025, 0.05, 0.075, 0.1, 0.15, 0.2, 0.25, 0.3 0.4, 0.6, 0.8, 1.0	A_s
Surface / cloud top pressure [hPa]	17	1063.10, 1037.90, 1013.30, 989.28, 965.83, 920.58, 876.98, 834.99, 795.01, 701.21, 616.60, 540.48, 411.05, 308.00, 226.99, 165.79, 121.11	p_s
AMF Wavelength	3	313, 326, 375	

5.2.3.1 Observation geometry

The LUT covers the full range of values for solar zenith angles, line-of-sight angles and relative azimuth angles that can be encountered in the TROPOMI measurements. The observation geometry is readily present in the L1b data of each satellite instrument.

5.2.3.2 Total ozone column

The measurement sensitivity at 313 nm is highly dependent on the total ozone absorption. The LUT covers a range of ozone column values from 200 to 500 DU for a set of typical ozone profiles. The total ozone column is directly available from the operational processing of S5P total ozone column product.

5.2.3.3 Surface albedo

The albedo value is very important for PBL anthropogenic SO_2 but less critical for volcanic SO_2 when it is higher in the atmosphere. For the surface albedo dimension, we use the climatological monthly minimum Lambertian equivalent reflector (minLER) data from Kleipool et al. (2008) at 328 nm for w1 and w2, and 376 m for w3. This database is based on OMI measurements and has a spatial resolution of 0.5° x 0.5°. Note that, other surface reflectance databases with improved spatial resolution (more appropriate for TROPOMI) will likely become available and these data sets will be considered for next algorithmic versions.

5.2.3.4 Clouds

The AMF calculations for TROPOMI partly cloudy scenes uses the cloud parameters (cloud fraction f_c, cloud albedo A_c, cloud pressure ctp) supplied by the nominal S5P cloud algorithm OCRA/ROCINN in its Clouds as Reflecting Boundaries (CRB) implementation [RD10]. The cloud surface is considered to be a Lambertian reflecting surface and the treatment of clouds is achieved through the independent pixel approximation (IPA; Martin et al., 2002) which considers a inhomogeneous satellite pixel as being composed (as for the radiance intensity) of two independent homogeneous scenes, one completely clear and the other completely cloudy. The box air mass factor is expressed as:

$$m(p) = \Phi m_{\text{cloud}}(p) + (1 - \Phi) m_{\text{clear}}(p) \qquad \text{5-7}$$

where Φ is the intensity-weighted cloud fraction or cloud radiance fraction

$$\Phi = \frac{f_c I_{cloud}}{f_c I_{cloud} + (1 - f_c) I_{clear}} \qquad \text{5-8}$$

The suffices clear and cloudy refers to the box air mass factor and Intensity calculation corresponding to a fully clear or cloudy pixel, respectively. The box air mass factor LUT is therefore accompanied by an intensity LUT with the same input grids. Both LUTs have been generated for a range of cloud cover fractions and cloud top pressures.

Note that the variations of the cloud albedo are directly related to the cloud optical thickness. Strictly speaking in a Lambertian (reflective) cloud model approach, only thick clouds can be represented. An effective cloud fraction corresponding to an effective cloud albedo of 0.8 ($f_{eff} = f_c \frac{A_c}{0.8}$) can be defined, in order to transform optically thin clouds into equivalent optically thick clouds of reduced extent. Note that in some cases (thick clouds with A_C>0.8) the effective cloud fraction can be larger than one and the algorithm assumes f_{eff}=1. In such altitude dependent air mass factor calculations, a single cloud top pressure is assumed within a given viewing scene. For low effective cloud fractions (f_{eff} lower than 10%), the current cloud top pressure output is highly unstable and it is therefore reasonable to consider the observation as a clear-sky pixel (i.e. the cloud fraction is set to 0 in Eq 5-8) in order to avoid unnecessary random error propagation through the retrievals, which can be as high as 100%. Moreover, it has been shown recently by Wang et al. (2017) using multi-axis DOAS (MAX-DOAS) observations to validate satellite data that in case of elevated aerosol loadings in the PBL (typically leading to apparent f_{eff} up to 10%), it is recommended to apply clear-sky AMFs rather than total AMFs (based on cloud parameters) that presumably correct implicitly for the aerosol effect on the measurement sensitivity.

It should be noted that the formulation of the pressure dependent air mass factor for a partly cloudy pixel implicitly includes a correction for the SO_2 column lying below the cloud and therefore not seen by the satellite, the so-called “ghost column”. Indeed, the total AMF calculation as expressed by Eqs. 5.6 and 5.7 assumes the same shape factor and implies an integration of the a priori profile from the top of atmosphere to the ground, for each fraction of the scene. The ghost column information is thus coming from the a priori profile shapes. For this reason, only observations with moderate cloud fractions (f_{eff} lower than 30%) should be used, unless it can be assumed that the cloud cover is entirely situated below the SO_2 layer, i.e. a typical situation for volcanic plumes. Whereas the SO_2 plume height is typically unknown for volcanic cases, a slightly different approach is used for the AMF calculation as described in section 5.2.3.6.

.

5.2.3.5 Surface height

The surface height (z_s) is determined for each pixel by interpolating the values of a high resolution digital elevation map, GMTED2010 (Danielson et al., 2011).

5.2.3.6 Profile shapes

It is generally not possible to know at the time of observation what is the SO_2 vertical profile and whether the observed SO_2 is of volcanic origin or from pollution (or both). Therefore, the algorithm computes four vertical columns for different hypothetical SO_2 profiles (see Table 6-2).

Three box profiles of 1 km thickness, located in the boundary layer, upper-troposphere and lower-stratosphere, are used. The first box profile stands for typical conditions of well mixed SO_2 (from volcanic or anthropogenic emissions) in the boundary layer while the upper-troposphere and lower stratosphere box profiles are representative of volcanic SO_2 plumes from effusive and explosive eruptions, respectively.

In order to have more realistic SO_2 profiles for polluted scenes, daily forecasts calculated with the global TM5 chemical transport model (Huijnen et al., 2010) is used. TM5 operates with a spatial resolution of 1°x1° in latitude and longitude, and with 34 sigma pressure levels up to 0.1 hPa in the vertical direction. TM5 uses 3-hourly meteorological fields from the European

Centre for Medium Range Weather Forecast (ECMWF) operational model (ERA-Interim reanalysis data for reprocessing, and the operational archive for real time applications and forecasts). These fields include global distributions of wind, temperature, surface pressure, humidity, and (liquid and ice) water content, and precipitation. A more detailed description of the TM5 model is given in [URL06] and in the NO_2 ATBD [RD11].
For NRT processing, the daily forecast of the TM5 model (located at KNMI) are ingested by the UPAS operational processor. For the calculation of the air mass factors, the profiles are linearly interpolated in space and time, at pixel centre and S5P local overpass time, through a model time step of 30 minutes. To obtain an AMF representative for the troposphere, the integral of the box air mass factor multiplied by the TM5 profile (Eq 5-6) is performed from the ground to the tropopause.

To reduce the errors associated to topography and the lower spatial resolution of the model compared to the TROPOMI spatial resolution, the a priori profiles need to be rescaled to effective surface elevation of the satellite pixel. TM5 surface pressure is converted by applying the hypsometric equation and the assumption that temperature changes linearly with height (Zhou et al., 2009):

$$p_s = p_{TM5}\left(\frac{T_{TM5}}{(T_{TM5} + \Gamma(z_{TM5} - z_s))}\right)^{-\frac{g}{R\Gamma}} \quad \text{5-9}$$

Where $p_{TM5}\ and\ T_{TM5}$ are the TM5 surface pressure and temperature, $\Gamma = 6.5\mathrm{Kkm}^{-1}$ the lapse rate, z_{TM5} the TM5 terrain height, and z_s surface elevation for the satellite ground pixel. The TM5 SO_2 profile is shifted to start at p_s and scaled so that volume mixing ratios are preserved (see Zhou et al., 2009).

As mentioned in section 5.2.3.4, the AMF calculation for the three box profiles is slightly different than for the TM5 modeled profile. The rationale behind this comes from the assumption made of a Lambertian equivant reflector for the clouds which are typically characterized by high reflectivity. Hence, below the cloud, the box air mass factor is zero (no sensitivity to SO_2 or shielding effect) and above the cloud, higher sensitivity (enhancing effect). We anticipate users from the volcanic community to use one of the 3 box VCDs or possibly interpolate the VCDs for a given SO_2 plume height (coming from an external source). While this is reasonable for clear-sky pixels, for a fully cloudy pixel, this approach will fail because of interpolation errors. In particular the SO_2 VCD for the box profile lying below the cloud is undetermined (AMF=0). The proposed workaround is to calculate the box air mass factors, for the cases where the cloud is above the box profile, by forcing the cloudy air mass factor to be equal to the 'above cloud AMF' (i.e., integrate/average m_{cloud} from to the cloud top to the top of atmosphere). The motivation is that if SO_2 is being detected then it is reasonably above (or at) the cloud height and therefore it is sensible to use the corresponding measurement sensitivity factors. One could argue that this choice will possibly lead to an underestimation of the SO_2 VCD and that the VCD is not strictly valid for the box profile anymore. However, it is arguably better doing so than dividing by a small number and end up with unrealistically large VCDs. The users are encouraged to use the box VCDs with cautious and consider the cloud product variables (cloud fraction and cloud height) for their investigations.

