Fiducial Reference Measurements for Ground-Based DOAS Air-Quality Observations

FRM4DOAS

ESA Contract No. 4000118181/16/I-EF

Deliverable D6:

MAXDOAS Algorithm ATBD

Date: 30/01/2018

Version: 0.2

Contributing authors:

Francois Hendrick (BIRA-IASB)

Caroline Fayt (BIRA-IASB)

Martina Friedrich (BIRA-IASB)

Udo Frieß (UHeid)

Andreas Richter (UBremen)

Steffen Beirle (MPIC)

Thomas Wagner (MPIC)

Yang Wang (MPIC)

+…..

Document change record

issue	date	item	comments
0.1	2016-11-08	–	Initial version
0.2	2018-01-30	–	Preliminary complete version

Contents

1 Introduction

1.1 Purpose of the ATBD

This Algorithm Theoretical Basis Document (ATBD) describes the algorithms to be implemented in the FRM$_4$DOAS centralised processing system for the retrieval of tropospheric NO_2 and HCHO vertical profiles, total O_3 columns, and stratospheric NO_2 vertical profiles. These algorithms have been selected through a dedicated Round-Robin exercise [RD-1].

….

1.2 Document overview

Following reference sections on applicable documentation (Section 2) and acronyms and abbreviations (Section 3), ground-based DOAS technique main principles and retrieval methods related to the algorithms implemented in the FRM$_4$DOAS processing system are described in Section 4. The algorithms descriptions are found in Section 5, which also contains sub-sections on error analyses, vertical resolution, spatial representativeness, and determination of cloud conditions. Conclusions are drawn in Section 6 and References are given in Section 7.

1.3 Acknowledgements

To be added if any.

2 Applicable and reference documents

2.1 Applicable documents

[AD-1] ESA/ESRIN Statement of Work, ref. Fiducial Reference Measurements for Ground-Based DOAS Air-Quality Observations, ENVI-SPPA-EOPG-SW-14-0003, July 2015.

…..

2.2 Standard documents

……

2.3 Reference documents

[RD-1] FRM$_4$DOAS Deliverable D5 ‘MAXDOAS Algorithm Round-Robin Definition and Results Document’, January 2018.

[RD-2] FRM$_4$DOAS Deliverable D13 ‘Intercomparison Campaign Planning Document’, October 2016.

[RD-3] QA4ECV Deliverable D3.9 ‘Quality indicators on uncertainties and representativity of atmospheric reference data’, November 2016 – See http://www.qa4ecv.eu/sites/default/files/D3.9.pdf

……

2.4 Digital references

[DR-1] NDACC: Network for the Detection of Atmospheric Composition Change: http://www.ndacc.org.

[DR-2] CINDI-2: 2nd Cabauw Intercomparison of Nitrogen Dioxide measuring Instruments: http://www.tropomi.eu/data-products/cindi-2.

[DR-3] NDACC standard approach for total O_3 VCD retrieval: http://ndacc-uvvis-wg.aeronomie.be/tools.php.

……

3 Acronyms and abbreviations

Acronyms and abbreviated terms that are specific for this document can be found below.

AFGL	Air Force Geophysics Laboratory
AK	Averaging Kernel
AMF	Air Mass Factor
ASCII	American Standard Code for Information Interchange
ATBD	Algorithm Theoretical Basis Document
CINDI-1	1st Cabauw Intercomparaison of Nitrogen Dioxide measuring Instruments
CINDI-2	2nd Cabauw Intercomparaison of Nitrogen Dioxide measuring Instruments
CTM	Chemistry Transport Model
DOAS	Differential Optical Absorption Spectroscopy
DOASIS	DOAS Intelligent System
DOFS	Number of Degrees of Freedom for Signal
DSCD	Differential Slant Column Density
ENVISAT	Environmental Satellite
FT	Fourier Transform
FRM_4DOAS	Fiducial Reference Measurements for Ground-Based DOAS Air-Quality Observations
FWHM	Full Width at Half Maximum
GAW	Global Atmosphere Watch
GOME-2	The Global Ozone Monitoring Experiment–2
GUI	Graphical User Interface
HDF-EOS5	Hierarchical Data Format - Earth Observing System 5
HWHM	Half Width at Half Maximum
IO3C	International Ozone Commission

LM	Levenberg-Marquardt
MAPA	MAinz Profile Algorithm
MAXDOAS	Multi-Axis Differential Optical Absorption Spectroscopy
MC	Monte Carlo
MFC	
NDACC	Network for the Detection of Atmospheric Composition Change
netCDF	Network Common Data Form
OEM	Optimal Estimation Method
OMI	Ozone Monitoring Instrument
PSCs	Polar Stratospheric Clouds
RCD	Residual column density in the reference spectrum
RMS	Root mean square
RTM	Radiative transfer model
SAOZ	Système d'Analyse par Observation Zénithale
SCD	Absolute Slant Column Density
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CartograpHY
SZA	Solar zenith angle
TV8	TOMS Version 8 ozone and temperature profile climatology
UNEP	United Nations Environment Programme
VCD	Vertical Column Density
WF	Weighting Function
WG	Working Group
WMO	World Meteorological Organisation

4 The ground-based DOAS technique

4.1 Introduction

For about three decades, stratospheric ozone (O_3) and related trace gases such as nitrogen dioxide (NO_2), bromine monoxide (BrO), and chlorine dioxide (OClO) have been monitored using zenith scattered sunlight DOAS observations at twilight [e.g., Pommereau and Goutail, 1988; Solomon et al., 1989; McKenzie et al., 1991; Fiedler et al., 1993; Kreher et al., 1997; Richter et al., 1999; Van Roozendael et al., 1998; Struthers et al., 2004; Hendrick et al., 2008]. Most of these measurements, which exploit the increase of the slant light paths in the stratosphere in low sun conditions, have been carried out at about 30 stations belonging to the Network for the Detection of Atmospheric Composition Change (NDACC). NDACC is an international research network of more than 70 globally-distributed remote sounding stations aiming at observing and understanding the composition and structure of the stratosphere and troposphere, and contributing to scientific programmes from WMO/GAW, UNEP, and IO3C [DR-1; De Mazière et al., 2017]. Since early 2000's, a new technique called MAXDOAS has been developed for detecting absorbers present in the lowest few kilometres of the troposphere like NO_2, formaldehyde (HCHO), sulphur dioxide (SO_2), nitrous acid (HONO), glyoxal (CHOCHO), BrO, as well as aerosols. MAXDOAS spectrometers measure the absorption of sunlight scattered by molecules and aerosols in the UV-visible range not only at zenith but also at different elevation angles towards the horizon (the so-called off-axis geometry; see Figure 1), increasing therefore the sensitivity to absorbers present close to the ground. During the recent years, NDACC stations equipped with MAXDOAS instruments are growing in number and the operational generation of quality-assessed tropospheric data products is being developed for progressive implementation in NDACC. This task is performed under the responsibility of the NDACC UV-vis WG, notably through the organisation of intercomparison campaigns of MAXDOAS instruments, like CINDI-1 and -2 [Roscoe et al., 2010; DR2].

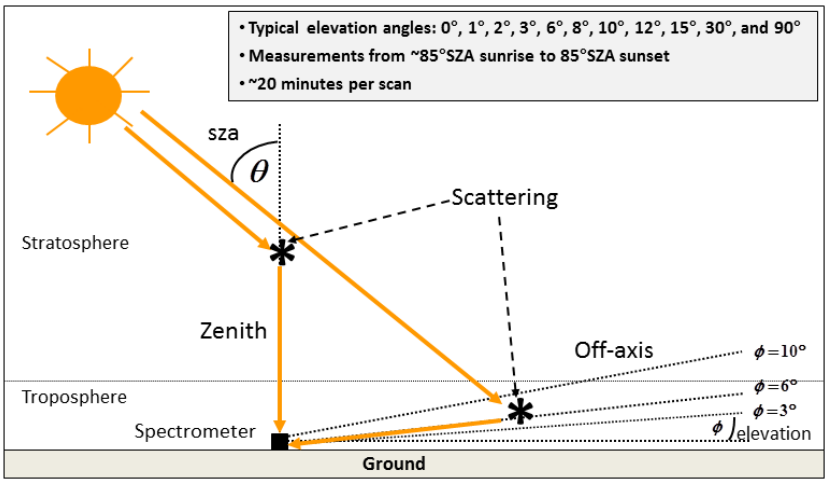

Figure 1: Sketch illustrating the MAXDOAS viewing geometries (off-axis and zenith). Typical elevation angle values used during a scan are given in the top right grey frame.

4.2 Principle of the DOAS method

The determination of the abundance of an atmospheric trace gas using the DOAS method involves a two-step procedure [Platt and Stutz, 2008; see also Hönninger et al., 2004]. The first step, called DOAS spectral fitting, is based on the Beer-Lambert extinction law and consists in (1) separating trace gas fine-scale absorption features from broad-band absorption due to Rayleigh and Mie scattering and instrumental effects, and (2) analysing the remaining absorber narrow-band structures by least-squares spectral fitting on laboratory cross-sections. It provides as end product the so-called differential slant column densities (DSCDs) which are defined as the difference between the trace-gas concentration integrated along the effective light path and the amount of absorber in a measured reference spectrum. In the second step, resulting DSCDs are converted into more physically interpretable variables, namely vertical column densities (VCDs) or vertical profiles, using appropriate radiative transfer calculations and profile inversion schemes.

Both steps are described in Sections 4.2.1 and 4.2.2 below.

4.2.1 DOAS spectral fit

In DOAS, the atmospheric model is the Beer-Lambert extinction law for trace gas absorbers, that describes the relationship between the incident light intensity at the location of the light source (or, in case of scattered light, at the top of the atmosphere) and the transmitted intensity at the location of the instrument detector:

$$I(\lambda) = I_0(\lambda)\exp(-\int_0^L \left[\sum_i \sigma_{a,i}(\lambda) n_i + \sigma_R(\lambda) n_R + k_M(\lambda)\right] ds) \qquad (1)$$

where $I_0(\lambda)$ and $I(\lambda)$ are the incident and transmitted intensity, respectively, at wavelength λ, L is the length of the effective light path, $\sigma_{a,i}(\lambda)$ and n_i are the absorption cross section at wavelength λ and number density of the absorber i, respectively, $\sigma_R(\lambda)$ is the Rayleigh scattering cross section at wavelength λ, n_R is the air density, $k_M(\lambda)$ is the Mie extinction coefficient at wavelength λ, and ds is the thickness of the atmospheric layer traversed by the light.

The optical depth $\tau(\lambda)$ can be introduced

$$\tau(\lambda) = -ln\left[\frac{I(\lambda)}{I_0(\lambda)}\right] = \sum_i (\int_0^L \sigma_{a,i}(\lambda) n_i ds) + \int_0^L \sigma_R(\lambda) n_R ds + \int_0^L k_M(\lambda) ds \qquad (2)$$

and Eq. (1) becomes

$$\tau(\lambda) = \sum_i (\int_0^L \sigma_{a,i}(\lambda) n_i ds) + \tau_R(\lambda) + \tau_M(\lambda) \qquad (3)$$

where $\tau_R(\lambda)$ and $\tau_M(\lambda)$ are the Rayleigh and Mie scattering optical depths, respectively.

Assuming that the height dependency of the absorption cross sections $\sigma_{a,i}(\lambda)$ can be neglected, i.e. absorption cross-sections are weakly dependent on the temperature and pressure along the effective light path, the summation and the integral in Eq. (3) can be exchanged and $\tau(\lambda)$ can be expressed as:

$$\tau(\lambda) = \sum_i \sigma_{a,i}(\lambda)\, SCD_i + \tau_{MR}(\lambda) + \tau_M(\lambda) \text{ with } SCD_i = \int_0^L n_i ds \qquad (4)$$

SCD_i is the so-called slant column density of the absorber *i* and is the direct product of the DOAS analysis. It represents the total amount of the absorber *i* per unit area integrated along the effective light path.

At this point, the so-called DOAS concept is introduced into the derivation. It based on the fact that Rayleigh and Mie scattering vary weakly with wavelength, while the molecular absorption cross section $\sigma_{a,i}(\lambda)$ of a certain species can be separated into a low and a high frequency component or a broad-band and a differential absorption cross section $\sigma^{S}_{a,i}(\lambda)$ and $\sigma'_{a,i}(\lambda)$, respectively:

$$\sigma_{a,i}(\lambda) = \sigma'_{a,i}(\lambda) + \sigma^{S}_{a,i}(\lambda) \quad (5)$$

By replacing $\sigma_{a,i}(\lambda)$ by (5) in Eq. (4) and by substituting components with low frequency variations ($\sigma^{S}_{a,i}(\lambda)$, Rayleigh and Mie scattering), which follow simple power laws, by a polynomial $P = \sum_P a_P \lambda^P$, the DOAS equation is then obtained:

$$\tau(\lambda) = -ln\left[\frac{I(\lambda)}{I_0(\lambda)}\right] = \sum_i \sigma'_{a,i}(\lambda)\, SCD_i + \sum_P a_P \lambda^P \quad (6)$$

Knowing the reference spectrum $I_0(\lambda)$, the slant column densities SCD_i and coefficient a_p can be retrieved by adjusting the differential absorption cross sections $\sigma'_{a,i}(\lambda)$ (determined from laboratory measurements) to the measured optical depth by means of a linear least squares fit.

For ground-based observations, it is a common method to use a measured spectrum $I'_0(\lambda)$ with the smallest absorptions as reference spectrum, since it is not possible, in contrast to satellite measurements, to record an extraterrestrial solar spectrum with the same instrument. For all viewing directions of the MAXDOAS measurements, the zenith spectrum of the scan is usually chosen as reference. In the case of the zenith-sky twilight observations, a daily noon zenith spectrum is used.