5.2.3.7 Temperature correction

The SO_2 absorption cross-sections of Bogumil et al. (2003) show a clear temperature dependence which has an impact on the retrieved SO_2 SCDs depending on the fitting window used. However, only one temperature (203K) is used for the DOAS fit, therefore a temperature correction needs to be applied: SCD'=C_{temp}.SCD. While the SO_2 algorithm provides vertical column results for a set of a-priori profiles, applying this correction to the slant column is not simple and as a workaround it is preferred to apply the correction directly to the AMFs (or box-AMFs to be precise) while keeping the (retrieved) SCD unchanged: AMF'=AMF/$C_{temp.}$ This formulation implicitly assumes that the AMF is not strongly affected by temperature, which is a reasonable approximation (optically thin atmosphere). The correction to be applied require a temperature profile for each pixel (which is obtained from the TM5 model):

$$C_{temp} = 1/[1-\alpha.(T[K]-203)] \qquad 5\text{-}10$$

where α equals 0.002, 0.0038 and 0, for the fitting windows 312-326 nm, 325-335 nm and 360-390 nm, respectively. The parameter α has been determined empirically by fitting eq. 5-10 through a set of data points (Figure 5-6), for each fitting window. Each value in Fig 5-6 is the slope of the fitting line between the SO_2 differential cross-sections at 203K vs the cross-section at a given temperature. In the fitting window 360-390 nm, no temperature correction is applied (α=0) because the cross-sections are quite uncertain. Moreover, the 360-390 nm is meant for extreme cases (strong volcanic eruptions) for SO_2 plumes in the lower-stratosphere where a temperature of 203K is a good baseline.

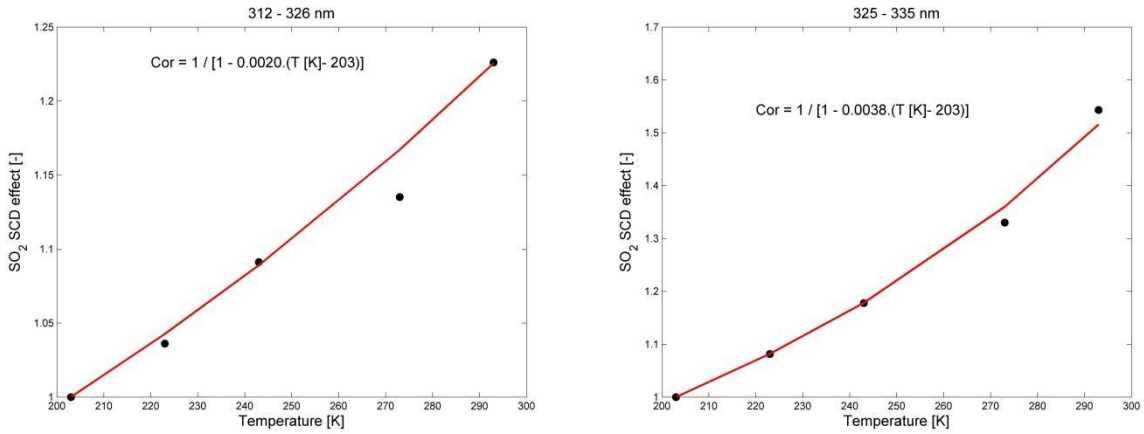

Figure 5-6 *Effect of temperature (relative to 203K) on SO_2 retrieved SCD for fitting windows 312-326 nm (left) and 325-335 nm (right). The red lines show the adopted formulation of C_{temp} (Eq 5.10). Note that, for the 312-326 nm window, the result at 273K has been discarded from the fit as it is seems rather inconsistent with the dependence at other temperatures.*

5.2.3.8 Aerosols

The presence of aerosol in the observed scene (likely when observing volcanic events), may affect the quality of the SO_2 retrieval (e.g. Yang et al., 2010).

No explicit treatment of aerosols (absorbing or not) is foreseen in the NRT algorithm as there is no general and easy way to treat the aerosols effect on the retrieval. At computing time, the aerosol parameters (extinction profile, single scattering albedo,..) are unknown. However, the

information on the AAI is included in the L2 SO_2 files as it gives information to the users on the presence of aerosols both for anthropogenic and volcanic SO_2 and that data should be used/interpreted with care. If, however, reliable information on absorbing aerosol can be obtained from the AAI and the aerosol height product, absorbing aerosols might be included in the forward model, in an offline future version of the product.

5.3 Description of SO_2 detection flag

The recognition of enhanced SO_2 values in operational S5p data is essential in order to detect and monitor volcanic eruptions and anthropogenic pollution sources. Furthermore, it facilitates the assimilation of SO_2 from S5P in order to allow the forecasting of the movement of the volcanic plume.

The detection algorithm has its heritage in the detection algorithm described in Brenot et al. (2014). It has been further developed and successfully applied to the operational GOME-2 products provided in the framework of EUMETSAT AC-SAF. The GOME-2 volcanic detection algorithm was further developed for TROPOMI; even weak emissions can be now recorded and attributed to potential emissions by volcanoes or known anthropogenic sources.

The detection algorithm works as follows:

- Each successfully retrieved pixel with a SZA < 75° is checked whether one of the following conditions apply (hereafter SO_2 VCD stands for the vertical column assuming a box profile at 15km):
    - SO_2VCD > SO_2threshold & SO_2VCD error < 60%
      or
    - SO_2VCD > SO_2threshold & SO_2fitwindow >1

  For SZA < 70° the SO2 VCD threshold is 0.35 DU
  For SZA > 70° the SO2 VCD threshold is 3 DU
  Note that the SO_2 VCD error is the sum of the random and systematic SO_2 error
- If more than 40% of the surrounding pixels within a radius of 15km also fulfill these criteria, the pixel is flagged as 'enhanced SO_2 detection' (flag value =1)
- For high SZA (> 70°), at least 66% of the surrounding pixels must fulfill these criteria. In this case the flag value = 4 to identify possible false-positive detections at high SZA
- Finally, each flagged pixel is checked for vicinity to a known volcanic or anthropogenic pollution source (based on the Fioletov et al. 2016 and Carn et al. 2017 emission catalogues)
    - If the pixel is within a radius of 300km to a known volcano, the flag value = 2 (volcanic detection)
    - If the pixel is within a radius of 100km to a known anthropogenic pollution source, the flag value =3 (pollution detection)

Each pixel in the L2 product can thus have the following flag value:

- 0: No detection
- 1: enhanced SO_2 detection
- 2: enhanced SO_2 detection in vicinity of a known volcano

- 3: enhanced SO_2 detection in vicinity of a known anthropogenic pollution source
- 4: enhanced SO_2 detection at high SZA (>70°), possibly false-positive detection

The flag is stored in the L2 variable sulfurdioxide_detection_flag

Accordingly, the qa_flag values *so2_volcanic_origin_certain_warning* and *so2_volcanic_origin_likely_warning* are set if the flag value is 2 or 3, respectively

6 Input-Output file description

6.1 S5P SO2 product description and size

In addition to the main product results, such as SO_2 slant column, vertical column and air mass factor, the level 2 data files contain several additional parameters and diagnostic information. Table 6-1 gives a minimum set of data fields present in the Level 2 data. A 1-orbit SO_2 column Level 2 file is about 740 MB (up to 950 MB from 6 August 2019 on). More details about the level 2 data format are provided in the Product User Model [RD16**Error! Reference source not found.**].

Table 6-1 *List of output fields in the TROPOMI SO_2 products. nAlong x nAcross mean the number of pixels in an orbit along track and across track, respectively.*

Name/Data	Symbol	Unit	Description	Data type	Number of entries per observation
Date		n.u.	Date and time of the measurement YYMMDDHHMMSS.MS	characters	nAlong
Latitudes	*lat*	degree	Latitudes of the four pixel corners + center	float	5 x nAlong x nAcross
Longitudes	*lon*	degree	Longitudes of the four pixel corners + center	float	5 x nAlong x nAcross
SZA	θ_0	degree	Solar zenith angle	float	nAlong x nAcross
VZA	θ	degree	Viewing zenith angle	float	nAlong x nAcross
RAA	φ	degree	Relative azimuth angle	float	nAlong x nAcross
SCD	N_s	mol.m^{-2}	SO2 slant column density	float	nAlong x nAcross
SCDcorr	N_s^c	mol.m^{-2}	SO2 slant column density background corrected	float	nAlong x nAcross
VCD	N_v	mol.m^{-2}	SO2 vertical column density (4values)	float	4 x nAlong x nAcross
Wdow flag	*Wflag*	n.u.	Flag for the fitting window used (1,2,3)	integer	nAlong x nAcross
Detection flag	*Detection_flag*	n.u.	Flag for the detection of enhanced SO2 VCD (0,1,2,3,4)	integer	nAlong x nAcross
AMF	*M*	n.u.	Air mass factor (4values)	float	4 x nAlong x nAcross
Cloud free AMF	M_{clear}	n.u.	Cloud Free Air mass factor (4values)	float	4 x nAlong x nAcross

Cloudy AMF	M_{cloud}	n.u.	Fully Cloudy Air mass factor (4values)	float	4 x nAlong x nAcross
CF	f_c	n.u.	Cloud fraction	float	nAlong x nAcross
CRF	Φ	n.u.	Cloud radiance fraction	float	nAlong x nAcross
CP	p_{cloud}	Pa	Cloud top pressure	float	nAlong x nAcross
CH	z_{cloud}	m	Cloud top height	float	nAlong x nAcross
CA	A_{cloud}	n.u.	Cloud top albedo	float	nAlong x nAcross
Albedo	A_S	n.u.	Surface albedo	float	nAlong x nAcross
Aerosol index	AAI	n.u.	Absorbing Aerosol Index	float	nAlong x nAcross
Chi-squared	Chi^2	n.u.	Chi-squared of the fit	float	nAlong x nAcross
VCD error	σ_N_v	mol.m^{-2}	Total error on the vertical column (individual measurement)	float	4x nAlong x nAcross
SCD random error	$\sigma\_N_{s\_rand}$	mol.m^{-2}	Random error on the slant column	float	nAlong x nAcross
SCD systematic error	$\sigma\_N_{s\_syst}$	mol.m^{-2}	Systematic error on the slant column	float	nAlong x nAcross
AMF random error	σ_M_{rand}	n.u.	Random error on the air mass factor (4values)	float	4x nAlong x nAcross
AMF systematic error	σ_M_{syst}	n.u.	Systematic error on the air mass factor (4 values)	float	4x nAlong x nAcross
AMF systematic error kernel	$\sigma\_M_{syst\_kernel}$	n.u.	Systematic error on the air mass factor if averaging kernels are used (4 values)	float	4x nAlong x nAcross
Averaging kernel	*AK*	n.u.	Total column averaging kernel (for a-priori profile from CTM)	float	34 x nAlong x nAcross
Averaging kernel scalings for box profiles	*Scaling box*	n.u.	Factors to apply to the averaging kernel function to obtain the corresponding averaging kernels for the 3 box profiles	float	3x nAlong x nAcross
SO_2 profile	n_a	n.u.	A priori profile from CTM (volume mixing ratio)	float	34 x nAlong x nAcross
Surface altitude	z_s	m	Digital elevation map	float	nAlong x nAcross

Surface pressure	p_s	Pa	Effective surface pressure of the satellite pixel	float	nAlong x nAcross
TM5 level coefficient a	A_i	Pa	TM5 pressure level coefficients that effectively define the mid-layer levels (from ECMWF)	float	24
TM5 level coefficient b	A_i	n.u.		float	24
TM5 tropopause layer index	n.u.	n.u.	Index of the layer defined as the tropopause limit in the TM5 vertical profiles.	Int	nAlong x nAcross

It should be noted that the averaging kernels are given only for the a priori profiles from the TM5 CTM (to save space). The averaging kernels for the box profiles can be estimated by scaling the provided averaging kernel (corresponding to TM5 profiles): $AK_{box}(p) = AK(p)$.Scaling box. Following the AK formulation of Eskes and Boersma (2004), the scaling factor is given simply by AMFs ratios: AMF_{TM5}/AMF_{box}.

We note that, strictly, the derived averaging kernels for the box profiles are not fully defined because AMF_{box} is not consistently calculated by integrating the box air mass factor. However, if a modeled SO_2 profile is convolved with AK_{box}, the resulting SO_2 VCD can directly be compared to the corresponding measured SO_2 VCD box.

6.2 Auxiliary information needs

The algorithm relies on several external data sets. These can be either static or dynamic. An overview is given in Table 6- and Table 6- below

6.2.1 Static data

See also the documentS5P/TROPOMI Static input for Level 2 processors [RD15]

Table 6-2 *Static auxiliary data for the algorithm.*

Name/Data	Symbol	Unit	Source	Pre-process needs	Comments
Absorption cross-sections					
SO2	σ_{SO2}	$cm^2molec.^{-1}$	Bogumil et al. (2003), 203K, 223K, 243K, 293K Hermans et al. (2009), all temperatures	Convolution at the instrumental spectral resolution using the provided slit function	
Ozone	σ_{o3218} σ_{o3243}	$cm^2molec.^{-1}$	Brion et al. (1998) ; 218K and 243K.		

BrO	σ_{BrO}	$cm^2molec.^{-1}$	Fleischmann et al. (2004), 223K		-
NO_2	σ_{NO2}	$cm^2molec.^{-1}$	Vandaele et al. (1998), 220K		
O_4 (O_2-O_2)	σ_{O4}	$cm^5molec.^{-2}$	Greenblatt et al. (1990)		
High resolution reference solar spectrum	E_s	$W\ m^{-2}nm^{-1}$	Chance and Kurucz, 2010	-	-
Ring effect	$\sigma_{ringev1}$ $\sigma_{ringev2}$	$cm^2molec.^{-1}$	2 Ring cross-sections generated internally.	A high-resolution reference solar spectrum and the instrument slit function are needed to generate the data set.	Calculated in an ozone containing atmosphere for low and high SZA, using LIDORT_RRS (Spurr et al., 2008b) and a standard atmosphere (Camelot European Pollution atmospheric profile).
Non-linear O_3 absorption effect	σ_{o3l} σ_{o3sq}	$nm.cm^2molec.^{-1}$ $cm^4molec.^{-2}$	2 pseudo-cross sections generated internally.	The O_3 cross-section at 218 K is needed.	Calculated from the Taylor expansion of the wavelength and the O_3 optical depth (Puķīte et al., 2010).
Instrument slit function	SF	n.u.	Slit Function by wavelength/detector.	-	Values between 300 and 400nm.
Surface Albedo	A_s	n.u.	OMI-based monthly minimum LER (update of Kleipool et al., 2008)	-	
Digital elevation map	z_s	m	GMTED2010 (Danielson et al., 2011)		Average over the ground pixel area.
SO2 profile	n_a	n.u.	One kilometer thick box profiles, with three different peak altitudes, representing different altitude regimes: Boundary layer: from the surface altitude to 1km above it Free troposphere: centered around 7 km altitude.	-	TM5 profiles from the last available day in case theTM5 profiles of the current day are not available Note 1: for the different fitting windows (312-326 nm, 325-335 nm, 360-390 nm), the assumed vertical column is 5 DU, 100 DU, 500 DU, respectively.

			Lower stratosphere: centered around 15 km altitude. Daily SO_2 profiles forecast from TM5		
Look-up table of pressure-resolved box-AMFs	m	n.u.	Calculated internally with the LIDORTv3.3 RTM (Spurr et al., 2008a).	-	-
Temperature correction parameters	α	K^{-1}	Bogumil et al. (2003),	-	-

6.2.2 Dynamic data

Table 6-3 *Dynamic auxiliary data for the algorithm*.

Name/Data	Symbol	Unit	Source	Pre-process needs	Backup if not available
S5P level 1B Earth radiance	I	mol s^{-1} $m^{-2}nm^{-1}sr^{-1}$	S5P L1b product	-	No retrieval
S5P level 1B sun irradiance	E_0	mol s^{-1} $m^{-2}nm^{-1}$	S5P L1b product	Wavelength recalibrated using a high-resolution reference solar spectrum	Use previous measurement
Cloud fraction	f_c	n.u.	S5 P operational cloud product based on a Lambertian cloud model [RD10]. UPAS processor.	-	No retrieval
Cloud top pressure	p_{cloud}	Pa			
Cloud top albedo	A_{cloud}	n.u.			
SO2 profile	n_a	n.u.	Daily forecast from TM5 CTM run at KNMI.	-	Use TM5 CTM profile from last available day

Temperature profile	T	K	Daily forecast from TM5 CTM run at KNMI.	-	Use TM5 CTM profile from last available day
Absorbing aerosol index	AAI	n.u.	S5P operational AAI product [RD09]. Used in offline mode only. KNMI processor.	-	Missing information flag.
Snow-ice flag		n.u.	Near real-time global Ice and Snow Extent (NISE) data from NASA.	-	Use snow/ice climatology.

7 Error analyses

7.1 Introduction

The total uncertainty (accuracy and precision) on the SO_2 column products is composed of many sources of error (see also e.g., Lee et al., 2009). Several of them are related to the measuring instrument, such as uncertainties due to noise or knowledge of the slit function. These instrumental errors propagate into the uncertainty on the slant column. Other types of error can be considered as model errors and are related to the representation of the physics in the algorithm. Examples of model errors are uncertainties on the trace gas absorption cross-sections and the treatment of clouds. Model errors can affect the slant column results or the air mass factors.

The total retrieval uncertainty on the SO_2 vertical columns can be derived by error propagation, starting from Eq. 5-1 and if one assumes uncorrelated retrieval steps (Boersma et al., 2004; De Smedt et al., 2008):

$$\sigma_{N_V}^2 = \left(\frac{\sigma_{N_S}}{M}\right)^2 + \left(\frac{\sigma_{N_S^{\text{back}}}}{M}\right)^2 + \left(\frac{\left(N_S - N_S^{\text{back}}\right)\sigma_M}{M^2}\right)^2 \quad \text{7-1}$$

where σ_{Ns} and σ_{Ns}^{back} are the errors on the slant column N_S and on the background correction N_S^{back}, respectively.

The error analysis is complemented by the total column averaging kernel (AK) as described in Eskes and Boersma (2003):

$$AK(p) = \frac{m'(p)}{M} \quad \text{7-2}$$

which is if often used to characterize the sensitivity of the retrieved column to a change in the true profile.

7.2 Error components

The following sections describe and characterize 20 error contributions to the total SO_2 vertical column uncertainty. These different error components and their values (where possible) are summarized in Tables 7-1 and 7-2.

A difficulty in the error formulation presented above comes from the fact that it assumes the different error sources/steps of the algorithm to be independent and uncorrelated, which is not strictly valid. For example, the background correction is designed to overcome systematic features/deficiencies of the DOAS slant column fitting and these two steps cannot be considered as independent. Hence, summing up all the corresponding error estimates would lead to overestimated error bars. Therefore, several error sources will be discussed in the following sub-sections without giving actual values at this point. Their impact is included and described in later sub-sections.

Another important point to note is that one should also (be able to) discriminate systematic and random components of a given error source V:

$$\sigma_V^2 = \frac{\sigma_{V(rand)}^2}{n} + \sigma_{V(syst)}^2 \qquad 7\text{-}3$$

here *n* is the number of pixels considered. However, they are hard to separate in practice. Therefore, each of the 20 error contributions are (tentatively) qualified as either “random” or “systematic” errors, depending on their tendencies to average out in space/time or not.

7.2.1 Errors on the slant column

Error sources that contribute to the total uncertainty on the slant column originate both from instrument characteristic and uncertainties/limitations on the representation of the physics in the DOAS slant column fitting algorithm. For the systematic errors on the slant column, the numbers provided in Table 7-1 have been determined based on sensitivity tests (using the QDOAS software).

All effects summed in quadrature, the various contributions are estimated to account for a systematic error of about 20% +0.2DU of the background-corrected slant column ($\sigma_{N_s,syst} = 0.2 * (N_s - N_s^{back})$+0.2DU).

For the random component of the slant column errors, the error on the slant columns provided by the DOAS fit (hereafter referred to as SCDE) as it is assumed to be dominated by and representative for the different random sources of error.