Thus, Eq. (6) should be rewrote as:

$$\tau(\lambda) = -ln\left[\frac{I(\lambda)}{I'_0(\lambda)}\right] = \sum_i \sigma'_{a,i}(\lambda)\,(SCD_i - SCD_i^{ref}) + \sum_P a_P \lambda^P \quad (7)$$

where $(SCD_i - SCD_i^{ref}) = DSCD_i$ are the so-called differential slant column densities of the absorber *i*. The absorber amount in the reference spectrum is then determined by an extrapolation method (e.g. a Langley-plot analysis) and added to the retrieved $DSCD_i$ to obtain the absolute slant column densities SCD_i.

4.2.2 VCD and vertical profile inversion methods

The different methods used in FRM$_4$DOAS for converting DSCDs into VCDs and/or vertical profiles are described in this Section. Since these methods are product-specific, a separate description is given based on FRM$_4$DOAS end-product (NO_2 and HCHO lower tropospheric vertical profiles, total O_3 columns, and stratospheric NO_2 vertical profiles).

4.2.2.1 NO_2 and HCHO lower tropospheric vertical profiles

NO_2 and HCHO lower tropospheric vertical profiles are retrieved for each MAXDOAS scan by applying a profiling algorithm to the corresponding sets of DSCDs measured at the different elevation angles. The zenith spectrum of the scan is usually chosen as reference, in this way also removing the

contribution of the stratosphere in off-axis DSCDs, which can be significant for a species like NO_2. The MAXDOAS profiling technique is based on the fact that the mean scattering height rises into the atmosphere with the increase of the elevation angle and probes the layers where the tropospheric absorber is present. So, each measured DSCD of a MAXDOAS scan is representative of the absorber concentration in a given altitude range and therefore the observed DSCD variation as a function of the elevation angle depends on the vertical distribution of the absorber.

Two different MAXDOAS profiling tools have been selected through a Round-Robin exercise for implementation in the FRM_4DOAS processing system: the BIRA MMF OEM-based algorithm that gives vertical distributions of NO_2 and HCHO on a defined altitude grid (e.g., 0 to 4 km by 200m step) and the MPIC MAPA MC algorithm relying on a parameterized retrieval scheme that gives a first order description of typical vertical distributions (see Figure 2).

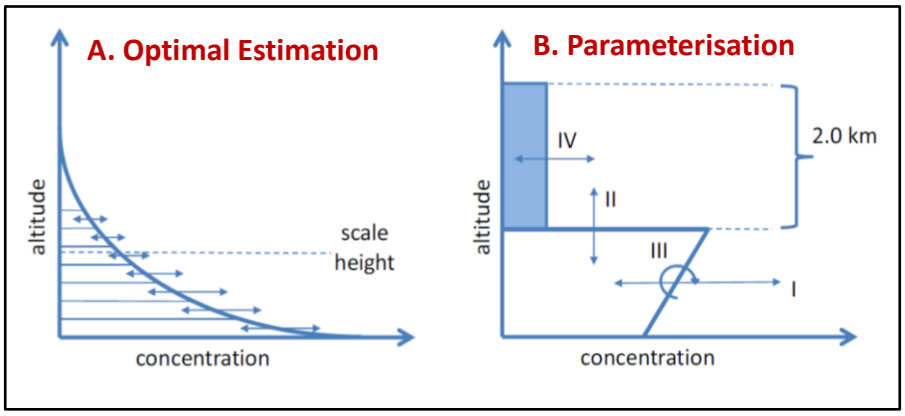

Figure 2: Schematic of profile parameterizations used in the OEM (A/left) and parametrized (B/right) approaches. In the case of method B, only a limited number of parameters are retrieved, e. g. (I) tropospheric VCD or AOD, (II) mixing layer top height, (III) shape parameter, (IV) free tropospheric VCD. Adapted from Vlemmix et al. [2015].

Both methods use a two-step procedure. First, the aerosol extinction vertical profiles are retrieved for each MAXDOAS scan from the corresponding measured oxygen dimer O_4 DSCDs. The principle of this retrieval is the following: since the O_4 vertical profile is well-known and nearly constant (it varies with the square of the O_2 monomer concentration), O_4 DSCD measurements can provide information on the vertical distribution of aerosols [Wagner et al. 2004; Friess et al., 2006]. This first step is required since the light path length through the atmosphere (and thus the measured NO_2 and HCHO DSCD) strongly depends on the aerosols and therefore a good estimate of the vertical distribution of the aerosols is needed to calculate appropriate trace gas WFs or SCDs. In the second step, the OEM-based or parameterized method is applied to the measured NO_2 and HCHO DSCDs in order to retrieve vertical profiles.

In the OEM formalism [see Rodgers 2000] used in MMF, a profile x is retrieved given an a priori profile x_a, the measurements y (DSCDs as a function of scan elevations for tropospheric NO_2 or HCHO profiles), their respective uncertainty covariance matrices (S_a and S_ε, respectively), and the matrix K of the WFs that indicate the sensitivity of the differential slant column abundances at each SZA to a change in the vertical profile:

$$x = x_a + S_a K^T (K S_a K^T + S_\varepsilon)^{-1} (y - K x_a) \text{ with } K = \frac{\partial y}{\partial x} \text{ and } K^T \text{ is the transpose of } K \quad (8)$$

The weighting functions have been determined by consecutively perturbing each layer of the a priori profile and recalculating the set of measurements using a RTM as forward model. The OEM for linear case is considered for optically thin absorbers like NO_2 and HCHO because measured DSCDs depend linearly on the concentrations in each profile layer. Therefore the WF matrix *K* is independent of the state [Heskes and Boersma, 2003] and a single inversion step is sufficient. In case of strong absorbers like O_4, a non-linear iterative approach based on a Gauss-Newton or Levenberg-Marquart minimisation scheme must be used [Rodgers, 2000; see also Friess et al. (2004)]:

$$x_{i+1} = x_i + (S_a^{-1} + K_i^T S_\varepsilon^{-1} K_i)^{-1}[K_i^T S_\varepsilon^{-1}\big(y - F(x_i)\big) - S_a^{-1}(x_i - x_a)] \quad (9)$$

where *F* is the forward model operator, K_i is the sequentially updated WF at step *i*, and the superscripts T and -1 denote the transposed and inversed matrix, respectively.

The a priori profile x_a and the covariance matrices of uncertainties in the a priori profile and in the measurements (S_a and S_ε, respectively) are key parameters for the retrieval. Because the inversion problem is ill-posed (there are more elements in the state vector *x* than independently measured elements in the vector *y* and therefore no unique solution to Eq. 8 and 9), a priori constraints are necessary to reject unrealistic solutions that might be consistent with the measurements. A priori profiles can be built by imposing a profile shape (e.g. decreasing exponential) or taken from the output of a 3D-CTM model.

Since the residuals from the DOAS fitting are in most cases dominated by the random noise of the detector, the measurement covariance matrix S_ε is chosen diagonal with values corresponding to the statistical errors on the trace gas DOAS fitting. The S_ε matrix being fixed, the a priori covariance matrix S_a can act like a tuning parameter [see e.g. Schofield et al., 2004]. The variance value to be placed on the diagonal of the S_a matrix will be empirically determined for each {trace gas, FRM_4DOAS station} pair based on dedicated sensitivity tests. The selected variance value should maximize the information content (DOFS; see Section 5.6 below) without leading to undesired oscillations in the retrieved profiles. S_a also contains extra-diagonal terms in order to account for correlations between trace gas concentrations at different altitude levels. These terms are usually added as Gaussian functions as follows [Barret et al., 2002]:

$$S_{a\,ij} = \sqrt{S_{a\,ii} S_{a\,jj} e^{-\ln(2)(\frac{z_i - z_j}{\gamma})^2}} \quad (10)$$

where z_i and z_j are the altitudes of i^{th} and j^{th} levels, respectively, and γ is the HWHM (half of the correlation length). In the case of MAXDOAS retrievals, the correlation length usually varies between 50 and 200m.

In contrast to the OEM, the parameterized approach makes use of a profile shape parameterization with just a few free parameters (e.g. layer height, column amount, shape parameter; see Figure 2) which are retrieved by combining off-line forward simulations of AMFs and SCDs and a least-square minimisation strategy [see e.g. Li et al., 2010; Wagner et al., 2011; Vlemmix et al., 2011 and 2015]. The main reason to use a low number of free parameters is that the information content of MAXDOAS observations with respect to the vertical distribution of aerosols and trace gases is quite

limited, and therefore, a sufficiently wide range of possible profile shapes can be retrieved with a limited but appropriate choice of free parameters [Vlemmix et al., 2015].

Regarding the minimization strategy, standard least-square methods like Levenberg-Marquart can be used [see e.g. Li et al., 2010] or more faster approaches which do not required the calculation of derivatives, like the simplex method [see Vlemmix et al., 2011]. In the case of the MAPA MC algorithm, forward model simulations are compared to the measured DSCDs by use of a Monte-Carlo scheme where vertical aerosol/trace gas profiles are derived from the best matching parameters [see Section 5.2.1].

4.2.3 Total O_3 column retrieval

Total O_3 VCDs are retrieved from twilight zenith-sky UV-vis observations using the following expression [see e.g. Hendrick et al., 2011]:

$$VCD(\theta) = \frac{DSCD(\theta) + RCD}{AMF(\theta)} \qquad (11)$$

where *VCD(θ)* is the O_3 VCD at the effective SZA *θ*, *DSCD(θ)* is the O_3 DSCD at SZA *θ*, *RCD* is the residual O_3 amount in the reference measurement (generally a fixed spectrum usually recorded at high sun around local noon), and *AMF(θ)* is the AMF at SZA *θ*.

Given expression (11), the retrieval of vertical columns consists of four steps: (1) slant column spectral fitting, (2) determination of residual amount in the reference spectrum, (3) conversion of absolute slant columns into vertical columns using appropriate AMFs, and (4) averaging of the vertical columns over a limited SZA range around 90° SZA. The DOAS spectral fit giving *DSCD(θ)* is described in Section 4.2.1. In step (2), *RCD* is derived using the so-called Langley plot method, which consists in rearranging Eq. (11) and plotting *DSCD(θ)* as a function of *AMF(θ)*, the intercept at AMF = 0 giving *RCD* [Roscoe et al., 1994; Vaughan et al., 1997]. *VCD(θ)* is then derived in step (3) by dividing the absolute O_3 SCD at SZA θ (*DSCD(θ)+RCD*) by an appropriate *AMF(θ)*. AMFs are computed at a single wavelength chosen around 500 nm with a RTM initialized with O_3, pressure, temperature, and aerosol extinction profiles representative, as much as possible, of the atmosphere at the location of the station.

Regarding step (4), sunrise and sunset O_3 column data are derived by averaging VCDs estimated with Eq. (11) over a limited SZA range around 90°SZA.

4.2.4 Stratospheric NO_2 profile retrieval

The inversion of stratospheric NO_2 vertical profiles from twilight zenith-sky NO_2 DSCDs observation is based on the dependence of the mean scattering height with SZA: during twilight, as the SZA increases, the mean scattering height rises into the absorber layer (20 to 40 km for NO_2), so each slant column measurement is mainly representative of the absorber concentration in a given altitude range. Thus, the observed slant column variation during twilight depends on the vertical distribution of the trace gas and the latter can be retrieved by applying the linear OEM approach described by Eq.

(8) in Section 4.2.2.1 [see Preston et al., 1997; Hendrick et al., 2004] and where the measurement vector *y* consists in a set of measured twilight (sunrise or sunset) NO_2 DSCDs usually measured in the 75-94°SZA range.

In the case of NO_2, the simple geometric dependence on which the profiling approach is based, is complicated by the rapid variation of NO_2 concentrations with local SZA at twilight due to photochemistry (see Figure 3). In order to calculate appropriate WFs, it is therefore important to use a RTM that accounts for the 2D (altitude, SZA) photochemical variation of NO_2, at least for the incident beam (first order approximation). The 2D NO_2 concentrations fields which need to be implemented in the RTM are generally taken from stacked box photochemical model simulations initialized with 3D-CTM meteorological and chemical fields extracted at the location of the station [Hendrick et al., 2004].

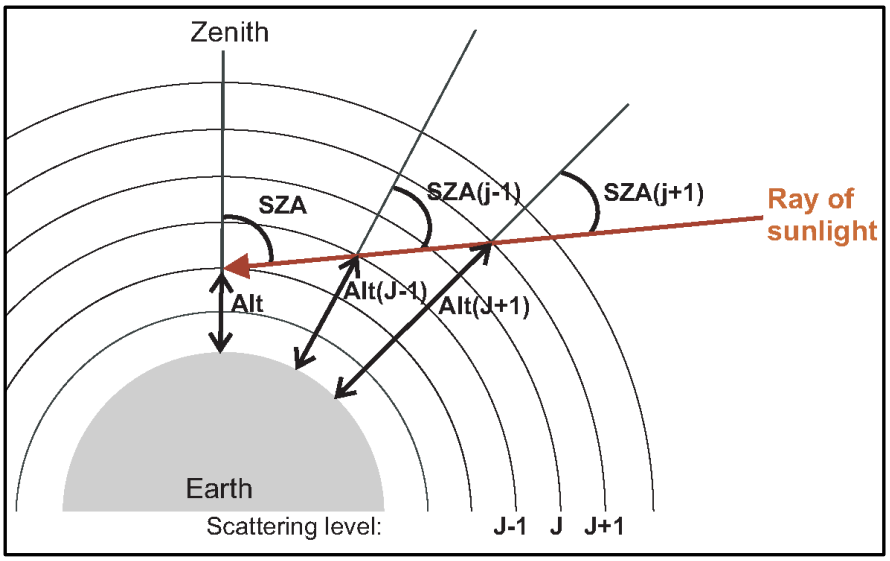

Figure 3: Variation of the local SZA along the path of a given ray. SZA(j)=local SZA of the ray in the j^{th} scatter level.

5 Algorithms description

In this Section, we describe the main algorithms which will be implemented in the FRM$_4$DOAS processing system.

5.1 DOAS spectral fitting algorithm

5.1.1 QDOAS software

The spectral fitting in the FRM$_4$DOAS centralized processing system will be done using the QDOAS cross-platform application developed at BIRA-IASB (see http://uv-vis.aeronomie.be/software/QDOAS/). It performs DOAS slant columns retrievals of trace gases from spectral measurements (satellite, ground-based, mobile or aircraft-based instruments). The software is available for Windows, Linux and Mac OS operating systems. It evolves along with the measurement techniques and DOAS research, and has been extensively validated in different campaigns and algorithm comparisons. The flexibility of the user-friendly interface, and the robustness of the algorithms largely contribute to QDOAS' success in the world (about 400 registered users in 50 countries in mid-2017).

QDOAS is written in C/C++. The GUI is built on the open-source version of the QT framework (QT4 or QT5). The software package includes five independent modules:

- A main user interface to configure and run applications
- A powerful convolution tool
- A tool to calculate Ring effect cross sections (Rotational Raman Scattering approach)
- A tool to generate undersampling cross sections
- A command line tool very useful to run applications from batches

The software has the following main features in terms of:

- Analysis:
  - ✓ DOAS/intensity fitting modes
  - ✓ Shift/stretch (in nm) fully configurable for any spectral item (cross section or spectrum)
  - ✓ Possibility to constrain the slant column density of a molecule to the value found in a previous window
  - ✓ Possibility to define gaps or automatically remove spikes within fitting intervals (e.g. to eliminate bad pixels)
- Convolution:
  - ✓ Standard and I_0-corrected convolutions are supported.
  - ✓ Possibility to create an effective slit function taking into account the (finite) resolution of the source spectrum (using a FT deconvolution method)
  - ✓ Asymmetric line shapes, wavelength dependent slit functions
- Calibration:

- ✓ Wavelength calibration and instrumental slit function characterization using a Non-Linear Least Squares fitting approach where measured intensities are fitted to a high resolution solar spectrum degraded to the resolution of the instrument.
- ✓ Possibility to correct for atmospheric absorption and Ring effect
- ✓ Possibility to approximate large shifts by correlation

- File formats:
  - ✓ The program supports a large number of spectrum file formats:
    - Satellites : GOME2, SCIAMACHY (ENVISAT), OMI (AURA)
    - Ground-based : SAOZ, MFC (binary and standard) generated by DOASIS,…
  - ✓ It accounts for specificities of input measurements (e.g., sequential analysis for MAXDOAS measurements).
  - ✓ The output is fully configurable. Analysis results and various data related to the measurements can be saved to ASCII, HDF-EOS5 or netCDF files.

5.1.2 DOAS settings

The DOAS settings and corresponding cross-sections used in the FRM_4DOAS QDOAS processor, which are for a large extent based on those used during the CINDI-2 campaign [see RD-2], are described below. It should be noted that they are baseline settings, meaning that they can be slightly modified based on the characteristics and performance of each individual instrument included in the FRM_4DOAS network.

5.1.2.1 Tropospheric NO_2

The DOAs settings for tropospheric NO_2 in the visible and UV ranges are described in Tables 1 and 2 below.

Table 1: DOAS settings for tropospheric NO_2 and O_4 for the visible range.

Wavelength range	425-490 nm (alternative range: 411-445 nm)
Wavelength calibration method	Calibration based on reference solar atlas [Chance and Kurucz, 2010]
Cross-sections:	
NO_2 (298 K)	Vandaele et al. [1998]
O_3 (223 K)	Serdyuchenko et al. [2014]
O_4 (293 K)	Thalman and Volkamer [2013]
H_2O	HITEMP [Rothman et al., 2010]
Ring	QDOAS Ring cross-sections based on Chance and Kurucz [2010]
Polynomial degree	Order 5 (6 coefficients)
Intensity off-set	Constant

Table 2: DOAS settings for tropospheric NO_2 and O_4 for the UV range.

Wavelength range	338-370 nm
Wavelength calibration method	Calibration based on reference solar atlas [Chance and Kurucz, 2010]

Cross-sections:	
NO_2 (298 K)	Vandaele et al. [1998]
O_3 (223 K)	Serdyuchenko et al. [2014]
O_4 (293 K)	Thalman and Volkamer [2013]
HCHO (297 K)	Meller and Moortgat [2000]
BrO (223 K)	Fleischmann et al. [2004]
Ring	QDOAS Ring cross-sections based on Chance and Kurucz [2010]
Polynomial degree	Order 5 (6 coefficients)
Intensity off-set	Constant

5.1.2.2 Stratospheric NO_2

The DOAs settings for stratospheric NO_2 in the visible and UV ranges are described in Tables 3 and 4 below.

Table 3: DOAS settings for stratospheric NO_2 and O_4 for the visible range.

Wavelength range	425-490 nm (alternative range: 411-445 nm)
Wavelength calibration method	Calibration based on reference solar atlas [Chance and Kurucz, 2010]
Cross-sections:	
NO_2 (220K)	Vandaele et al. [1998]
O_3 (223 K)	Serdyuchenko et al. [2014]
O_4 (293 K)	Thalman and Volkamer [2013]
H_2O	HITEMP [Rothman et al., 2010]
Ring	QDOAS Ring cross-sections based on Chance and Kurucz [2010]
Polynomial degree	Order 5 (6 coefficients)
Intensity off-set	Order 1

Table 4: DOAS settings for NO_2 and O_4 for the UV range.

Wavelength range	338-370 nm
Wavelength calibration method	Calibration based on reference solar atlas [Chance and Kurucz, 2010]
Cross-sections:	
NO_2 (220 K)	Vandaele et al. [1998]
O_3 (223 K)	Serdyuchenko et al. [2014]
O_4 (293 K)	Thalman and Volkamer [2013]
HCHO (297 K)	Meller and Moortgat [2000]
BrO (223 K)	Fleischmann et al. [2004]
Ring	QDOAS Ring cross-sections based on Chance and Kurucz [2010]
Polynomial degree	Order 5 (6 coefficients)
Intensity off-set	Order

5.1.2.3 HCHO

The DOAs settings for HCHO are described in Table 5 below.

Table 5: DOAS settings for HCHO.

Wavelength range	336.5-359 nm
Wavelength calibration method	Calibration based on reference solar atlas [Chance and Kurucz, 2010]
Cross-sections:	
HCHO (297 K)	Meller and Moortgat [2000]
NO_2 (298 K)	Vandaele et al. [1998]
O_3 (223 K)	Serdyuchenko et al. [2014]
O_4 (293 K)	Thalman and Volkamer [2013]
BrO (223 K)	Fleischmann et al. [2004]
Ring	QDOAS Ring cross-sections based on Chance and Kurucz [2010]
Polynomial degree	Order 5 (6 coefficients)
Intensity off-set	Order 1

5.1.2.4 Total O_3

The DOAs settings for total O_3 in the Chappuis bands are described in Table 6 below.

Table 6: DOAS settings for O_3 in the Chappuis band

Wavelength range	450-550 nm (alternative range: 450-520 nm)
Wavelength calibration method	Calibration based on reference solar atlas [Chance and Kurucz, 2010]
Cross-sections:	
O_3 (223 K)	Serdyuchenko et al. [2014]
NO_2 (220 K)	Vandaele et al. [1998]
O_4 (296 K)	Thalman and Volkamer [2013]
H_2O	HITEMP [Rothman et al., 2010]
Ring	QDOAS Ring cross-sections based on Chance and Kurucz [2010]
Polynomial degree	Order 5 (6 coefficients)
Intensity off-set	Order 1

5.2 Lower troposphere profiling algorithms

The parameterized and OEM-based MAXDOAS profiling algorithms selected through the FRM_4DOAS Round-Robin exercise (MAPA MC and MMF, respectively; see [RD-1]) are described in this Section.

5.2.1 MAPA MC parameterized algorithm

The Monte Carlo MAinz Profile Algorithm (MAPA MC) derives vertical profiles of aerosol extinction and trace gas concentration profiles from MAX-DOAS measurements. The algorithm is based on a parameterization of the vertical profile by three parameters (see below; see also Wagner et al., [2011]).

5.2.1.1 Parameterization

Profiles are parameterized by 3 parameters:

1. Integrated column *c*, i.e. AOD (aerosols) or VCD (trace gases),

2. Layer height *h*

3. Shape parameter *s*.

The shape parameter *s* determines the shape of the profile. *s*=1 corresponds to a box profile with height *h*, containing column *c*. For *s*<1, the fraction *s***c* is within the box of height *h*, and (1-*s*)**c* above (exponentially decreasing). *s*>1 represents elevated layers. For details, see Wagner et al. [2011].

5.2.1.2 Forward model

The forward model is realized as a look-up table *L* which relates the parameters *c*, *h*, *s* to an AMF for given geometry (SZA, RAA and elevation angle). *L* is calculated by McArtim [Deutschmann et al., 2011]. MAPA also allows for the usage of LUTs derived from other RTMs.

5.2.1.3 Profile retrieval

In the profile inversion described in Wagner et al. [2011], best matching parameters (i.e. the parameters where the difference between measured and modeled SCDs is minimal in a least-squares sense) are fitted using a Levenberg-Marquardt (LM) algorithm. As LM results generally depend on initial parameters if the parameter space has multiple minima in RMS, a multi-stage approach is used where LM fits are performed on different subsets of the parameter space. This turned out to be a very stable, but computational expensive approach.

The LM profile inversion yields best matching parameters and their confidence intervals. From the parameters, the respective profile can be constructed, but errors on the resulting extinction/concentration are not provided.

In order to define profile errors, the confidence interval of the fitted parameter have been used to define an ensemble of profile parameters rather than only one set by calculating randomly distributed parameters for the given mean and uncertainty. From this parameter ensemble, a mean profile and the respective standard deviation can be derived. This approach has been used for the first FRM$_4$DOAS intercomparison (spring 2017).

This approach, however, has some disadvantages. In particular, the LM confidence interval is symmetric by definition. For best parameters close to the physical limits of the parameter space, the random numbers might thus be unphysical (e.g., *c*<0, *h*<0, *s*<0, *s*>2), and if these parameters are skipped, the resulting mean is biased.

Thus, a revised MAPA version has been developed in autumn 2017 which is completely based on Monte Carlo (MC). Ensembles of matching parameter sets are gained by trial&error, i.e. by covering the parameter space with random numbers and keeping those with the best RMS. The detailed procedure is as follows:

- For the aerosol retrieval, up to four free parameters can be considered: *c*, *h*, *s*, and a scaling factor for O_4. Note that any of these parameters might also be set to a constant value via configuration files.

- For each free parameter, random numbers are generated, covering the valid range as specified in the configuration. Note that in many cases, physical or mathematical bounds exist (c>0, h>0, s>0, s<2).
- For N (default: 100000) sets of random parameters, the corresponding DSCDs are calculated.
- The best match (in terms of RMS) is identified.
- All matches within a defined tolerance (default: 1.3*RMS_{min}) are selected
- If less than n (default: 100) parameter sets fulfil this criterion, a new set of N random parameters is determined for a subrange of the parameter space specified by the hits of the first iteration.
- From all matches, the best one as well as n random hits are kept (to reduce computational costs).
- Profiles are calculated for all matching parameter sets
- The mean and standard deviation of all parameters and profiles are calculated.
- For trace gases, the parameters h and s are determined in the same way (default N: 10000). The VCD is subsequently calculated as the mean of the individual $VCD_i = SCD_i/AMF_i$ for all elevation angles ea_i.

MAPA MC needs about 2 seconds per profile retrieval on a single core PC.

5.2.2 MMF OEM-based algorithm

Mexican MAXDOAS Fit (MMF) is an inversion code for MAXDOAS measurements. It was developed at Universidad Nacional Autónoma de México (UNAM) from 2013-2015. During 2017, it has been improved and further developed at UNAM and the Royal Belgian Institute for Space Aeronomy (BIRA).

MMF consists of a forward model step, very briefly described in Sect. 5.2.2.1 and an inversion step, described in Sect. 5.2.2.4. The calculation of the input parameters for the forward model is presented in Sect. 5.2.2.2 and the calculation of the input parameters for the inversion step is described in Sect. 5.2.2.3.

5.2.2.1 Forward model

MMF uses as forward model the lidort part of the VLIDORT [e.g. Spurr et al., 2001; Spurr, 2006 and 2013] algorithm version 2.7 released in August 2014. In MMF, only the intensity information and not the whole set of Stokes parameters is used. (V)LIDORT is a discrete ordinate finite difference algorithm based on finding solutions for an atmosphere consisting of homogeneous layers using quadrature approximations to the integro-differential radiative transfer equation expressed using Legendre polynomials for the phase function and Fourier series for the intensity. It also provides analytic jacobians. In MMF, (V)LIDORT is configured to return profile jacobians w.r.t. gas absorption layer properties or aerosol total extinction layer properties. Details are given below. The input to the forward model are total layer quantities (i.e. combined gas and aerosol): total optical depth (τ) and single scattering albedo (ω) in the layer, and phase function coefficients β_l and derivatives of these quantities w.r.t. the variation variable x. In case of trace gas retrieval, $x = a_{gas}\ \Delta h$ and in case of aerosol retrieval, $x = \sigma_{aer}\ \Delta h + a_{aer}\ \Delta h$. Here, Δh is the layer thickness, a_{aer} denotes the aerosol absorption coefficient, a_{gas} the trace-gas absorption coefficient, and σ_{aer} the aerosol scattering coefficient.