1) SO_2 cross-section

Systematic errors on slant columns due to SO_2 cross-sections uncertainties are estimated to be around 6% (Vandaele et al., 2009) in window 1 (312-326 nm) and window 2 (325-335 nm) and unknown in window 3 (360-390 nm). In addition, the effect of the temperature on the SO_2 cross-sections has to be considered as well. Using cross-sections at different temperature than the DOAS baseline (Bogumil at 203K) has an effect on the SCD results. We refer to see section 7.2.2 for a discussion of this source of error.

2) SO_2 and O_3 absorption

Non-linear effects due to SO_2 and O_3 absorption are to a large extent accounted for using the triple windows retrievals and the Taylor expansion of the O_3 optical depth (Pukīţe et al., 2010). Remaining systematic biases are then removed using the background correction; hence residual systematic features are believed to be small (please read also the discussion on errors 9 and 10). The random compenent of the slant column error contributes to SCDE.

Non-linear effects due to SO_2 absorption itself (mostly for volcanic plumes) are largely handled by the triple windows retrievals but the transition between the different fitting windows is a compromise and there are cases where saturation can still lead to rather large uncertainties. However, those are difficult to assess on a pixel to pixel basis.

3) Other atmospheric absorption /interferences

In some regions on Earth, several systematic features in the slant columns remain after the background correction procedure (see discussion on error 9: background correction error) and are attributed to spectral interferences not fully accounted for in the DOAS analysis, such as incomplete treatment of the Ring effect.This effect has also a random component and contributes to the retrieved SCD error (SCDE).

4) Radiance shot noise

It has a major contribution to SCDE and it can be estimated from typical S/N values of S5P in UV band 3 (800-1000, according to Veefkind et al., 2012). This translates to typical SCD detection limits of about 0.3 -0.5, 5 and 60 DU for window 1, 2 and 3, respectively.

5) DOAS settings

Tests on the effect of changing the lower and upper limits of the fitting windows by 1 nm and the order of the closure polynomial (4 instead of 5) have been performed. Based on a selection of orbits for the Kasatochi eruption (wide range of measured SCDs), the corresponding SCD errors are less than 11, 6 and 8 % for window 1, 2 and 3, respectively.

6) Wavelength and radiometric calibration

Tests on the effect of uncertainties in the wavelength calibration have been performed in the ESA CAMELOT study. The numbers are for a shift of 1/20th of the spectral sampling in the solar spectrum and 1/100th of the spectral sampling in the Earthshine spectrum. The shift can be corrected for, but interpolation errors can still lead to a remaining uncertainty of a few percent.
Regarding radiometric calibration, the retrieval result is in principle insensitive to flat (spectrally constant) offsets on the measured radiance because the algorithm includes an intensity offset correction. From the ESA ONTRAQ study it was found that additive error signals should remain within 2% of the measured spectrum.

7) Spectral response function

Uncertainties in the instrumental slit functions can lead to systematic errors on the retrieved SO_2 slant columns. The instrumental slit function has been investigated and results have been optimized during commissioning phase. Overall, the instrumental slit function is very well characterized and this error source is believed to be small. However, it is needed to monitor the stability of the instrument and check for changes in slit function during the full TROPOMI mission, and possibly adopt mitigation strategies if needed.

8) Other spectral features

Unknown or untreated instrumental characteristics such as stray light and polarization sensitivity can introduce spectral features that may lead to bias in the retrieved slant column data. To certain extend these can be prevented by the DOAS polynomial and the intensity offset correction settings, as long as the perturbing signals are a smooth function of wavelength. Conversely, high-frequency spectral structures can have potentially a large impact on SO_2 retrievals depending on their amplitude and whether they interfere with SO_2 absorption structures. At the time of writing, it is hard to evaluate the impact of these measurement errors (if any).

In the ONTRAQ study, testing sinusoidal perturbation signals showed that effect on the retrieval result depends strongly on the frequency of the signal. Additives signals with an

amplitude of 0.05 % of the measurement affect the retrieved SO_2 slant column up to 30%. The effect scales more or less linearly with the signal amplitude.

9) Background correction

This error source is mostly systematic and important for anthropogenic SO_2 or for monitoring degassing volcanoes. Based on OMI and GOME-2 test retrievals, the uncertainty on the background correction is estimated to be < 0.2 DU. This value accounts for limitations of the background correction and is compatible with residual slant columns values typically found (after correction) in some clean areas (e.g. above the Sahara), or for a possible contamination by volcanic SO_2, after a strong eruption.

Table 7-1 *Systematic and random error components contributing to the total uncertainty on the SO_2 slant column*

	Error source	**Type ***	**Parameter uncertainty**	**Typical uncertainty on SO_2 SCD**
1	SO_2 absorption cross section	S	6% (window 1) 6% (window 2) unknown (window 3)	6%
2	SO_2 and O_3 absorption	S & R		Errors 9 & 10
3	Other atmospheric absorption or interference	S & R		Error 9
4	Radiance shot noise	R	S/N=800-1000	0.3-0.5 DU (window 1) 5 DU (window 2) 60 DU (window 3)
5	DOAS settings	S	1 nm, polynomial order	<11% (window 1) <6% (window 2) <8% (window 3)
6	Wavelength and radiometric calibration	S	Wavelength Calibration: Radiometric calibration.	Wavelength calibration and spectral shifts be corrected by the algorithm to less than 5 % effect on the slant column.

			additive errors should remain below 2 %.	Intensity offset correction in principle treats (small) radiometric calibration errors
7	Spectral response function		-	TROPOMI-specific Uncertainty: few %
8	Other spectral features		Strongly dependent on interfering signal	-
9	Background correction	S & R		0.2 DU

* R: random, S: systematic

7.2.2 Errors on the air mass factor

The error estimates on the AMF are listed in Table 7-2 and are based on simulations and closed-loop tests using the radiative transfer code LIDORT. One can identify two sources of errors on the AMF. First, the adopted LUT approach has limitations in reproducing the radiative transfer in the atmosphere (forward model errors).

Secondly, the error on the AMF depend on input parameter uncertainties. This contribution can be broken down into the squared sum (Boersma et al., 2004):

$$\sigma_M^2 = \left(\frac{\partial M}{\partial \text{alb}} \cdot \sigma_{\text{alb}}\right)^2 + \left(\frac{\partial M}{\partial \text{ctp}} \cdot \sigma_{\text{ctp}}\right)^2 + \left(\frac{\partial M}{\partial feff} \cdot \sigma_{feff}\right)^2 + \left(\frac{\partial M}{\partial s} \cdot \sigma_s\right)^2 \quad \text{7-4}$$

where σ_{alb}, σ_{ctp}, σ_{feff}, σ_s are typical uncertainties on the albedo, cloud top pressure, cloud fraction and profile shape, respectively.

The contribution of each parameter to the total air mass factor error depends on the observation conditions. The air mass factor sensitivities ($\frac{\partial M}{\partial parameter}$), i.e. the air mass factor derivatives with respect to the different input parameters, can be derived for any particular condition of observation using the altitude-dependent AMF LUT, created with LIDORTv3.3, and using the a priori profile shapes. In practice, a LUT of AMF sensitivities has been created using reduced grids from the AMF LUT and a parametrisation of the profile shapes based on the profile shape height.

10) AMF wavelength dependence

Sensitivity tests have been performed and showed that AMF calculations at 313, 326 and 375 nm (for window 1, 2 and 3, respectively) provide reasonable results and compensate to some extent for non–linear spectral features in DOAS. Figure 7-1 gives an illustration of these sensitivity tests in the baseline window and shows an excellent correlation and slope close to 1 for the scatter plot of retrieved versus simulated slant columns using an effective wavelength of 313 nm for the AMF. Overall, for low solar zenith angles the deviations from the truth, observed in the closed-loop retrievals, remain less than 5% in most cases, except for boundary layer (BL) SO_2 at a 1 DU column level and for low albedo scenes (deviations up to 20%). For high solar zenith angles deviations are less than 10% in most cases; again: except for BL SO_2 at a 1 DU column level and for low albedo scenes (underestimation up to a factor of 2).

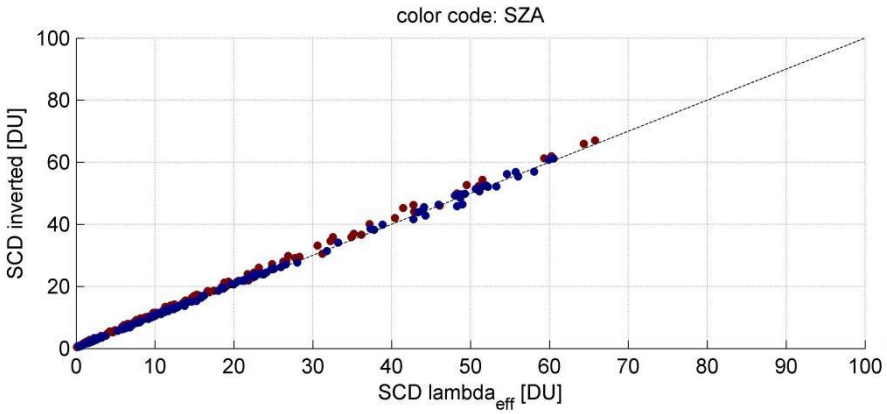

Figure 7-1 *SO_2 slant columns retrieved versus simulated at an effective wavelength of 313 nm from synthetic spectra (SZA: 30°/70°) in the spectral range 312-326 nm and for SO2 layers in the boundary layer, upper troposphere and lower stratosphere. The different points correspond to different values for the line-of-sight angle (0, 45°), surface albedo (0.06, 0.8), surface height (0, 5 km) and total ozone column (350, 500 DU). SO_2 vertical columns as input of the RT simulations are maximum of 25 DU.*

11) Model atmosphere

This error relates to uncertainties in the atmospheric profiles used as input of LIDORT for the box air mass factor look-up-table calculations.

Although the effect of O_3 absorption on the AMF is treated in the algorithm, the O_3 profiles used as input of LIDORT are not fully representative of the real profiles and typical error (including error due to interpolation) of 5-10% can occur.

A test has been achieved by replacing the US standard atmosphere pressure and temperature profiles by high latitude winter profiles and the impact on the results is found small.

Table 7-2 *Systematic and random error components contributing to the total uncertainty on the SO_2 air mass factor.*

	Error	Type *	Parameter uncertainty	Typical uncertainty on the AMF
10	AMF wavelength dependence	S		10%
11	model atmosphere	S	O_3 profile P,T profiles (see text)	~5-10% <5% (troposphere) <1% (stratosphere) small
12	Forward model	S	< 5%	<5%
13	Surface albedo†	S	0.02	15% (PBL) 5% (FT) 1% (LS)
14	Cloud fraction†	R	0.05	5% (PBL) 15% (FT) 1% (LS)
15	Cloud top pressure†	R	50 hPa (~0.5 km)	50% (PBL) 50% (FT) 1% (LS)
16	Cloud correction	R		< 5% on yearly averaged data
17	Cloud model		TBD	
18	SO_2 profile shape	S	50 hPa	anthropogenic SO_2 20%-35% volcanic SO_2 large (low albedo), < 50% (high albedo)
19	Aerosol	S & R		Anthropogenic SO_2 ~ 15% (Nowlan et al., 2011). Volcanic SO_2 (aerosols: ash/sulphate) : ~ 20% (Yang et al., 2010)
20	Temperature correction	R		~5%

* R: random, S: systematic) † Effect on the AMF estimated from Figure 7-2

12) Forward model

It is expected to be small: less than 5% (Hendrick et al., 2006; Wagner et al., 2007).

13) Surface albedo

A typical uncertainty on the albedo is of 0.02 (Kleipool et al., 2008). This translates to an error on the air mass factor using the slope of the air mass factor as a function of the albedo (Eq.7-4) and can be evaluated for each satellite pixel. As an illustration, Figure 7-2 shows the typical and expected dependence of the AMF with albedo but also with the cloud conditions. From Figure 7-2a, one can conclude that the retrievals of SO_2 in the BL is much more sensitive to the exact albedo than for SO_2 higher up in the atmosphere, for this particular example.

More substantial errors can be introduced if the real albedo differs considerably from what is expected, for example in the case of the sudden snowfall or ice cover. A Snow/ice cover flag is therefore used to identify such cases.

14) Cloud fraction

An uncertainty on the cloud fraction of 0.05 is considered. The corresponding AMF error can be estimated through Eq.7-4 (see Figure 7-2b) or by analytic derivation from Eqs. 5.6-5.8.

15) Cloud top pressure

An uncertainty on the cloud top height of 0.5 km (~50 hPa) is assumed. The corresponding AMF error can be estimated through Eq. 7-4. Figure 7.2c illustrates the typical behaviour of signal amplification /shielding for a cloud below/ above the SO_2 layer. One can see that the error (slope) dramatically increases when the cloud is at a height similar as the SO_2 bulk altitude.

16) Cloud correction

Sensitivity tests showed that applying the independent pixel approximation or assuming cloud-free pixels makes a difference of only 5% on yearly averaged data (for anthropogenic BL SO_2 VC with cloud fractions less than 40%).

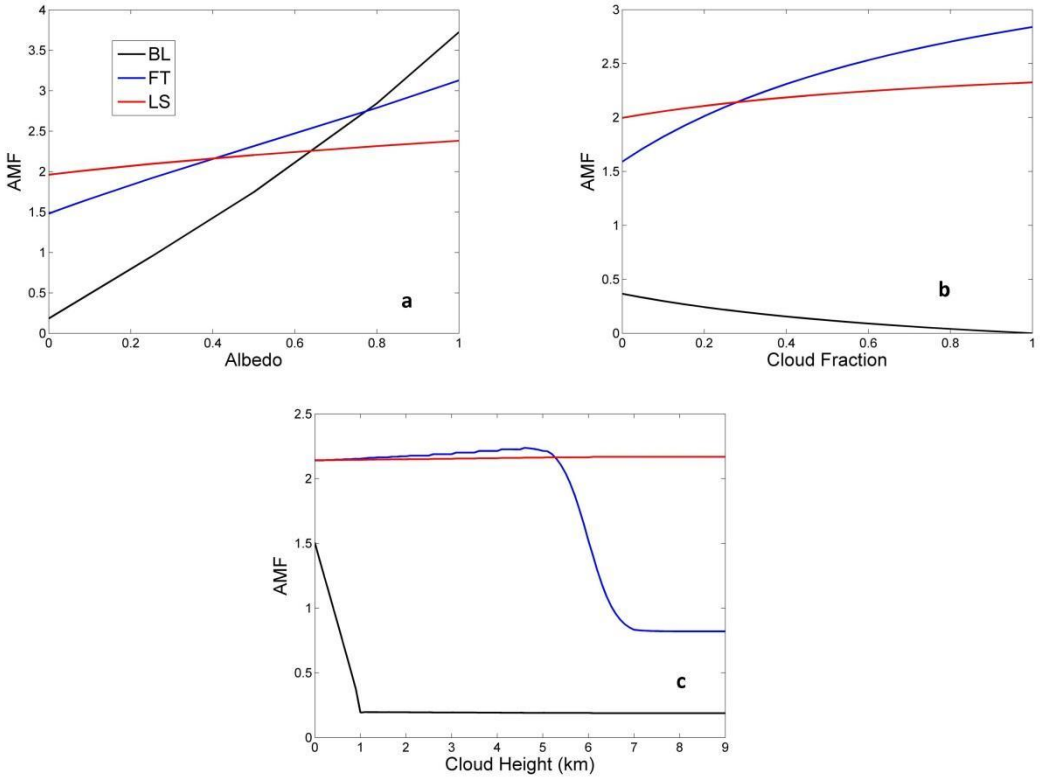

Figure 7-2 *Air mass factors at 313 nm for SO2 in the boundary layer (BL:0-1 km), free-troposphere and lower stratosphere (FT, LS: Gaussian profiles with maximum height at 6,15 km and FWHM: 1 km). Calculations are for SZA=40°, Los=10°, RAA=0° and surfh=0 km. AMFs are displayed as a function of the (a) albedo for clear-sky conditions, (b) cloud fraction for alb=0.06 and cth=2km and (c) cloud top height for alb=0.06 and cf=0.3.*

17) Cloud model

Cloud As Layer (CAL) is the baseline of the cloud algorithm, but a Lambertian Equivalent Reflector (LER) implementation is used for NO2, SO2 and HCHO retrievals. The error due to the choice of the cloud model will be evaluated during the operational phase.

18) Profile shape

A major source of systematic uncertainty for virtually every SO_2 scene is the shape of the vertical SO_2 distribution. The corresponding AMF error can be calculated through Eq. 7-4 and estimation of uncertainty on the profile shape. Note that vertical columns are provided with their averaging kernels, so that vertical column data might be improved for particular locations by using more accurate SO_2 profile shapes based on input from models or observations. For this reason, two values for the systematic uncertainty on the AMF are given in the output file, one considering all systematic error sources in Eq. 7-4 and the other neglecting the error from the SO_2 profile shape (the data fields are easily identified by the suffix '_kernel' in the output file).

For anthropogenic SO_2 under clear-sky, sensitivity tests using a box profile from 0 to 1±0.5 km above ground level, or using the different profiles from the CAMELOT study [RD07], give differences in AMFs in the range of 20-35%. Note that for particular conditions SO_2 may also be uplifted above the top of the boundary layer and sometimes reach upper-tropospheric levels (e.g., Clarisse et al., 2011). SO_2 box air mass factors displayed in Figure 5-5 show that the measurement sensitivity is then increased up to factor of 3 and therefore constitutes a major source of error.

In the SO_2 algorithm, the uncertainty on the profile shape is estimated using one parameter describing the shape of the TM5 profile: the profile height, i.e. the altitude (pressure) below which resides 75% of the integrated SO_2 profile. $\frac{\partial M}{\partial s}$ is approached by $\frac{\partial M}{\partial s_h}$ where s_h is half of the profile height. Relatively small variations of this parameter have a strong impact on the total air mass factors for low albedo scenes, because altitude-resolved air mass factors decrease strongly in the lower troposphere, where the SO_2 profiles peak (see e.g. Figure 5-5). The error due to the profile shape uncertainty is estimated by taking σ_{sh}=50 hPa.

For volcanic SO_2, the effect of the profile shape uncertainty depends on the surface or cloud albedo. For low albedo scenes (Fig 5-5 a), if no external information on the SO_2 plume height is available, it is a major source of error at all wavelengths. Vertical columns may vary up to a factor of 5. For high albedo scenes (Fig 5-5 b), the error is less than 50%. It should be noted that these conditions are often encountered for strong eruptions injecting SO_2 well above the cloud deck (high reflectivity). Further uncertainty on the retrieved SO_2 column may arise if the vertical distribution shows distinct layers at different altitude, due do the different nature of successive phases of the eruption.

In the SO_2 algorithm, three 1km thick box profiles are used in the AMF calculation mostly to represent typical volcanic SO_2 profiles. The error due to the profile shape uncertainty is estimated by varying the box centre levels by 50 hPa.

19) Aerosols

The effect of aerosols on the air mass factors are not explicitly considered in the SO_2 retrieval algorithm. To some extent, however, the effect of the non-absorbing part of the aerosol extinction is implicitly included in the cloud correction (Boersma et al., 2004). Indeed, in the presence of aerosols, the cloud detection algorithm is expected to overestimate the cloud fraction, resulting partly in a compensation effect for cases where aerosols and clouds are similar heights. Absorbing aerosols have a different effect on the air mass factors, and can lead to significant errors for high AODs. In the TROPOMI SO_2 product, the absorbing aerosol index field can help identifying observations with elevated absorbing aerosols.