5.2.2.2 Forward model input parameter calculation

The total layer optical depth is the product of the layer thickness and the sum of all extinction and scattering coefficients:

$$\tau = \Delta h \,.\left(\sigma_{air} + \sigma_{aer} + a_{gas} + a_{aer}\right) \quad (12)$$

where *a*air is the air scattering coefficient due to Raleigh scattering.

The layer combined single scattering albedo can be calculated as:

$$\omega = \left(\sigma_{gas} + \sigma_{aer}\right).\Delta h\ /\tau \quad (13)$$

The combined expansion coefficients are calculated as follows:

$$\beta_l = \left(\beta_{air,l}\,.\,\sigma_{air} + \beta_{aer,l}\,.\,\sigma_{aer}\right)/(\sigma_{air} + \sigma_{aer}) \quad (14)$$

where $\beta_{air,l}$ and $\beta_{aer,l}$ denote the phase function moments of the air and aerosol scattering, respectively.

The gaseous absorption coefficient a_{gas} is calculated according to

$$a_{gas} = VMR\,.\,\xi_\lambda.\rho \quad (15)$$

where *VMR* is the volume mixing ratio of the considered trace gas, ξ_λ is the cross section of the trace gas in question at the middle of the wavelength window that was used for the DOAS retrieval. ρ is the number density of air in the layer in question (calculated at the middle height of the layer).

The gaseous scattering coefficient can be calculated as:

$$\sigma_{air} = Q_{Ray}.\rho \quad (16)$$

and the Rayleigh scattering expansion coefficients β_l as (Spurr et al., 2001):

$$\beta_0 = 1, \beta_1 = 0, \beta_2 = (1-\Delta)/(2+\Delta) \quad (17)$$

For the calculation of the depolarization ratio Δ, the main contributions are from N_2, CO_2, O_2 and Ar. Our implementation follows Bates [1984]. For the calculation of the refractive index of air *n*, we also follow Ciddor (1996). For the calculation of the Raleigh cross section Q_{Ray} we follow the implementation of Goody and Yung [1989; their equation 7.37; see also Platt et al., 2007].

The a-priori total aerosol extinction in each layer is a direct input parameter to MMF, as are the values for the aerosol single scattering albedo ω_{aer}. The product of the two yield directly the aerosol scattering coefficient σ_{aer}. The aerosol absorption coefficient a_{aer} is the layer aerosol extinction times $(1-\omega_{ae})$ and the aerosol phase function coefficients $\beta_{aer,l}$ can be calculated via the Henyey phase function and the asymmetry parameter *g* [see e.g. Hess et al., 1998], where *l* denotes the moment.

$$\beta^l = (2l+1)\,.\,g^l \quad (18)$$

Since jacobians with respect to changes of quantity *x* in each layer are required, the normalized derivatives of the total primary optical quantities, τ, ω, and β with respect to this quantity need to be supplied. In case of trace gas retrieval, $x = a_{gas}\,\Delta h$ and in case of aerosol retrieval, $x = \sigma_{aer}\,\Delta h + a_{aer}\,\Delta h$.

Therefore, what need to be done is to calculate the derivatives of Eq. 12 - Eq. 14 with respect to these quantities. It should be remembered that the β is in fact a vector. The three quantities for aerosol are:

$$\frac{x}{\tau}\frac{d\tau}{dx} = \frac{(\sigma_{aer}+a_{aer})\Delta h}{\tau}\frac{d\tau}{d(\sigma_{aer}+a_{aer})\Delta h} = \frac{\sigma_{aer}+a_{aer}}{\sigma_{aer}+\sigma_{gas}+a_{aer}+a_{gas}} \tag{19}$$

$$\frac{x}{\omega}\frac{d\omega}{dx} = \frac{(\sigma_{aer}+a_{aer})\Delta h}{\omega}\cdot\left[\frac{\partial\omega}{\partial\sigma_{aer}}\frac{d\sigma_{aer}}{d[x(\sigma_{aer}+a_{aer})\delta h]}+\frac{\partial\omega}{\partial a_{aer}}\frac{da_{aer}}{d[x(\sigma_{aer}+a_{aer})\delta h]}\right]$$

$$= \frac{(\sigma_{aer}+a_{aer})\Delta h}{\omega}\cdot\left[\frac{1-\omega}{e}\frac{\omega^*}{\Delta h}+\frac{-\omega}{e}\frac{(1-\omega^*)}{\Delta h}\right] = \frac{(\sigma_{aer}+a_{aer})}{\omega e}[(1-\omega)\omega^*-\omega+\omega\omega^*]$$

$$= \frac{\sigma_{aer}+a_{aer}}{\omega e}(\omega^*-\omega) = \frac{\sigma_{aer}a_{gas}-a_{aer}\sigma_g}{\omega e^2} \tag{20}$$

$$\frac{x}{\beta}\frac{d\beta}{dx} = \frac{\sigma_{aer}}{\beta}\frac{\beta_{aer}-\beta}{\sigma_{aer}+\sigma_{gas}} \tag{21}$$

where e is the total extinction coefficient: $\tau = e\,\Delta h$ and $\omega^* = \frac{\sigma_{aer}}{\sigma_{aer}+a_{aer}}$ is the aerosol single scattering coefficient.

For the trace gas jacobian calculation, the corresponding quantities are:

$$\frac{x}{\tau}\frac{d\tau}{dx} = \frac{a_g\Delta h}{\tau}\frac{d\tau}{da_g} = \frac{a_g}{e} \tag{22}$$

$$\frac{x}{\omega}\frac{d\omega}{dx} = \frac{a_g\Delta h}{\tau\omega}\frac{d\omega}{da_g} = -\frac{a_g}{e} \tag{23}$$

$$\frac{x}{\beta}\frac{d\beta}{dx} = 0 \tag{24}$$

5.2.2.3 Inverse input parameter calculation

The forward model output needs to be converted into differential slant column densities (dSCD). For each set of dSCD, the forward model is run 2 times: with and without the gas absorption included. Each of these two sets consists of simulations in the desired off-axis telescope directions and one towards the zenith. If we call the intensity and jacobian of the simulation of the off-axis angles without gas absorption $I_0^\alpha$and K_0^α, the intensity (jacobian) with gas absorption towards the zenith I_g^{zen} $\left(K_g^{zen}\right)$, the intensity (jacobian) with gas absorption towards the desired angle I_g^α $\left(K_g^\alpha\right)$, and the intensity (jacobian) without gas absorption towards the zenith I_0^α (K_0^α), then the dSCD and weighting function K can be calculated as

$$dSCD = \log\left(\frac{I_0^\alpha . I_g^{zen}}{I_g^\alpha . I_0^{zen}}\right)/\xi(\lambda) \tag{25}$$

and

$$K = \frac{\left(K_0^{\alpha} I_g^{zen} I_g^{\alpha} I_0^{zen} + I_0^{\alpha} K_g^{zen} I_g^{\alpha} I_0^{zen} - I_0^{\alpha} I_g^{zen} K_g^{\alpha} I_0^{zen} - I_0^{\alpha} I_g^{zen} I_g^{\alpha} K_0^{zen}\right)}{\left(I_0^{\alpha} I_g^{zen} I_g^{\alpha} I_0^{zen} \xi(\lambda)\right)} \quad (26)$$

5.2.2.4 Inversion

The inversion in MMF is based on a Levenberg-Marquardt (LM) iteration scheme and uses least square fitting with either optimal estimation (OE) or Tikhonov regularization. This, as well as whether to read in an uncertainty covariance matrix S_a or to calculate this from the specified a-priori profile according to Eq. 27 is specified via an input setting.

$$S_{a_{i,j}} = \sqrt{S_{a_{i,i}} S_{a_{j,j}} exp^{-\ln(2)\frac{(h_i - h_j)^2}{c_l}}} \quad (27)$$

For the Tikhonov setting, the regularization matrix is constructed from the discrete first derivative operator. Alternatively, a regularization matrix can be provided directly via an input file. Only the OE setting with construction of S_a according to Eq. (27) is used in the FRM$_4$DOAS project. The code can operate in retrieval log-space or retrieval lin-space. Only log-space settings are used for the FRM$_4$DOAS project. A-priori profiles and values for g and ω are provided via input files. A temperature-pressure profile needs to be provided as well. The trace gas a-priori profile can be either provided in terms of VMR in ppb or in terms of number densities.

The a-priori values, as well as the temperature-pressure profile can be provided on different grids. All values are internally interpolated on the simulation grid. This grid, as well as a specification of the number of layers (from the surface) which are used as the retrieval grid, need to be provided via an input file. The layers do not need to be equidistant. For example, applied for MAXDOAS measurements in Mexico City, an almost equidistant pressure grid has been used.

For both aerosol and trace-gas retrieval, the equation that is solved is the iterative LM equation for non-linear inversion problems (Rodgers, 2000, Equation 5.36):

$$x_{i+1} = x_i + \left[(1+\gamma)S_a^{-1} + K_i^T S_\varepsilon^{-1} K\right]\left[K_i^T S_\varepsilon^{-1}\left(y - F(x_i)\right) - S_a^{-1}(x_i - x_a)\right] \quad (28)$$

Here, superscript T denotes transposed, superscript -1 denotes the inverse. The index *i* is the iteration index, S_ε the measurement error covariance matrix, *y* the measurement vector (for MMF this is the vector of differential slant column densities*)*, $F(x_i)$ the simulated differential slant column densities, calculated as in Eq. (25) using the forward model with input profile x_i and K_i is the jacobian matrix at the *i*-th iteration, calculated according to Eq. (26). If the retrieval is to be performed in log-space, i.e. to work with *ln(x)* instead of with *x* as the retrieval parameter, *K* in Eq. (26) is multiplied by *x* and the uncertainty covariance matrix is left and right multiplied by *1/x*.

5.3 Total O_3 column algorithm

Total O_3 VCDs are retrieved from twilight zenith-sky UV-vis observations using the standard NDACC approach [see DR-3; Hendrick et al., 2011]. Since the main theoretical principles of this method are

given in Section 4.2.3, only the NDACC settings are described here, in particular the AMF climatology used to convert O_3 slant columns into vertical columns.

The standard NDACC O_3 AMF climatology is based on the TOMS version 8 (TV8) ozone and temperature profile climatology [Barthia et al., 2004]. It consists of 18 LUTs generated using the UVSPEC/DISORT RTM [Mayer and Kylling, 2005] initialised with TV8, each of these LUT corresponding to one TV8 latitude (10° latitude bands between 90°S and 90°N) [DR-3; Hendrick et al., 2011]. The other entry parameters are: wavelength, ground albedo, altitude of the station, day of the year, and SZA. The extraction of appropriate O_3 AMFs for the FRM_4DOAS stations is done by using the dedicated interpolation routine in Fortran developed in the framework of the NDACC UV-vis WG [DR-3].

For the selection of the SZA range representative of twilight conditions, the best compromise between accuracy and precision is achieved in the 86-91° SZA range. The recommended NDACC approach is to apply a linear fit on VCDs in the above SZA range and then derive the column value at an effective SZA, which is usually fixed at 90°.

5.4 Stratospheric NO_2 profiling algorithm

The OEM-based profiling algorithm developed by BIRA for the inversion of NO_2 vertical profiles from twilight zenith-sky NO_2 DSCDs [see Hendrick et al., 2004] will be implemented in the FRM_4DOAS processing system. The main theoretical principles on which this algorithm is based are given in Section 4.2.4. Only the settings specific to the BIRA tool will be described here.

The forward model consists of the stacked box photochemical model PSCBOX [Errera and Fonteyn, 2000] coupled to the RTM package UVSPEC/DISORT [Mayer and Kylling, 2005]. As mentioned in Section 4.2.4, a photochemical model is required to reproduce the effect of the rapid variation of NO_2 concentrations at twilight and the RTM is used to calculate SCDs/WFs from the NO_2 concentrations predicted by the photochemical model.

The PSCBOX model includes 48 variable species, 141 gas-phase photochemical reactions as well as heterogeneous reactions on liquid sulfuric acid aerosols and on solid nitric acid trihydrate (NAT) and ice particles. It is initialised daily at 20 independent altitude levels (between ~10 and ~55 km of altitude) with 0h UT meteorological and chemical fields from the 3D CTM SLIMCAT [Chipperfield, 2006]. The chemical timestep is 6 minutes; no family and photochemical equilibrium assumptions are made. Updated kinetic and photochemical data are taken from the JPL 2006 compilation [Sander et al., 2006]. Photolysis rates are computed off-line by using the radiative transfer module of the two-dimensional model SOCRATES [Huang et al., 1998].

The UVSPEC/DISORT RTM is based the discrete ordinate method and includes a treatment of multiple scattering in a pseudo-spherical geometry approximation. All WF calculations are performed in multiple scattering mode, accounting for the 2D (altitude, SZA) photochemical variation of NO_2 (incident beam only; see Figure 3). The standard altitude grid is 0-90 km with a layer thickness of 2 km. The aerosol extinction profile has been constructed from the aerosol model of Shettle [1989] included in UVSPEC/DISORT. In the stratosphere, the aerosol settings correspond to typical background conditions while in the boundary layer and troposphere, settings representative of a

rural environment are used. It should be noted that the UVSPEC/DISORT RTM has been validated through several comparison exercises [Hendrick et al., 2006; Wagner et al., 2007].