Generally speaking, the effect on AMF is highly variable and strongly dependent on aerosols properties (AOD, height and size distribution, single scattering albedo, scattering phase function, etc.). Typical AMFs uncertainties due to aerosols found in the literature are given in Table 7-2. As aerosols affect cloud fraction, cloud top height and to some extend the albedo database used, correlations between uncertainties on these parameters are to be expected.

20) Temperature correction

Differences to the temperature of the DOAS baseline cross section (Bogumil at 203K) have an effect on the SCD results. This effect is in principle accounted for in the temperature correction (which is applied in practice to the AMFs, see section 5.2.3.7) but with a certain error associated to it (~5%).

8 Validation

The validation of TROPOMI SO_2 product is covered by various initiatives including the TROPOMI Mission Performance Centre. The reader is referred to the MPC validation report:

http://mpc-vdaf.tropomi.eu/index.php/sulphur-dioxide

9 Conclusions

Based on the heritage from GOME, SCIAMACHY, GOME-2 and OMI, a DOAS retrieval algorithm has been developed for the operational retrieval of SO_2 vertical columns from TROPOMI Level1b measurements in the UV spectral range. Here we describe its main features.

In addition to the traditionally used fitting window of 312-326 nm, the new algorithm allows for the selection of two additional fitting windows (325-335 nm and 360-390nm), reducing the risk of saturation and ensuring accurate SO_2 column retrieval even for extreme SO_2 concentrations as observed for major volcanic events. The spectral fitting procedure also includes an advanced wavelength calibration scheme and a spectral spike removal algorithm.

After the slant column retrieval, the next step is a background correction, which is empirically based on the O_3 slant column (for the baseline fitting window) and across-track position, and accounts for possible across-track dependencies and instrumental degradation.

The SO_2 slant columns are then converted into vertical columns by the means of air mass factor calculations. The latter is based on box air mass factor look-up-tables with dependencies on the viewing geometry, clouds, surface pressure, albedo, ozone, and is applied to pre-defined box profiles and TM5 CTM forecast profiles. In addition, the algorithm computes DOAS-type averaging kernels and a full error analysis of the retrieved columns.

Based on the first two years of data, the TROPOMI SO_2 algorithm has a comparable level of accuracy than its predecessors. In terms of retrieval precision, TROPOMI is not as good as OMI on the base of individual pixels but is better for averaged data due to a much finer spatial resolution of 5.5x3.5 km² (at best). For single measurements, the user requirements for tropospheric SO_2 concentrations is not met, but improved monitoring of strong pollution and volcanic events is possible by spatial and temporal averaging the increased number of observations of TROPOMI. Nevertheless, it require significant validation work which is not covered by the present document.

References

Afe, O. T., Richter, A., Sierk, B., Wittrock, F., and Burrows, J.P.: BrO emissions from volcanoes: a survey using GOME and SCIAMACHY measurements, Geophys. Res. Lett., 31, L24113, 2004.

Barbu,A.L., Segers, A.J., Schaap, M., Heemink, A.W., and Builtjes, P.J.H.: A multi-component data assimilation experiment directed to sulphur dioxide and sulphate over Europe, Atm. Env., Vol. 43, 9, P1622-1631, 2009.

Beirle, S., Hörmann, C., Penning de Vries, M., Dörner, S., Kern, C., and Wagner, T.: Estimating the volcanic emission rate and atmospheric lifetime of SO_2 from space: a case study for Kīlauea volcano, Hawai, Atmos. Chem. Phys., 14, 8309-8322, doi:10.5194/acp-14-8309-2014, 2014.

Bobrowski, N., Kern, C., Platt, U., Hörmann, C., and Wagner, T.: Novel SO_2 spectral evaluation scheme using the 360–390 nm wavelength range, Atmos. Meas. Tech., 3, 879-891, doi:10.5194/amt-3-879-2010, 2010.

Boersma, K. F., Eskes, H. J., and Brinksma, E. J.: Error analysis for tropospheric NO_2 retrieval from space, J. Geophys. Res., 109, D04311, doi: 10.1029/2003JD003962, 2004.

Bogumil, K., J. Orphal, T. Homann, S. Voigt, P. Spietz, O. Fleischmann, A. Vogel, M. Hartmann, H. Bovensmann, J. Frerick, and J. Burrows: Measurements of molecular absorption spectra with the SCIAMACHY Pre-Flight Model: instrument characterization and reference data for atmospheric remote-sensing in the 230--2380 nm region, Journal of Photochemistry and Photobiology A, 157, 167-184, 2003.

Brenot, H., Theys, N., Clarisse, L., van Geffen, J., van Gent, J., Van Roozendael, M., van der A, R., Hurtmans, D., Coheur, P.-F., Clerbaux, C., Valks, P., Hedelt, P., Prata, F., Rasson, O., Sievers, K., and Zehner, C.: Support to Aviation Control Service (SACS): an online service for near real-time satellite monitoring of volcanic plumes, Nat. Hazards Earth Syst. Sci., 14, 1099-1123, doi:10.5194/nhess-14-1099-2014, 2014.

Brion, J., et al.: Absorption spectra measurements for the ozone molecule in the 350-830 nm region, J. Atmos. Chem., 30, 291-299, 1998.

Carn, S.A. and T.M. Lopez, Opportunistic validation of sulfur dioxide in the Sarychev peak volcanic eruption cloud, Atmos. Meas. Tech., 4, 1705-1712, 2011.

Carn, S.A., V.E. Fioletov, C.A. McLinden, C. Li, and N.A. Krotkov (2017), A decade of global volcanic SO_2 emissions measured from space, *Sci. Rep.*, 7, 44095; doi:10.1038/srep44095.

Chance, K. and R. J. Spurr: Ring effect studies: Rayleigh scattering including molecular parameters for rotational Raman scattering, and the Fraunhofer spectrum, Applied Optics, 36, 5224-5230, 1997.

Chance, K. and Kurucz, R. L.: An improved high-resolution solar reference spectrum for earth's atmosphere measurements in the ultraviolet, visible, and near infrared, J. Quant. Spectrosc. Radiat. Transf., 111(9), 1289-1295, 2010.

Clarisse, L., Fromm, M., Ngadi, Y., Emmons, L., Clerbaux, C., Hurtmans, D., and Coheur, P.-F.: Intercontinental transport of anthropogenic sulfur dioxide and other pollutants; an infrared remote sensing case study, Geophys. Res. Lett., 38, L19806, doi:10.1029/2011GL048976, 2011.

Danckaert, T., Fayt, C., Van Roozendael, M., De Smedt, I., Letocart, V., Merlaud, A., Pinardi, G: Qdoas Software User Manual, Version 2.1, http://uv-vis.aeronomie.be/software/QDOAS/QDOAS_manual_2.1_201212.pdf, 2012.

Danielson, J.J., and Gesch, D.B.: Global multi-resolution terrain elevation data 2010 (GMTED2010): U.S. Geological Survey Open-File Report 2011–1073, 26 p, 2011.

De Smedt, I., J.-F. Müller, T. Stavrakou, R. van der A, Eskes, H., and Van Roozendael, M ., Twelve years of global observation of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, Atmos. Chem. Phys., 8, 4947-4963, 2008.

Eisinger, M., and Burrows, J.P.: Tropospheric sulfur dioxide observed by the ERS-2 GOME instrument, Geophys. Res. Lett., Vol. 25, pp. 4177-4180, 1998.

Eskes, H. J. and Boersma, K. F.: Averaging kernels for DOAS total column satellite retrievals, Atmos. Chem. Phys., 3, 1285–1291, 2003.

Faloona, I., Sulfur processing in the marine atmospheric boundary layer: A review and critical assessment of modeling uncertainties, Atmos. Env., 43, 2841-2854, doi:10.1016/j.atmosenv.2009.02.043, 2009.

Fayt, C. and M. Van Roozendael, Windoas 2.1, Software User Manual, BIRA-IASB, 2001.

Fioletov, V. E., C. A. McLinden, N. Krotkov, K. Yang, D. G. Loyola, P. Valks, N. Theys, M. Van Roozendael, C. R. Nowlan, K. Chance, X. Liu, C. Lee, and R. V. Martin: Application of OMI, SCIAMACHY, and GOME-2 satellite SO2 retrievals for detection of large emission sources, J. Geophys. Res. Atmos., 118, 11,399–11,418, doi:10.1002/jgrd.50826, 2013.

Fioletov, V., C. McLinden, N. Krotkov, C. Li, J. Joiner, N. Theys, S. Carn, and M. Moran: A global catalogue of large SO2 sources and emissions derived from the Ozone Monitoring Instrument, Atmos. Chem. Phys., 16, 11497-11519, doi:10.5194/acp-16-11497-2016, 2016.

Galle, B., Johansson, M., Rivera, C., Zhang, Y., Kihlman, M., Kern, C., Lehmann, T., Platt, U., Arellano, S., and Hidalgo, S.: Network for Observation of Volcanic and Atmospheric Change (NOVAC) – A global network for volcanic gas monitoring: Network layout and instrument description, J. Geophys. Res., 115, D05304, doi:10.1029/2009JD011823, 2010.

Heue K.-P., C.A.M. Brenninkmeijer, T. Wagner, K. Mies, B. Dix, U. Frieß, B. G. Martinsson, F. Šlemr, and P.F.J. van Velthoven, Observations of the 2008 Kasatochi volcanic SO_2 plume by CARIBIC aircraft DOAS and the GOME-2 satellite. Atmos. Chem. Phys., 10, 4699-4713, doi:10.5194/acp-10-4699-2010, 2010.

Heue, K., Brenninkmeijer, C. A. M., Baker, 935 A. K., Rauthe-Scho¨ch, A., Walter, D., Wagner, T., Hörmann, C., Sihler, H., Dix, B., Frieß, U., Platt, U., Martinsson, B. G., van Velthoven, P. F. J., Zahn, A., and Ebinghaus, R.: SO2 and BrO observation in the plume of the Eyjafjallajökull volcano 2010: CARIBIC and GOME-2 retrievals, Atmos. Chem. Phys., 11, 2973–2989, doi:10.5194/acp-11-2973-2011, 2011.