In order to retrieve stratospheric NO_2 vertical profiles for globally-distributed stations, LUTs of WFs and 2D (SZA, altitude) fields of NO_2 concentrations are generated for 10° latitude bands between 90°S and 90°N (longitude is fixed to 0°) using the UVSPEC/DISORT-PSCBOX model package initialized with corresponding 3D-CTM meteorological and chemical fields. The other entry parameters are day of the year, altitude of the station, and ground albedo. The a priori profiles needed for the OEM inversion method are extracted from the LUTs of 2D NO_2 concentration fields. Since the standard retrieval altitude range is 0-90km and photochemical model simulations give profile data only from ~10 to ~55 km altitude, the a priori profiles should be complemented below and above this altitude range. Below the lowest model altitude level, a priori profile values are calculated using the following expression: $x_{a,\ level\ i} = 0.5\ x_{a,\ \text{level i+1}}$. Thus, the NO_2 tropospheric content in the a priori profile is made negligible for all retrievals. Above the highest model altitude level, the AFGL standard atmosphere completes the NO_2 profile.

Regarding the construction of the a priori covariance matrix S_a, the diagonal values will correspond to a given percentage of the a priori profile and this percentage will be empirically determined for each FRM_4DOAS station based on sensitivity tests. The extra-diagonal terms will be added as Gaussian function with a correlation length of 8km, as described in Section 4.2.2.1.

5.5 Error analyses

In this section, the error analysis methods used for the different FRM_4DOAS data products are described.

This Section should be updated depending on the error assessment methods that will be implemented in practice in the processing system. A sub-section on systematic errors related to the XS could be also added.

5.5.1 Tropospheric products by the MC-based parameterized approach

When a Monte Carlo approach is used as in the MAPA algorithm, an ensemble of parameter sets is derived and from this ensemble, the mean and standard deviation of parameters, as well as the corresponding extinction/concentration profiles, are determined. The error of AOD and VCD is thus provided by the standard deviation of the column parameter *c*.

The fit RMS holds information on the fit quality and is used to define cases where the profile inversion is likely affected by systematic errors, which might be caused e.g. by clouds, or profile shapes which are not covered by the chosen parameterization.

The comparison of the parameters mean with the best matching parameter set, as well as the standard deviation of the parameter ensemble, provide additional information on the consistency of the resulting profile and is used for flagging purpose. The thresholds for flagging are specified via configuration files.

5.5.2 Tropospheric and stratospheric products by the OEM approach

In the OEM, the total error of the retrieved profile is the sum of three errors [Rodgers, 2000]: the error due to the smoothing of the true profile or smoothing error, the error due to random and systematic errors in the measurements, and the error due to systematic errors in the forward model.

The smoothing error covariance matrix S_s can be calculated using the following expression [Rodgers, 2000]:

$$S_s = (A - I)\, S_x\, (A - I)^T \quad (29)$$

where S_x is a realistic covariance matrix of the true profile, A is the averaging kernels matrix (see Sect. 5.6) and I is the identity matrix. S_x also contains extra-diagonal terms in order to account for correlations between trace gas concentration values at different altitude levels. In a first approximation, the S_x matrix is generally chosen identical to S_a.

The retrieval noise S_m which is the retrieval error covariance resulting from measurement error is defined as [Rodgers, 2000]:

$$S_m = G\, S_\varepsilon\, G^T \text{ with } G = \frac{\partial \hat{x}}{\partial y} = S_a K^T (K S_a K^T + S_\varepsilon)^{-1} \quad (30)$$

where S_ε is the measurement error covariance matrix, and G is the contribution functions matrix expressing the sensitivity of the retrieved profile to changes in the measured trace gas slant column abundances. As mentioned in Sect. 4.2.2.1, the S_ε matrix was chosen diagonal with values corresponding to the statistical errors on the DOAS fitting.

The forward model parameter error S_f is the retrieval error due to errors in the forward model parameters (e.g., errors on the rate constants in the photochemical model). S_f is given by the following expression [Rodgers, 2000]:

$$S_f = G\, K_b\, S_b\, K_b^T\, G^T \quad (31)$$

where G is the contribution functions matrix (see Eq. 30), K_b is the sensitivity of the forward model to perturbations of forward model parameters b, and S_b is the covariance matrix of b. S_f cannot be determined easily due to the large number of forward model parameters.

In the FRM$_4$DOASprocessing system, the S_f matrix for stratospheric NO_2 profile will be taken from Preston et al. [1997]. This one is derived by calculating the sensitivity of the slant column measurements to large impact forward model (photochemical and/or RT models) parameters like O_3, HNO_3, N_2O_5, aerosol, temperature profiles and ground albedo. In the case of MAXDOAS trace gas profiling, the uncertainty on retrieved aerosol profiles is generally the main error source related to the forward model parameters and is estimated using sensitivity tests [e.g. Clémer et al., 2010; Hendrick et al., 2014; Wang et al., 2017].

5.5.3 Total O_3 column

The error budget of total O_3 columns is obtained by considering error sources affecting the determination of the DSCD, RCD, and AMF (see Eq. 11). These error sources are summarized below, based on Hendrick et al. [2011].

5.5.3.1 Error on DSCD

Errors associated to the least-squares spectral fit are due to detector noise, instrumental imperfections (small wavelength scale and resolution changes, etaloning and non-linearities of the detector, stray-light, polarisation effects, ...) as well as errors or unknowns in the signal modelling (Ring effect, unknown absorbers, wavelength dependence of the AMF, etc.). To some extent, such errors are pseudo-random in nature and, as such, can be estimated statistically from the least-squares fit procedure. Fitting errors derived from the least-squares analysis typically give small uncertainties of the order 5 DU for O_3 DSCDs. However, results from intercomparisons exercises [e.g. Van Roozendael et al., 1998; Vandaele et al., 2005; Roscoe et al., 2010] show that state-of-the-art instruments hardly agree to better than a few percents, even using standardised analysis procedures, which indicates that the actual accuracy on the DSCDs is limited by uncontrolled instrumental and/or analysis factors. Based on experience and results from intercomparison campaigns, the random uncertainty on O_3 DSCD is of the order of 3 %. This error adds up to systematic uncertainties on ozone absorption cross sections in the Chappuis bands and on their limited temperature dependence which is of the order of 3 % in the 450-550 nm spectral range [Orphal, 2003].

5.5.3.2 Error on RCD

The uncertainty on the determination of RCD is limited by the method used to derive the O_3 VCD at the time of the reference spectrum acquisition. Using a Langley-plot approach leads to an uncertainty of the order of 2% [Van Roozendael et al., 1994; Sarkissian et al., 1997].

5.5.3.3 Error on AMF

A first source of uncertainty in the NDACC O_3 AMF LUTs is related to the use of the TV8 O_3 profile climatology, originally designed for nadir backscatter measurements from space. In order to test the validity of this climatology, mean O_3 AMFs in the 86-91°SZA range ($AMF_{86\text{-}91°SZA}$) extracted from the LUTs have been compared to calculations performed using O_3 profiles measured with ozonesondes and/or lidar observations at 9 NDACC stations representative of a wide range of conditions (tropics, mid- and high-latitudes). The mean relative difference between O_3 $AMF_{86\text{-}91°SZA}$ extracted from the LUTs and those calculated with the measured O_3 profiles is of the order of -1 ± 1.3%. Although some residual seasonalities in the mean relative differences are found at some stations, these results show that the TV8 climatology reproduces well on average the mean O_3 profiles latitudinal and seasonal variations, so that sufficiently accurate O_3 AMFs can be calculated when using it.

Clouds are not accounted for in the NDACC O_3 AMF calculations but their impact has been investigated using the water clouds model included in UVSPEC/DISORT. O_3 AMFs have been calculated for cloudy and non-cloudy conditions for TV8 climatology O_3 profiles corresponding to 25°N/275DU, 45°N/325DU, and 65°N/325DU in June. For cloudy conditions, the cloud model parameters values are fixed as follows: water content: 0.3 g/m^3, effective droplet radius: 5 µm, cloud layer thickness and altitude: 1 km between 1 and 2 km. Cloudy AMFs are found to be systematically larger than non-cloudy AMFs by about 5-8% at 86°SZA and 2% at 91°SZA. The relative difference in the 86-91°SZA range for the three selected O_3 profiles is of 3.3% on average. Similar comparison results are obtained for winter O_3 profiles. The small impact of clouds on zenith-sky ozone UV-vis measurements at twilight is due to the fact that the mean scattering layer is generally located at higher altitude than that of the clouds. However, there are two exceptions: in the tropics where

thunderstorms accompanied by heavy rainfall can reach 15-16 km, and at high latitude in the winter where PSCs are sometimes present, disturbing the ozone measurements. These episodes can be easily filtered out from the ground-based data series by detecting enhancements larger than 70% or more of the absorption by O_4 and H_2O in the tropics in the presence of thick clouds and rainfall, and by the use of a color index (ratio between irradiances at 550 and 350 nm) in case of PSCs [Sarkissian et al., 1991].

Another source of uncertainty we have tested is the impact of surface albedo. O_3 AMFs corresponding to the same three TV8 climatology O_3 profiles as above have been calculated using the UVSPEC/DISORT RTM with albedo fixed to 0.04 (ice free sea) and 1 (fresh snow, sea ice or thunderstorm anvils). Within the 86-91°SZA range, the impact of surface albedo is rather low with a mean relative difference between albedo 1.0 and 0.04 cases of 0.7% for the three selected O_3 profiles.

A last source of uncertainty is the impact of the RTM used for AMF calculations. Although previous studies [e.g., Hendrick et al., 2006; Wagner et al., 2007] have demonstrated that, for AMF calculation, the UVSPEC/DISORT model shows very good consistency with others RTMs, a verification exercise has been carried out to firmly assess the reliability of the present O_3 AMF calculations. It consisted in comparing O_3 AMFs calculated using the UVSPEC/DISORT and SCIATRAN v2.2 [Rozanov et al., 2005] RTMs initialized in the same way. Regarding the O_3 profile, the following cases have been considered: polar latitude in January and June (65°N and S with a total column of 325 DU, mid-latitude in January and June (45°N and S, 325 DU), and tropics in January and June (25°N and S, 275 DU). It is found that, within the 86-91°SZA range, the mean relative difference is about 0.7%, indicating the reliability of the UVSPEC/DISORT RTM for O_3 AMF calculation.

5.5.3.4 Total error on VCD

Typical total error budget of twilight zenith-sky visible total O_3 total columns measurements is summarized in Table 7. The precision by adding quadratic random errors is 4.7 % to which the largest contribution comes from the AMF and from the error on the slant column estimated to be 3 % at twilight, including the impact of unknown instrumental and systematic misfit effects. The total uncertainty, important for comparison with other instruments, is ~5.9 %.

Table 7: Error budget on total O_3 column.

Error source	Error (%)
a) Random	
Slant column spectral fit, including interference effects	3
O_3 AMF	
TV8 climatology	1.0
Clouds	3.3
Aerosols	0.6
Albedo	0.7
RTM	0.7
Precision	**4.7**

b) Systematic	
O_3 cross sections	3.0
Residual column	2.0
Total Accuracy	**5.9**

5.6 Vertical resolution/Averaging kernels

The vertical sensitivity of the ground-based MAX(DOAS) measurements to the vertical distribution of trace gases is usually characterized using the AKs [Rodgers, 2000], which correspond to the rows of the averaging kernel matrix *A*. This matrix is derived using the following expression [Rodgers, 2000]:

$$A = \frac{\partial \hat{x}}{\partial x} = (K^T S_\varepsilon^{-1} K + S_a^{-1})^{-1} K^T S_\varepsilon^{-1} K \quad (32)$$

AKs express the relationship between the retrieved profile $\hat{x}$ and the true atmospheric profile x through the following equation:

$$\hat{x} = x_a + A(x - x_a) + \text{Error terms} \quad (33)$$

Following Eq. (33), the retrieval of any profile point is an average of the entire true profile weighted by the row of the *A* matrix corresponding to the altitude of the retrieved profile point. For an ideal observing system, the *A* matrix would be therefore equal to the identity matrix. In the reality, the retrieved profile is only a smoothed perception of the true profile. The vertical resolution of this smoothed information at a given altitude can be estimated by taking the FWHM of the main peak of the corresponding averaging kernel. Another important characterization parameter which can be derived from the *A* matrix is the number of degrees of freedom for signal (DOFS) providing an estimate of the number of independent pieces of information that can be retrieved from the measurements. This parameter is given by the trace of *A* [Rodgers, 2000].

Figures 4 and 5 show typical examples of AKs for stratospheric and tropospheric NO_2 OEM-based retrievals, respectively.

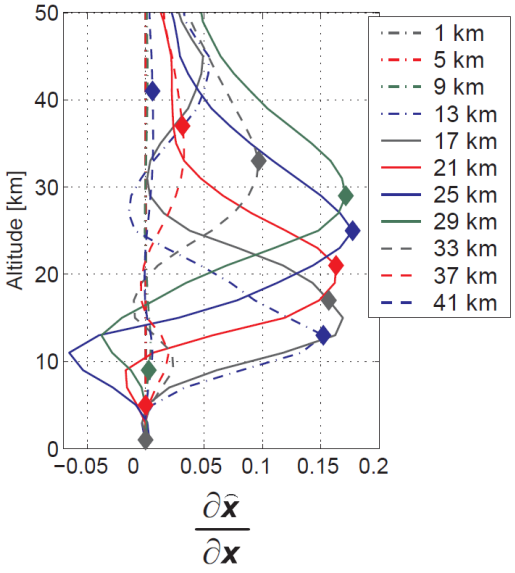

Figure 4: Typical example of AKs for stratospheric NO_2 OEM-based retrieval. Plain diamonds indicate the nominal altitude at which each averaging kernel should peak. [from Hendrick et al., 2004].