Hendrick, F., M. Van Roozendael, A. Kylling, A. Petritoli, A. Rozanov, S. Sanghavi, R. Schofield, C. von Friedeburg, T. Wagner, F. Wittrock, D. Fonteyn, and M. De Mazière : Intercomparison exercise between different radiative transfer models used for the interpretation of ground-based zenith-sky and multi-axis DOAS observations, Atmos. Chem. Phys., 6, 93-108, 2006.

Hermans, C., A.C. Vandaele, and S. Fally (2009), Fourier transform measurements of SO_2 absorption cross sections: I. Temperature dependence in the 24 000–29 000 cm^{-1} (345–420 nm) region, J. Quant Spectrosc. Radiat. Transfer, 110, 756–765, doi: 10.1016/j.jqsrt.2009.10.031.

Hörmann, C., H. Sihler, N. Bobrowski, S. Beirle, M. Penning de Vries, U. Platt, and T. Wagner, Systematic investigation of bromine monoxide in volcanic plumes from space by using the GOME-2 instrument, Atmos. Chem. Phys., 13, 4749-4781, doi:10.5194/acp-13-4749-2013, 2013

Huijnen, V., Williams, J., van Weele, M., van Noije, T., Krol, M., Dentener, F., Segers, A., Houweling, S., Peters, W., de Laat, J., Boersma, F., Bergamaschi, P., van Velthoven, P., Le Sager, P., Eskes, H., Alkemade, F., Scheele, R., Nédélec, P., and Pätz, H.-W., The global chemistry transport model tm5: description and evaluation of the tropospheric chemistry version 3.0., Geoscientific Model Development, 3(2):445-473, 2010.

Jin, J., Ma, J., Lin, W., Zhao, H., Shaiganfar, R., Beirle, S., and Wagner, T.: MAX-DOAS measurements and satellite validation of tropospheric NO2 and SO2 vertical column densities at a rural site of North China, Atmospheric Environment, 133, 12–25, 2016.

Kern, C., Deutschmann, T., Vogel, A., Wöhrbach, M., Wagner, T., and Platt, U.: Radiative transfer corrections for accurate spectroscopic measurements of volcanic gas emissions, Bull. Volcanol., 72,233-247, 2009.

Khokhar, M. F., C. Frankenberg, M. Van Roozendael, S. Beirle, S. Kühl, A. Richter, U. Platt and T. Wagner, Satellite Observations of Atmospheric SO_2 from Volcanic Eruptions during the Time Period of 1996 to 2002, J. Adv. Space Res., 36(5), 879-887, 10.1016/j.asr.2005.04.114, 2005.

Kleipool, Q. L., Dobber, M. R., de Haan, J. F. and Levelt, P. F.: Earth surface reflectance climatology from 3 years of OMI data, J. Geophys. Res., 113(D18), D18308, doi:10.1029/2008JD010290, 2008

Koelemeijer, R. B. A., Stammes, P., Hovenier, J. W. and de Haan, J. F.: A fast method for retrieval of cloud parameters using oxygen A band measurements from the Global Ozone Monitoring Experiment, J. Geophys. Res., 106(D4), 3475-3490, doi:10.1029/2000JD900657 2001.

Koelemeijer, R.B.A., de Haan, J.F. and Stammes, P. , A database of spectral surface reflectivity in the range 335-772 nm derived from 5.5 years of GOME observations. Journal of Geophysical Research 108: doi: 10.1029/2002JD002429. issn: 0148-0227, 2003

Krotkov, N. A., Carn, S. A., Krueger, A. J., Bhartia, P. K., Yang, K.: Band residual difference algorithm for retrieval of SO_2 from the Aura Ozone Monitoring Instrument (OMI), IEEE Trans. Geosci. Remote Sensing, AURA Special Issue, 44(5), 1259-1266, doi:10.1109/TGRS.2005.861932, 2006.

Krotkov, N. A., et al., Validation of SO_2 retrievals from the Ozone Monitoring Instrument over NE China, J. Geophys. Res., 113, D16S40, doi:10.1029/2007JD008818, 2008.

Krueger A.J.: Sighting of El Chichon sulfur dioxide clouds with the Nimbus 7 total ozone mapping spectrometer, Science, 220, 1377–1379, 1983.

Li, C., J. Joiner, N. A. Krotkov, and P. K. Bhartia (2013), A fast and sensitive new satellite SO2 retrieval algorithm based on principal component analysis: Application to the ozone monitoring instrument, Geophys. Res. Lett., 40, 6314–6318, doi:10.1002/2013GL058134.

Lee, C., R. V. Martin, A. van Donkelaar, G. O'Byrne, N. Krotkov, A. Richter, L. G. Huey, and J. S. Holloway (2009), Retrieval of vertical columns of sulfur dioxide from SCIAMACHY and OMI: Air mass factor algorithm development, validation, and error analysis, J. Geophys. Res., 114, D22303, doi:10.1029/2009JD012123.

Martin, R. V., Chance, K., Jacob, D. J., Kurosu, T. P., Spurr, R. J. D., Bucsela, E., Gleason, J.F., Palmer, P.I., Bey, I., Fiore, A.M., Li, Q., Yantosca, R.M., and Koelemeijer, R.B.A..: An improved retrieval of tropospheric nitrogen dioxide from GOME, J. Geophys. Res., 107(D20), 4437, doi:10.1029/2001JD001027, 2002.

Nowlan, C.R., Liu, X., Chance, K., Cai, Z., Kurosu, T.P., Lee, C., and Martin, R.V., Retrievals of sulfur dioxide from the Global Ozone Monitoring Experiment 2 (GOME-2) using an optimal estimation approach: Algorithm and initial validation, J. Geophys. Res., 116, D18301, doi:10.1029/2011JD015808, 2011.

Palmer, P. I., Jacob, D. J., Chance, K. V., Martin, R. V., D, R. J., Kurosu, T. P., Bey, I., Yantosca, R. and Fiore,A.: Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, Journal of Geophysical Research, 106(D13), 14539-14550, doi:10.1029/2000JD900772, 2001.
Platt, U., and Stutz, J.: Differential Optical Absorption Spectroscopy (DOAS), Principle and Applications, ISBN 3-340-21193-4, Springer Verlag, Heidelberg, 2008.

Prata, A. J., Satellite detection of hazardous volcanic clouds and the risk to global air trafc, Nat. Hazards, 51(2), 303–324, DOI: 10.1007/s11069-008-9273-z, 2009.

Puķīte, J., Kühl, S., Deutschmann, T., Platt, U., and Wagner, T.: Extending differential optical absorption spectroscopy for limb measurements in the UV, Atmos. Meas. Tech., 3, 631-653, 2010.

Richter, A., Wittrock, F., Sch¨onhardt, A., and Burrows, J.: Quantifying volcanic SO2 emissions using GOME2 measurements, Geophys. Res. Abstr., EGU2009-7679, EGU General Assembly 2009, Vienna, Austria, 2009.

Richter, A., Begoin, M., Hilboll, A. and Burrows, J. P.: An improved NO2 retrieval for the GOME-2 satellite instrument, Atmos. Meas. Tech., 4(6), 213-246, doi:10.5194/amt-4-1147-2011, 2011.

Rix, M., P. Valks, N. Hao, D. G. Loyola, H. Schlager, H. H. Huntrieser, J. Flemming, U. Koehler, U. Schumann, and A. Inness, Volcanic SO2, BrO and plume height estimations using GOME-2 satellite measurements during the eruption of Eyjafjallajökull in May 2010, J. Geophys. Res., 117, D00U19, doi:10.1029/2011JD016718, 2012.

Robock, A.: Volcanic eruptions and climate, Rev. Geophys., 38, 191–219, 2000.

Rozanov, A., Rozanov, V., and Burrows, J. P.: A numerical radiative transfer model for a spherical planetary atmosphere: Combined differential integral approach involving the Piccard iterative approximation, J. Quant. Spectrosc. Radiat. Transfer, 69, 491–512, 2001.

Schumann, U., Weinzierl, B., Reitebuch, O., Schlager, H., Minikin, A., Forster, C., Baumann, R., Sailer, T., Graf, K., Mannstein, H., Voigt, C., Rahm, S., Simmet, R., Scheibe, M., Lichtenstern, M., Stock, P., R¨uba, H., Sch¨auble, D., Tafferner, A., Rautenhaus, M., Gerz, T., Ziereis, H., Krautstrunk, M., Mallaun, C., Gayet, J.-F., Lieke, K., Kandler, K., Ebert, M., Weinbruch, S., Stohl, A., Gasteiger, J., Groß, S., Freudenthaler, V., Wiegner, M., Ansmann, A., Tesche, M., Olafsson, H., and Sturm, K.:

Airborne observations of the Eyjafjalla volcano ash cloud over Europe during air space closure in April and May 2010, Atmos. Chem. Phys., 11, 2245-2279, doi:10.5194/acp-11-2245-2011, 2011.

Spinei, E., S.A. Carn, N.A. Krotkov, G.H. Mount, K. Yang, and A.J. Krueger, Validation of ozone monitoring instrument SO2 measurements in the Okmok volcanic plume over Pullman, WA in July 2008, J. Geophys. Res., Okmok-Kasatochi Special Issue, 115, D00L08, doi:10.1029/2009JD013492, 2010

Spurr, R., LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems. Light Scattering Reviews, Volume 3, ed. A. Kokhanovsky, Springer, 2008.

Spurr, R., J.F. de Haan, R. van Oss and A. Vasilkov, Discrete Ordinate Radiative Transfer in a Stratified Medium with First Order Rotational Raman Scattering, J. Quant. Spectros. Rad. Transf., 2008, 109, 3, 404-425, doi:10.1016/j.jqsrt.2007.08.011.