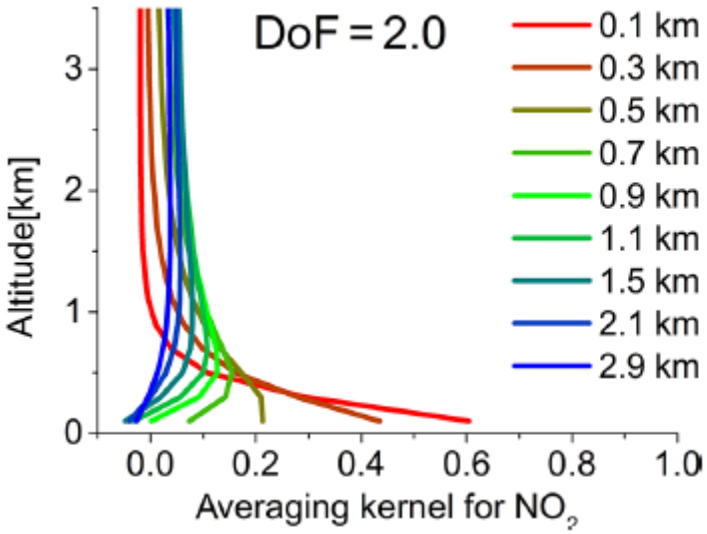

Figure 5: Typical example of AKs for tropospheric NO_2 OEM-based retrieval. [from Wang et al., 2017].

From Preston et al. [1997] and Hendrick et al. [2004], it has been found that 13-37 km is the altitude range where twilight zenith-sky measurements give significant information about the vertical distribution of NO_2. From Figure 4, we also see that the vertical resolution is about 8 km at 13 km of altitude and reaches 20 km at an altitude of 33 km. Typical values for the trace of *A* are about 2, meaning that there are about 2 independent pieces of information in the measurements.

The examination of the MAXDOAS NO_2 AKs in Figure 5 indicates that the inversions are sensitive to the layers from the surface up to 1-1.5 km. The corresponding DOFS is about 2. It should be noted that, inherent to the parameterized approach, AKs cannot be derived in the case of the MAPA MC algorithm and therefore the vertical sensitivity cannot be determined.

A typical example total O_3 column AK is shown in Figure 6. It has been computed according to Eskes and Boersma [2003], i.e. the AK of layer *l* can be approximated by the ratio between the box-AMF of layer *l* and the total AMF, using the UVSPEC/DISORT RTM [Mayer and Kylling, 2005; see also Section 5.4] initialized with O_3, pressure, and temperature profiles from the TV8 climatology. Since the mean scattering layer is located around 14 km altitude, the sensitivity of zenith sky twilight measurements to tropospheric ozone is limited, with averaging kernel value smaller than 0.5 below 8 km, and increases in the stratosphere where averaging kernel value is larger than 1 between 14-30 km altitude. So, the twilight zenith-sky UV-vis total column O_3 measurements are strongly weighted by the contribution of the stratosphere and therefore show very limited sensitivity to the uncertainties on parameters affecting tropospheric ozone like e.g. Mie scattering in a cloud layer. However, these measurements are sensitive to the tropospheric ozone column used for the AMF calculation which acts as a ghost column in the total column retrieval.

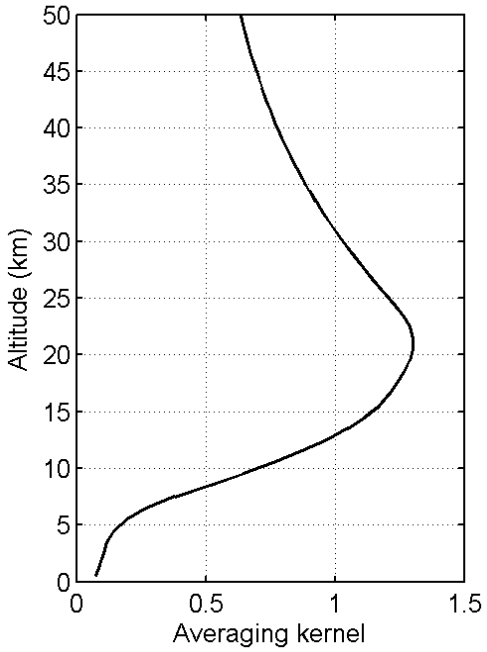

Figure 6: Column AK computed for 90°SZA in zenith-sky geometry using O_3 and temperature profiles corresponding to 45°N/325 DU in June extracted from the TV8 climatology. The wavelength is fixed to 500 nm [from Hendrick et al., 2011].

5.7 Horizontal representativeness

5.7.1 Introduction

In contrast to *in-situ* observations such as an ozone monitor located at the surface, remote sensing observations average over space and time providing integrated values representative for values averaged over a certain volume. When evaluating the air volume for which these measurements are representative, quantities like vertical and horizontal resolutions of the measurements and horizontal displacement of the measured air volume are usually considered. One can go even one step further by investigating for which area and volume the measurements are representative given

the history of the air masses and the atmospheric lifetime of the quantity of consideration [see RD-3].

In the context of FRM$_4$DOAS, the vertical resolution of the data products is evaluated and quantified using the AKs (see Section 5.6). Regarding the evaluation of the horizontal displacement, simple parameterized approaches are usually used and those are described below.

5.7.2 Twilight zenith-sky products (total O_3 column and stratospheric NO_2 profiles)

For zenith-sky observations, the horizontal displacement of the measurement volume depends mainly on *SZA* and the light path has in first approximation two separate parts: one from the sun to the scattering point in altitude *z* which is determined by the solar position, and a second one passing vertically through the atmosphere from altitude *z* to the instrument on the ground (Figure 7). The horizontal displacement *x* from the instrument location can be computed for each scattering height *z* by:

$$x(z,i) = (i - z).\tan(SZA) \qquad (34)$$

For the effective signal, the average over the horizontal displacements for all scattering altitudes *z* needs to be calculated, weighted by the respective radiation received from this altitude at the surface *WF(z)*:

$$d(i) = \frac{\sum_{z,i=0}^{top-atmosphere} x(z,i).WF(z)}{\sum_{z,i=0}^{top-atmosphere} WF(z)} \qquad (35)$$

The resulting distances *d,* calculated by using a simple ray-tracing model, are shown in Figure 8 and can be as large as 500 km in the stratosphere at twilight.

In FRM$_4$DOAS, LUTs of distance *d* with *SZA* and altitude as entries are used in combination of the azimuth solar angle to report the mean location of the effective air mass corresponding to the retrieved twilight total O_3 columns and stratospheric NO_2 profiles.

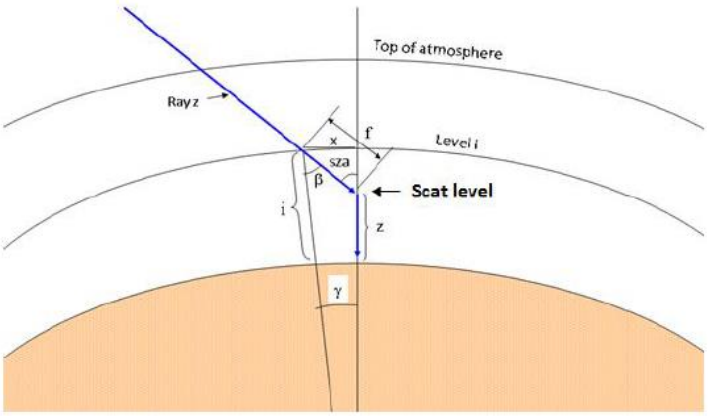

Figure 7: Sketch of the light path for zenith-sky observations. Indicated is the distance *x* that the measurement point has at altitude level *i* for a scattering altitude *z* at solar zenith angle SZA [from RD-3].

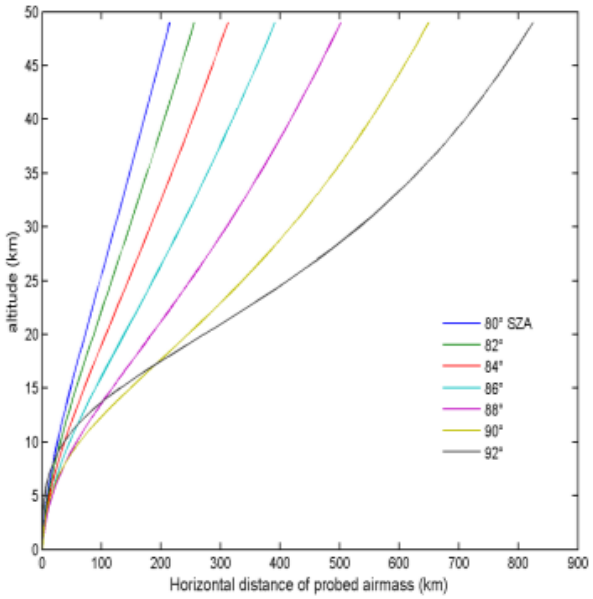

Figure 8: Average displacement of the measurement point in zenith viewing geometry for different altitudes at different solar zenith angles [from RD-3].

5.7.3 Tropospheric products

Compared to zenith-sky observations, the characterisation of horizontal displacement is further complicated. While in single scattering approximation the light path can still be divided into two parts, the first being determined by the solar position, the second part from the scattering point to the instrument depends on the viewing azimuth and elevation of the instrument. Also, the distance of the last scattering point from the instrument is strongly dependent on wavelength and aerosol load and profile and can only be determined from the measurements themselves using the O_4 absorption. This is illustrated in Figure 9, where 2d box air mass factors (box-AMF) are shown for two wavelengths without aerosols (left and middle) and with an AOD of 0.5 in the lowest 1 km. Without aerosol, horizontal displacements of up to 200 km are possible at 500 nm, of about 50 km at 360 nm and much less in the presence of aerosols.

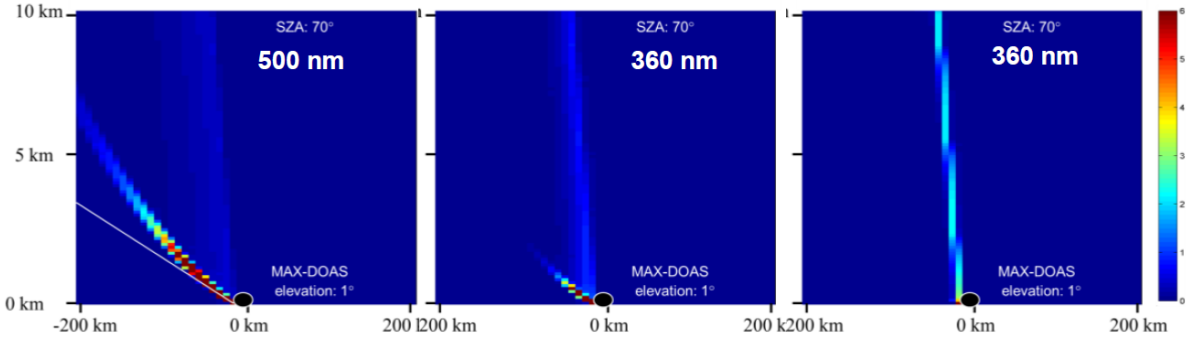

Figure 9: 2d box AMF for MAX-DOAS observations at 1° elevation and an SZA of 70°. Left: 500 nm, no aerosol. Middle: 360 nm, no aerosol. Right: 360 nm, AOD 0.5 in lowest km. All calculations have been done using a horizontal grid of 8 x 8 km^2 and a vertical grid of 100 m. The deviation of the line of high box-AMF from a straight line is caused by the Earth's curvature [from RD-3].

In FRM$_4$DOAS, the horizontal displacement is derived using a universal 2D polynomial function of the SZA and the relative azimuth angle (RAZI), which describes the ratio between the measured O_4 DAMF ($AM_{90°}$-$AMF_{off\text{-}axis}$) and the horizontal distance (in km). This 2D function has been determined based on RTM simulations of O_4 DAMF for different wavelength (360nm, 477nm, 577nm, and 630nm), SZA (0, 10, 20, 30, 40, 50, 60, 70, 80, 85, 88, 90°), RAZI (0, 5, 10, 20, 30, 50, 70, 90, 120, 150, 180), elevation angle (1, 2, 3, 6°), and AOD (0, 0.05, 0.1, 0.2, 0.5, 1, 2, 3) scenarios [Wagner and Beirle, 2016].

It should be noted that for retrievals based on elevation angles higher than 6° (e.g. 15° or 30° as often used for the geometric approximation), the effect of the atmospheric visibility is much weaker. Also, the horizontal distance, for which these observations are sensitive, is rather small: for a trace gas layer between the surface and 1km, the horizontal sensitivity range will e.g. be only between about 2 and 4 km for elevation angles of 30° and 15°, respectively. Thus in this case, the horizontal sensitivity range (D) for elevation angles (α) larger than about 10° is parameterized by the following linear relationship with the altitude (h):

$$D = h/\tan(\alpha) \qquad (36)$$

5.8 Cloud conditions

5.8.1 Introduction

MAXDOAS retrievals of trace-gas and aerosol profiles and columns typically assume clear-sky conditions in forward model simulations. However, MAXDOAS measurements, in contrast to twilight zenith-sky observations, are often strongly affected by clouds, leading to significant data quality degradation and larger uncertainties on the retrievals. This, in turn, strongly impairs the use of ground-based retrievals in the context of satellite and model validation.

Cloud-free sky conditions are ideal for the profile inversion, as under cloudy skies the atmospheric light paths are complicated, especially for rapidly changing cloud conditions. In principle, it would be possible to include clouds in the radiative transfer simulations, but usually the necessary information on cloud properties is not available [e.g., Erle et al., 1995; Wagner et al., 1998, 2002, 2004; Winterrath et al., 1999]. So, it is important to identify and classify clouds and aerosols for each measurement in order to characterise the quality of the retrieval.