Tanskanen, A. Lambertian Surface Albedo Climatology at 360 nm from TOMS Data Using Moving Time-Window Technique. In: Proceedings of the XX Quadrennial Ozone Symposium, 1-8 June 2004, Kos, Greece.

Theys, N., I. De Smedt, J. van Gent, T. Danckaert, T. Wang, F. Hendrick, T. Stavrakou, S. Bauduin, L. Clarisse, C. Li, N. A. Krotkov, H. Yu, M. Van Roozendael, Sulfur dioxide vertical column DOAS retrievals from the Ozone Monitoring Instrument: Global observations and comparison to ground-based and satellite data, J. Geophys. Res. Atmos., 120, doi:10.1002/2014JD022657, 2015.

Theys, N., I. De Smedt, H. Yu, T. Danckaert, J. van Gent, C. Hörmann, T. Wagner, P. Hedelt, H. Bauer, F. Romahn, M. Pedergnana, D. Loyola, M. Van Roozendael : Sulfur dioxide operational retrievals from TROPOMI onboard Sentinel-5 Precursor: Algorithm Theoretical Basis, Atmos. Meas. Tech., 10, 119-153, doi:10.5194/amt-10-119-2017, 2017.

Vandaele A.C., C. Hermans, P.C. Simon, M. Carleer, R. Colin, S. Fally, M.F. Mérienne, A. Jenouvrier, and B. Coquart, Measurements of the NO2 absorption cross-section from 42000 cm-1 to 10000 cm-1 (238-1000 nm) at 220 K and 294 K, J.Q.S.R.T., 59, 171-184, 1998

Vandaele, A. C., C. Hermans, and S. Fally : Fourier transform measurements of SO2 absorption cross sections: II. Temperature dependence in the 29000–44000 cm−1 (227–345 nm) region, J. Quant. Spectrosc. Radiat. Transfer, 110, 2115–2126, doi:10.1016/j.jqsrt.2009.05.006, 2009.

Van Geffen, J., M. van Roozendaal, M. Rix, and P. Valks, Initial Validation of GOME-2 GDP 4.2 SO2 Total Columns—ORR B, TN-IASB-GOME2-O3MSAF-SO2-01, Sep. 2008

Veefkind, J.P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H.J., de Haan, J.F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeven, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P.F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sensing of Environment, doi:10.1016/j.rse.2011.09.027, 2012.

Vountas, M., Rozanov, V. V. and Burrows, J. P.: Ring effect: impact of rotational Raman scattering on radiative transfer in earth's atmosphere, J. of Quant. Spec. and Rad. Trans., 60(6), 943-961, 36 1998.

Wang, Y., Beirle, S., Lampel, J., Koukouli, M., De Smedt, I., Theys, N., Li, A., Wu, D., Xie, P., Liu, C., Van Roozendael, M., Stavrakou, T., Müller, J.-F., and Wagner, T.: Validation of OMI, GOME-2A and GOME-2B tropospheric NO_2, SO_2 and HCHO products using MAX-DOAS observations from 2011 to 2014 in Wuxi, China: investigation of the effects of priori profiles and aerosols on the satellite products, Atmos. Chem. Phys., 17, 5007-5033, doi:10.5194/acp-17-5007-2017, 2017.

Wagner, T., J. P. Burrows, T. Deutschmann, B. Dix, C. von Friedeburg, U. Frieß, F. Hendrick, K.-P. Heue, H. Irie, H. Iwabuchi, Y. Kanaya, J. Keller, C. A. McLinden, H. Oetjen, E. Palazzi, A. Petritoli, U. Platt, O. Postylyakov, J. Pukite, A. Richter, M. van Roozendael, A. Rozanov, V. Rozanov, R. Sinreich, S. Sanghavi, and F. Wittrock: Comparison of box-air-mass-factors and radiances for Multiple-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) geometries calculated from different UV/visible radiative transfer models, Atmos. Chem. Phys., 7, 1809-1833, 2007.

Wagner, T., Beirle, S., & Deutschmann, T. (2009). Three-dimensional simulation of the Ring effect in observations of scattered sun light using Monte Carlo radiative transfer models. Atm. Meas. Tech., 2, 113-124.

Yang, K., N. Krotkov, A. Krueger, S. Carn, P. K. Bhartia, and P. Levelt, Retrieval of Large Volcanic SO2 columns from the Aura Ozone Monitoring Instrument (OMI): Comparisons and Limitations, J. Geophys. Res., 112, D24S43, doi:10.1029/2007JD008825, 2007.

Yang, K., Liu, X., Bhartia, P., Krotkov, N., Carn, S., Hughes, E., Krueger, A., Spurr, R., Trahan, S.: Direct retrieval of sulfur dioxide amount and altitude from spaceborne hyperspectral UV measurements: Theory and application, J. Geophys. Res., 115, D00L09, doi:10.1029/2010JD013982, 2010.

Yang, K., R. R. Dickerson, S. A. Carn, C. Ge, and J. Wang (2013), First observations of SO2 from the satellite Suomi NPP OMPS: Widespread air pollution events over China, Geophys. Res. Lett., 40, 4957–4962, doi:10.1002/grl.50952.

Zhou, Y., Brunner, D., Boersma, K. F., Dirksen, R., and Wang, P.: An improved tropospheric NO_2 retrieval for OMI observations in the vicinity of mountainous terrain, Atmos. Meas. Tech., 2, 401-416, doi:10.5194/amt-2-401-2009, 2009.

A Description of prototype software

The prototype software for the retrieval of total vertical SO_2 columns has been based on a long history of development and improvements of DOAS-type retrieval techniques. The processing flow of the current prototype software is outlined in Figure 5-1. The key elements of the retrieval process are threefold:

1. The derivation of a trace gas slant column, along the effective light path in the atmosphere. For this the QDOAS software environment is used, extended for the processing of TROPOMI data.
2. An offset correction, mitigating SO_2 signals originating from systematic biases in the algorithm.
3. Conversion of the slant column to a vertical column. For this we use look-up tables of pressure dependent box air mass factor s, modified by an appropriate vertical SO_2 profile shape. The box air mass factors have been calculated with the LIDORT v3.3 atmospheric radiative transfer model (RTM).

The different processing steps in the derivation of a vertical gas column from satellite L1b measurements have been outlined in detail in chapter 5. Here, we limit ourselves to describing the software modules, used in the processing flow, that have a known track record.

A.1 QDOAS

The SO_2 prototype uses the QDOAS software suite, a multi-purpose DOAS analysis environment developed at BIRA-IASB (Danckaert et al., 2012). The experience of BIRA-IASB in the development and improvement of algorithms for the retrieval of trace gas concentrations goes back to the early 1990s. WinDOAS, the first program developed at BIRA-IASB in 1997, knew a success story due to a friendly user interface complemented with some powerful DOAS tools (Fayt and Van Roozendael, 2001). This program, extensively validated through different campaigns, has been used worldwide and for many different DOAS applications (mainly for groundbased and satellite applications). QDOAS is a cross-platform implementation of WinDOAS *i.e.* that the software is portable to Windows and Unix-based operating systems. Besides its graphical user interface and visualization tools, a powerful command line tool (doas_cl) for batch processing is also available. The QDOAS fitting algorithm uses a combination of Singular Value Decomposition (SVD) and the Levenberg-Marquard methods to solve the linear and non-linear parts of the DOAS equation.

The main QDOAS features include:

Analysis

- DOAS/intensity fitting modes;
- shift/stretch fully configurable for any spectral item (cross-section or spectrum);
- possibility to filter spectra and cross-sections before analysis (supported filters include Kaiser, Gaussian, boxcar, Savitsky Golay…);
- possibility to define gaps within fitting intervals (e.g. to eliminate bad pixels);
- possibility to fit an instrumental offset;
- possibility to define several configurations of spectral windows under a project ;
- non-linear parameters (offset, shift and stretch) can be fitted using wavelength polynomials up to the second order.

Calibration and slit function characterization

- wavelength calibration and instrumental slit function characterization using a nonlinear least-squares (NLLS) fitting approach where measured intensities are fitted to a high resolution solar spectrum degraded to the resolution of the instrument. The fitting method (DOAS or intensity fitting) can be different from the method used in the analysis;
- possibility to correct for atmospheric absorption and Ring effect;
- supports different analytical line shapes.

Cross-sections handling

- possibility to calculate differential absorption cross-sections (by orthogonalisation or high-pass filtering);
- possibility to correct multiple cross-sections using a wavelength dependent AMF;
- possibility to fix the column density of any selected species;

More information on the QDOAS software, as well as the software user manual can be found online [URL01].

A.2 LIDORT

The linearized discrete ordinate radiative transfer code LIDORT is the radiative transfer model applied in the calculation of the box air mass factor look-up tables, used in the algorithm for the determination of the air mass factor. LIDORT is being developed and maintained by the company RT Solutions, Inc in Cambridge, MA in the United States. Many aspects of LIDORT environments have been realized through intensive cooperation between RT Solutions, Inc. and BIRA-IASB. As an example, LIDORT has been the core of the radiative transfer calculations in the direct fitting algorithm GODFIT, developed at BIRA-IASB for the derivation of improved total ozone columns from measurements of the GOME instrument on-board the ERS-2 platform. During this project, LIDORT was extended with the capability of calculating analytical box air mass factor for total gas columns, which plays a key role in GODFIT's O_3 column derivation. Later versions of GODFIT benefited from many improvements in LIDORT, both regarding performance and the treatment of physics. Currently GODFIT uses LIDORT v3.3 (Lerot et al., 2010; Spurr, 2008). This model is but one of several members of the LIDORT family of RTMs. Another family member, applied at BIRA-IASB for the generation and study of Ring spectra, is LIDORT-RRS (Spurr et al., 2008). This model incorporates rotational Raman scattering by atmospheric N_2 and O_2 molecules. More information on the LIDORT suite of RTM codes can be found on the RT Solutions, Inc. website [URL03].