Cloud information derived from MAXDOAS observations, instead of other sources like e.g. visual inspection or camera images, is very important, because it can be directly assigned to individual MAXDOAS observations without any spatio-temporal interpolation and without requiring the installation of additional instrumentation. This is especially important for future harmonised MAXDOAS data processing in global monitoring networks of tropospheric species.

Several cloud-screening methods using MAX-DOAS measurements have been developed recently [Wagner et al. 2014; Gielen et al. 2014; Wang et al. 2015; Wagner et al. 2016; Wang et al., 2017; see also RD-3], based on parameters such as the colour index and the absorption of the oxygen dimer. These cloud-screening schemes typically differentiate between a few primary sky conditions, such as:

clear sky (with low/high aerosol load), continuous clouds, cloud holes, broken/scattered clouds, and secondary conditions such as thick clouds or fog.

5.8.2 Cloud screening approach applied in FRM$_4$DOAS

In the FRM$_4$DOAS centralized processing system, a simplified version of the color-index-based cloud screening method described in Gielen et al. [2014] is used to identify the following primary sky conditions: clear-sky/thin clouds, thick clouds, and broken clouds. In order to make it applicable to all stations, only information derived from spectra is needed in this approach (so, e.g. no need for RTM simulations).

5.8.2.1 Color index calculation

The following steps are used to calculate the color index (*CI*) based on measured spectra:

- Selection of two wavelengths covering a large spectral range of your observed data, making sure that the selected wavelenghts are not affected by strong spectral features.
- Calculation of the color index as the ratio of the intensity of the lowest ($I_{\lambda\ low}$) over the highest ($I_{\lambda\ high}$) wavelength :

$$CI = I_{\lambda\ low} / I_{\lambda\ high} \qquad \text{(xx)}$$

- Using this definition, higher values for the *CI* points to clearer skies, whereas low values point to polluted skies (aerosols/clouds). For extremely polluted skies, the values are observed to cluster around a minimum value. It should be noted that measurements with *SZA* > 85° are not taken into account since they do no longer allow for a clear distinction between different sky conditions.
- Although applicable to other viewing directions, only the *CI* corresponding to the zenith elevation is used here.

5.8.2.2 General sky condition flagging

The general sky condition flagging is made by investigating the observed frequency distribution of the color index. If enough data is present and some sky conditions are more frequent, a peak in the frequency distribution is observed. For sites often experiencing thick clouds, this peak will correspond to cloudy condition; for sites where clouds are very unusual, the peak will correspond to clear-sky conditions. Some sites might show a double-peaked distribution, if clear skies and cloudy conditions both occur with a similar frequency.

The histogram analysis is based on the following steps (see Figure 10):

- Make a histogram of your color index values.
- Determine the peak of the *CI* distribution (CI_{max}) and calculate the full-width-at-half-maximum (FWHM).
- Define a region [*CImax* − *FWHM*,*CImax* + *FWHM*] (red region in Figure 10).
- Mark points in this region as 'Thick clouds/Extreme pollution', the other data points then correspond to 'Clear sky' or 'Thin clouds/pollution'.

- The distinction between green ('clear-sky') and orange ('thin clouds) regions in Figure xx was based on RTM simulations, which goes beyond the scope of the present simplified approach.

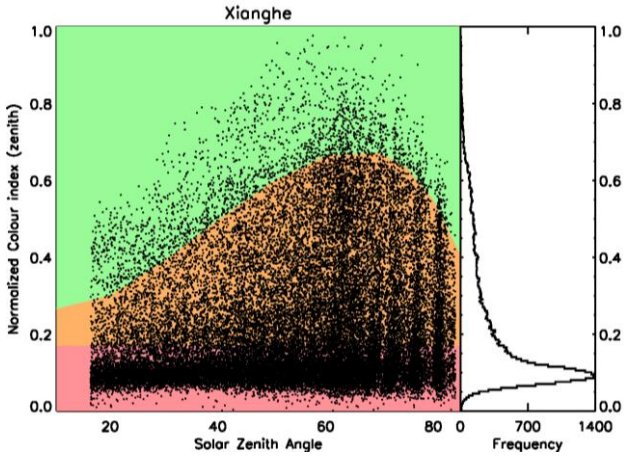

Figure 10: Normalized CI versus SZA distribution from BIRA MAXDOAS observations at the Xianghe station (Beijing suburban area). In color, we defined the regimes corresponding to clear-sky (green) and thick clouds (red). The red box was constructed around the maximum of the histogram, corresponding in this case to cloudy days. The green box (distinction between clear-sky and thin clouds conditions) was defined by a RTM simulation corresponding to aerosol AOD = 0.15 and cloud optical depth = 0.

It should be noted if a site experiences little cloudy conditions, the observed peak in the CI distribution will correspond to clear-sky MAXDOAS scans, and the method above will be used to detect clear-sky conditions, instead of cloudy conditions.

5.8.2.3 Broken cloud flagging

To determine the presence of broken clouds, the following model approach is used:

- Model the data with a double-sine function $f(x) = a + b\sin(cx - d) + e\sin(fx - g)$ – See Figure 11.
- Identify outliers: $|\ (CI - model)/model\ | > 0.1$. The threshold value of 10% is empirically chosen to include most outliers but not the small variations. This value can be adapted, depending on the station.
- Flag each outlier as being affected by broken clouds and leave this flag empty if there are not enough data points to perform a good fit.

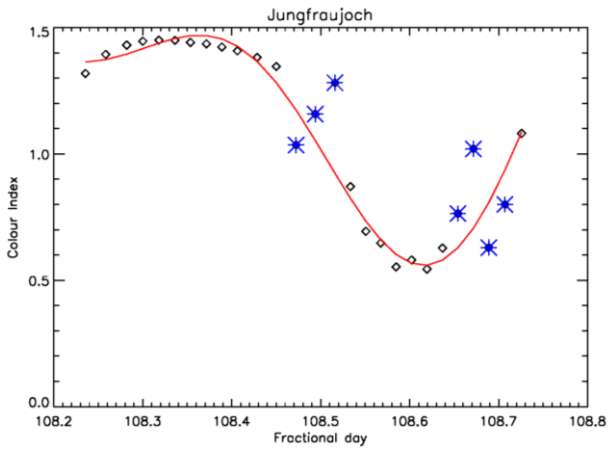

Figure 11: Measured CI (black) together with model fit (red), and outlier detection (blue) for MAXDOAS observations at the Jungfraujoch station.

6 Conclusions

The retrieval algorithms to be implemented in the FRM_4DOAS centralised processing are described in the present ATBD, together with the theoretical principles which are behind each of them.

Trace gas slant column densities are first derived from the calibrated radiance spectra using the QDOAS spectral fitting software. QDOAS output are then be used to retrieve the FRM_4DOAS vertical profile and column data products:

- Lower tropospheric profiles and vertical columns of NO_2 and HCHO are retrieved using the parameterized MAPA MC and OEM-based MMF algorithms selected through a dedicated Round-Robin exercise.
- Total ozone column are derived using the standard AMF-based NDACC approach.
- Stratospheric NO_2 vertical profiles are retrieved using the BIRA OEM-based profiling tool.

The methodologies which are used to derive the error budgets corresponding to these products are also extensively described, as well as their vertical resolution, horizontal representativeness, and sensitivity to the presence of clouds, which are three important parameters for the assessment of the quality of the (MAX)DOAS retrievals.

7 References

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

Clémer, K., M. Van Roozendael, C. Fayt, F. Hendrick, C. Hermans, G. Pinardi, R. Spurr, P. Wang, and M. De Mazière, Multiple wavelength retrieval of tropospheric aerosol optical properties from MAXDOAS measurements in Beijing, Atmos. Meas. Tech, 3, 863-878, 2010.

Barthia, P. K., C. G. Wellemeyer, S. L. Taylor, N. Nath, and A. Gopolan, Solar Backscatter (SBUV) Version 8 profile algorithm, Proceedings of the Quadrennial Ozone Symposium 2004, edited by C. Zerefos, pp. 295-296, Athens, Greece, ISBN, 960-630-103-6, 2004.

Barret, B., M. De Mazière, and P. Demoulin, Retrieval and characterisation of ozone profiles from solar infrared spectra at the Jungfraujoch, J. Geophys. Res., 107, doi:10.1029 / 2001JD001298, 2002.

Bates, D. R., Rayleigh scattering by air, Planetary and Space Science, 32, 785-790, 1984.

Chipperfield, M. P., New version of the TOMCAT/SLIMCAT off-line chemical transport model: Intercomparison of stratospheric tracer experiments, Q. J. R. Meteorol. Soc., 132, 1179-1203, doi:10.1256/qj.05.51, 2006.

Ciddor, P. E., Refractive index of air: New equations for the visible and near infrared, Applied Optics, 35,1566-1572, 1996.

De Mazière, M., A. M. Thompson, M. J. Kurylo, J. Wild, G. Bernhard, T. Blumenstock, J. Hannigan, J.-C. Lambert, T. Leblanc, T. J. McGee, G. Nedoluha, I. Petropavlovskikh, G. Seckmeyer, P. C. Simon, W. Steinbrecht, S. Strahan, and J. T. Sullivan, The Network for the Detection of Atmospheric Composition Change (NDACC): History, status and perspectives, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-402, in review, 2017.

Deutschmann, T., S. Beirle, U. Frieß, M. Grzegorski, C. Kern, L. Kritten, U. Platt, J. Pukite, T. Wagner, B. Werner, and K. Pfeilsticker, The Monte Carlo Atmospheric Radiative Transfer Model McArtim: Introduction and Validation of Jacobians and 3D Features, J. Quant. Spectr. Rad. Transf., doi:10.1016/j.jqsrt.2010.12.009, 2011.

Erle, F., K. Pfeilsticker, and U. Platt, On the influence of tropospheric clouds on zenith scattered light measurements of stratospheric species, Geophys. Res. Lett., 22, 2725-2728, 1995.

Errera, Q., and D. Fonteyn, Four-dimensional variational chemical assimilation of CRISTA stratospheric measurements, J. Geophys. Res., 106 (D11), 12253-12265, 2001.

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

Fiedler M., H. Frank, T. Gomer, M. Hausmann, K. Pfeilsticker, and U. Platt, Groundbased spectroscopic measurements of stratospheric NO_2 and OClO in arctic winter 1989/90, Geophys. Res. Lett., 20, 963-966, 1993.

Fleischmann, O. C., M. Hartmann, J. P., Burrows, and J. Orphal, New ultraviolet absorption cross-sections of BrO at atmospheric temperatures measured by time-windowing Fourier transform spectroscopy, J. Photochem. Photobiol. A, 168, 117–132, 2004.

Friess U., P. S. Monks, J. J. Remedios, A. Rozanov, R. Sinreich, T. Wagner, and U. Platt, MAX-DOAS O_4 measurements: A new technique to derive information on atmospheric aerosols: 2. Modeling studies, J. Geophys. Res. 111, D14203, doi:10.1029/2005JD006618, 2006.

Gielen, C., M. Van Roozendael, F. Hendrick, G. Pinardi, T. Vlemmix, V. De Bock, H. De Backer, C. Fayt, C. Hermans, D. Gillotay, and P. Wang, A simple and versatile cloud-screening method for MAX-DOAS retrievals, Atmos. Meas. Tech., 7, 3509-3527, doi:10.5194/amt-7-3509-2014, 2014.

Goody, R. M. and Y. L. Yung, Atmospheric radiation : theoretical basis, 1989.

Hendrick, F., B. Barret, M. Van Roozendael, H. Boesch, A. Butz, M. De Mazière, F. Goutail, C. Hermans, J.-C. Lambert, K. Pfeilsticker, and J.-P. Pommereau, Retrieval of nitrogen dioxide stratospheric profiles from ground-based zenith-sky UV-visible observations: Validation of the technique through correlative comparisons, Atmos. Chem. Phys., 4, 2091-2106, 2004.

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.

Hendrick, F., P.V. Johnston, K. Kreher, C. Hermans, M. De Mazière, and M. Van Roozendael, One decade trend analysis of stratospheric BrO over Harestua (60°N) and Lauder (44°S) reveals a decline, Geophys. Res. Lett., 35, L14801, doi:10.1029/2008GL034154, 2008.

Hendrick, F., J.-P. Pommereau, F. Goutail, R. D. Evans, D. Ionov, A. Pazmino, E. Kyrö, G. Held, P. Eriksen, V. Dorokhov, M. Gil, and M. Van Roozendael, NDACC/SAOZ UV-visible total ozone measurements: Improved retrieval and comparison with correlative ground-based and satellite observations, Atm. Chem. Phys., 11, 5975-5995, 2011.

Hendrick, F., J.-F. Müller, K. Clémer, M. De Mazière, C. Fayt, C. Gielen, C. Hermans, J. Z. Ma, G. Pinardi, T. Stavrakou, T. Vlemmix, P. Wang, and M. Van Roozendael, Four Years of Ground-based MAX-DOAS Observations of HONO and NO_2 in the Beijing Area, Atm. Chem. Phys., 14, 765-781, 2014.

Hess, M., P. Koepke, and I. Schult, Optical Properties of Aerosols and Clouds: The Software Package OPAC., Bulletin of the American Meteorological Society, 79, 831-844, 1998.

Hönninger, G., C. von Friedeburg, and U. Platt, Multi axis differential optical absorption spectroscopy (MAX-DOAS), Atmos. Chem. Phys., 4, 231–254, doi:10.5194/acp-4-231-2004, 2004.

Huang, T., S. Walters, G. Brasseur, D. Hauglustaine, W. Wu, S. Chabrillat, X. Tie, C. Granier, A. Smith, S. Madronich, and G. Kockarts, Description of SOCRATES-A Chemical Dynamical RadiativeTwo-Dimensional Model, NCAR Technical Note, NCAR/TN-440+EDD, 1998.

Kreher, K., P. V. Johnston, S. W. Wood, and U. Platt, Ground-based measurements of tropospheric and stratospheric BrO at Arrival Heights (78°S), Antarctica, Geophys. Res. Lett., 24, 3021-3024, 1997.

Li, X., T. Brauers, M.Shao, R. M. Garland, T. Wagner, T. Deutschmann, and A. Wahner, MAX-DOAS measurements in southern China: retrieval of aerosol extinctions and validation using ground-based in-situ data, Atmos. Chem. Phys., 10, 2079–2089, doi:10.5194/acp-10-2079-2010, 2010.

Mayer, B., and A. Kylling, Technical note: The libRadtran software package for radiative transfer calculations - description and examples of use, Atmos. Chem. Phys., 5, 1855-1877, 2005.

McKenzie, R., P. V. Johnston, C. T. McElroy, J. B. Kerr, and S. Solomon, Altitude distributions of stratospheric constituents from ground-based measurements at twilight, J. Geophys. Res., 96, 15499-15511, 1991.

Meller, R. and G. K. Moortgat, Temperature dependence of the absorption cross sections of formaldehyde between 223 and 323K in the wavelength range 225–375 nm, J. Geophys. Res., 105, 7089–7101, 2000.

Orphal, J.: A critical review of the absorption cross-sections of O_3 and NO_2 in the 240-790 nm region, J. of Photochem. and Photobiol. A: Chemistry, 157, 185-209, 2003.

Roscoe, H. K., J. A. C. Squires, D. J. Oldham, A. Sarkissian, J.-P. Pommereau, and F. Goutail, Improvements to the accuracy of zenith-sky measurements of total ozone by visible spectrometers, J. Quant. Spectrosc. Radiat. Transfer, 52(5), 639-648, 1994.

Platt, U., K. Pfeilsticker, K., and M. Vollmer, Radiation and Optics in the Atmosphere, p. 1165, doi:10.1007/978-0-387-30420-5 19, 2007.

Platt, U., and J. Stuz, Differential Optical Absorption Spectroscopy (DOAS), Principles and Applications, ISBN 978-3-540-21193-8, Springer, Berlin-Heidelberg, 2008.

Pommereau, J. P., and F. Goutail, O_3 and NO_2 ground-based measurements by visible spectrometry during Arctic winter and spring 1988, Geophys. Res. Lett., 15, 891-894, 1988.

Preston, K. E., R. L. Jones, and H. K. Roscoe, Retrieval of NO_2 vertical profiles from ground-based UV-visible measurements: Method and validation, J. Geophys. Res., 102 (D15), 19,089-19,097, 1997.

Richter, A., M. Eisinger, A. Ladstätter Weißenmayer, F. Wittrock, and J. P. Burrows, DOAS zenith-sky observations: 2.Seasonal variations of BrO over Bremen (53°N) 1994-1995, J. Atmos. Chem., 32, 83-99, 1999.

Rodgers, C. D., Inverse Methods for Atmospheric Sounding, Theory and Practice. World Scientific Publishing, Singapore-NewJersey-London-Hong Kong, 2000.

Roscoe, H. K., M. Van Roozendael, C. Fayt, A. du Piesanie, N. Abuhassan, C. Adams, M. Akrami, A. Cede, J. Chong, K. Clemer, U. Friess, M. Gil Ojeda, F. Goutail, R. Graves, A. Griesfeller, K. Grossmann, G. Hemerijckx, F. Hendrick, J. Herman, C. Hermans , H. Irie, Y. Kanaya, K. Kreher, P. Johnston, R. Leigh, A. Merlaud, G. H. Mount, M. Navarro, H. Oetjen, A. Pazmino, M. Perez-Camacho, E. Peters, G. Pinardi, O. Puentedura, A. Richter, A. Schönhardt, R. Shaiganfar, E. Spinei, K. Strong, H. Takashima, T. Vlemmix, M. Vrekoussis, T. Wagner, F. Wittrock, M. Yela, S.Yilmaz, F. Boersma, J. Hains, M. Kroon, and A. Piters, Intercomparison of slant column measurements of NO_2 and O_4 by MAX-DOAS and zenith-sky UV and visible spectrometers, Atmos. Meas. Tech., 3, 1629-1646, 2010.

Rothman, L. S., I.E. Gordon, R.J. Barber, H. Dothe, R.R. Gamache, A. Goldman, V.I. Perevalov, S.A. Tashkun, and J. Tennyson, HITEMP, the high-temperature molecular spectroscopic database, J. Quant. Spectrosc. Radiat. Transfer, 111, 2139-2150, 2010.

Rozanov, A., V. Rozanov, M. Buchwitz, A. Kokhanovsky, and J. P. Burrows, SCIATRAN 2.0. A new radiative transfer model for geophysical applications in the 175–2400 nm spectral region, Adv. Space Res., 36(5), 1015–1019, 2005.

Sander, S. P., et al., Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies, Evaluation n°15, NASA JPL Publication n° 06-2, 2006.

Sarkissian, A., J.-P. Pommereau, and F. Goutail, Identification of polar stratospheric clouds from the ground by visible spectrometry, Geophys. Res. Lett., 18, 779, 1991.

Sarkissian, A., G. Vaughan, H. K. Roscoe, L. M. Bartlett, F. M. O'Connor, D. G. Drew, P.A. Hughes, and D. M. Moore, Accuracy of measurements of total ozone by a SAOZ ground-based zenith-sky visible spectrometer, J. Geophys. Res., 102(D1), 1379-1390, 1997.

Schofield, R., K. Kreher, B. J. Connor, P. V. Johnston, A. Thomas, D. Shooter, M. P. Chipperfield, C. D. Rodgers, and G. H. Mount, Retrieved tropospheric and stratospheric BrO columns over Lauder, New Zealand, J. Geophys. Res., 109, D14304, doi:10.1029/2003JD004463, 2004.

Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W., and Burrows, J. P.: High spectral resolution ozone absorption cross-sections – Part 2: Temperature dependence, Atmos. Meas. Tech., 7, 625-636, doi:10.5194/amt-7-625-2014, 2014.

Shettle, E. P., Models of aerosols, clouds, and precipitation for atmospheric propagation studies, AGARD Conference Proceedings No. 454: Atmospheric propagation in the UV, visible, IR and mm-region and related system aspects, 1989.

Solomon, S., R. W. Sanders, M. A. Carroll, and A. L. Schmeltekopf, Visible and near-ultraviolet spectroscopy at McMurdo station, Antarctica, 5. Diurnal variation of OClO and BrO, J. Geophys. Res., 94, 11393-11403, 1989.

Spurr, R. J. D., T. P. Kurosu, and K. V. Chance, A linearized discrete ordinate radiative transfer model for atmospheric remote-sensing retrieval, Journal of Quantitative Spectroscopy and Radiative Transfer, 68, 689-735, doi:10.1016/S0022-4073(00)00055-8, 2001.

Spurr, R., VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media, Journal of Quantitative Spectroscopy & Radiative Transfer, pp. 316-342, 2006.

Spurr, R., User's Guide VLIDORT Version 2.6, RT Solutions, Inc., 2013.

Struthers, H., K. Kreher, J. Austin, R. Schofield, G. Bodeker, P. Johnston, H. Shiona, and A. Thomas, Past and future simulations of NO_2 from a coupled chemistry-climate model in comparison with observations, Atmos. Chem. Phys, 4, 2227-2239, 2004.

Thalman, R. and R. Volkamer, Temperature dependent absorption cross-sections of O_2-O_2 collision pairs between 340 and 630 nm and at atmospherically relevant pressure., Phys. Chem. Chem. Phys., 15(37), 15371–81, doi:10.1039/c3cp50968k, 2013.

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 NO_2 absorption cross section from 42000 cm-1 to 10000 cm-1 (238-1000 nm) at 220 K and 294 K, J. Quant. Spectrosc. Radiat. Transfer, 59, 171-184, 1998.

Vandaele, A. C., C. Fayt, F. Hendrick, C. Hermans, F. Humbled, M. Van Roozendael, M. Gil, M. Navarro, O. Puentedura, M. Yela, G. Braathen, K. Stebel, K.Tørnkvist, P. Johnston, K. Kreher, F. Goutail, A. Mieville, J.-P. Pommereau, S. Khaikine, A. Richter, H. Oetjen, F. Wittrock, S. Bugarski, U. Frieß, K. Pfeilsticker, R. Sinreich, T. Wagner, G. Corlett, and R. Leigh, An intercomparison campaign of ground-based UV-Visible measurements of NO_2, BrO, and OClO slant columns. Methods of analysis and results for NO_2, J. Geophys. Res., 110, D08305, doi:10.1029/2004JD005423, 2005.

Van Roozendael, M., M. De Mazière, and P. C. Simon, Ground-based visible measurements at the Jungfraujoch station since 1990, J. Quant. Spectrosc. Radiat. Transfer, 52 (3/4), 231-240, 1994.

Van Roozendael, M., P. Peters, H. K. Roscoe, H. De Backer, A. E. Jones, L. Bartlett, G. Vaughan, F. Goutail, J.-P. Pommereau, E. Kyrö, C. Wahlstrom, G. Braathen, and P. C. Simon, Validation of ground-based visible measurements of total ozone by comparison with Dobson and Brewer spectrophotometers, J. Atm. Chem., 29, 55-83, 1998.

Vlemmix, T., A. J. M.Piters, A. J. C. Berkhout, L. F. L. Gast, P. Wang, and P. F. Levelt, Ability of the MAX-DOAS method to derive profile information for NO_2: can the boundary layer and free troposphere be separated?, Atmos. Meas. Tech., 4, 2659–2684, doi:10.5194/amt-4-2659-2011, 2011.

Vlemmix, T., F. Hendrick, G. Pinardi, I. De Smedt, C. Fayt, C. Hermans, A. Piters, P. Levelt, and M. Van Roozendael, MAX-DOAS observations of aerosols, formaldehyde and nitrogen dioxide in the Beijing area: comparison of two profile retrieval approaches, Atmos. Meas. Tech., 8, 941-963, 2015.

Wagner, T., F. Erie, L. Marquard, C. Otten, K. Pfeilsticker, T. Senne, J. Stutz, and U. Platt, Cloudy sky optical paths as derived from differential optical absorption spectroscopy observations, J. Geophys. Res., 103, 25307, doi:10.1029/98JD01021, 1998.

Wagner, T., C. von Friedeburg, M. Wenig, C. Otten, and U. Platt, UV/vis observations of atmospheric O_4 absorptions using direct moon light and zenith scattered sunlight under clear and cloudy sky conditions, J. Geophys. Res., 107, D204424, doi:10.1029/2001JD001026, 2002.

Wagner T., B. Dix, C. von Friedeburg, U. Friess, S. Sanghavi, R. Sinreich, and U. Platt, MAX-DOAS O_4 measurements: A new technique to derive information on atmospheric aerosols – Principles and information content, J. Geophys. Res. 109, D22205, doi:10.1029/2004JD004904, 2004.

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. Mc Linden, H. Oetjen, E. Palazzi, A. Petritoli, U. Platt, O. Postylyakov, J. Pukite, A. Richter, M. Van Roozendael, A. Rozanov, R. Sinreich, S. Sanghavi, S., and Wittrock, F. : 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., S. Beirle, T. Brauers, T. Deutschmann, U. Frieß, C. Hak, J. D. Halla, K. P. Heue, W. Junkermann, X. Li, U. Platt, and I. Pundt-Gruber, Inversion of tropospheric profiles of aerosol

extinction and HCHO and NO_2 mixing ratios from MAX-DOAS observations in Milano during the summer of 2003 and comparison with independent data sets, Atmos. Meas. Tech., 4, 2685-2715, doi:10.5194/amt-4-2685-2011, 2011.

Wagner, T., A. Apituley, S. Beirle, S. Dörner, U. Frieß, J. Remmers, and R. Shaiganfar, Cloud detection and classification based on MAX-DOAS observations, Atmos. Meas. Tech., 7, 1289-1320, doi:10.5194/amt-7-1289-2014, 2014.

Wagner, T., and S. Beirle, QA_4ECV Technical Note: Estimation of the horizontal sensitivity range from MAX-DOAS O_4 observations, April 2016.

Wagner, T., S. Beirle, J. Remmers, R. Shaiganfar, and Y. Wang, Absolute calibration of the colour index and O_4 absorption derived from Multi AXis (MAX-)DOAS measurements and their application to a standardised cloud classification algorithm, Atmos. Meas. Tech., 9, 4803-4823, doi:10.5194/amt-9-4803-2016, 2016.

Wang, Y., M. Penning de Vries, P. Xie, S. Beirle, S. Dörner, J. Remmers, A. Li, and T. Wagner, Cloud and aerosol classification for 2.5 years of MAX-DOAS observations in Wuxi (China) and comparison to independent data sets, Atmos. Meas. Tech., 8, 5133-5156, doi:10.5194/amt-8-5133-2015, 2015.

Wang, Y., J. Lampel, P. Xie, S. beirle, A. Li, D. Wu, and T. Wagner, Ground-based MAX-DOAS observations of tropospheric aerosols, NO_2, SO_2, and HCHO in Wuxi, China, from 2011 to 2014, 17, 2189–2215, 2017, doi:10.5194/acp-17-2189-2017.

Winterrath, T., T. P. Koruso, A. Richter, and J. P. Burrows, Enhanced O_3 and NO_2 in thunderstorm clouds: convection or production, Geophys. Res. Lett., 26, 1291-1294, doi:10.1029/1999GL900243, 1999.