ECMWF COPERNICUS REPORT

Copernicus Climate Change Service

Algorithm Theoretical Basis Document

GOMOS - AerGOM

Issued by: BIRA / Christine Bingen and Charles Robert

Date: 30/09/2018

Ref: C3S_D312a_Lot5.2.1.9_201809_ATBD_GOMOS_AerGOM_v2

Official reference number service contract: 2016/C3S_312a_Lot5_DLR/SC1

IMPLEMENTED BY ECMWF

Contributors

BIRA-IASB
C. Bingen
C. Robert
F. Vanhellemont
N. Mateshvili

History of modifications

Version	Date	Description of modification	Chapters / Sections
1.0	30/09/2018	Adaptation of the Aerosol_cci ATBD in Version 5.0 into ECMWF template; Description of the new features of and products in the output dataset.	All chapters

Related documents

Reference ID	Document
D1	ESA Climate Change Initiative, aerosol_cci, Algorithm Theoretical Basis Document GOMOS AerGOM, Christine Bingen, Filip Vanhellemont, and Charles Robert, Doc. No.: ATBD0-AerGOM, issue 5.0, 30/03/2017, http://www.esa-aerosol-cci.org
D2	ESA. GOMOS Product Handbook, Issue 3.0, 31 May 2007.
D3	Algorithm Theoretical Basis Document, "AERGOM: Aerosol profile retrieval prototype for GOMOS", Filip Vanhellemont and Nina Mateshvili, ESA Doc. No.: AERGOM-ATBD-2010-07 Iss./Rev.: 2.0, 15/02/2013.

Acronyms

Acronym	Definition
ACE	Atmospheric Chemistry Experiment
AE	Angstrom exponent
AerGOM	GOMOS Retrieval algorithm optimized for aerosols
AOD	Aerosol optical density
APE-THESEO	Airborne Platform for Earth observations – contribution to the Third European Stratospheric Experiment on Ozone
AOT	Aerosol optical thickness
ATBD	Algorithm Theoretical Basis Document
BIRA-IASB	Royal Belgian Institute for Space Aeronomy
CCI	Climate Change Initiative
CLAES	Cryogenic Limb Etalon Array Spectrometer
CO_2	Carbon dioxide
E/B	Extinction-to-backscatter ratio
ECMWF	European Centre for Medium-range Weather Forecasts
ENVISAT	ESA's ENVIronmental SATellite
Eqn.	Equation
ESA	European Space Agengy
FCM	Full covariance matrix
FMI	Finish Meteorological Institute

GOM_EXT	GOMOS data files containing the residual extinctions and ECMWF analysis
GOMOS	Global Ozone Monitoring by Occultation of Stars
GOM_TRA	GOMOS data files containing the transmission spectra products
HNO_3	Nitric acid
H_2O	Water, water vapour
H_2SO_4	Sulfuric acid
IPF	Instrument Processor Facility
IR	Infrared
L1, L2, L3	Level 1, Level2, Level3 data
LM	Levenberg-Marquardt
LUT	Look-up table
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MAESTRO	Measurements of Aerosol Extinction in the Stratosphere and Troposphere Retrieved by Occultation
Na	Natrium
NAT	Nitric acid trihydrate
NO_2	Nitrogen dioxide
NO_3	Nitrate
O_2	Dioxygen
O_3	Ozone
OClO	Chlorine dioxide
POAM	Polar Ozone and Aerosol Measurement
OSIRIS	Optical, Spectroscopic, and Infrared Remote Imaging System
PSC	Polar stratospheric cloud
PSD	Particle size density
PSF	Point-spread function
SAGE	Stratospheric Aerosol and Gas Experiment
SNR	Signal-to-noise ratio
SPA	GOMOS spectrometer A
SPB, SPB1, SPB2	GOMOS spectrometer B, B1, B2
SSA	Single-scattering albedo
STS	Supercooled ternary solution
SZA	Solar zenith angle
TH	Tropopause height
UTLS	Upper troposphere – Lower stratosphere
UV	Ultraviolet
Vis	Visible
WGS84	World Geodetic System 1984
WMO	World Meteorological Organization

General definitions

Table of Contents

Scope of the document

This document describes the theoretical basis for the development of gridded time series of a suite of products describing the radiative and microphysical properties of stratospheric aerosols, derived from GOMOS measurements. In this work initiated in the framework of the Aerosol_cci project, we make use of **AerGOM**, a retrieval algorithm developed for GOMOS by the BIRA-IASB and the FMI as an alternative to the operational GOMOS algorithm (Instrument Processor Facility, IPF) to optimize the aerosol retrieval in the stratosphere and the upper troposphere.

AerGOM and its performance have been extensively described in a twin paper [Vanhellemont et al., 2016, Robert et al., 2016], and a report [D3] written for the ESA-AERGOM project. The development of aerosol extinction time series for CCI has been described in two papers ([Popp et al., 2016; Bingen et al., 2017). The present ATBD aims to provide an overview of the algorithm and its use in the development of aerosol products in the framework of the Aerosol_cci project and beyond, with detailed references to the above-mentioned documents, and with an overview of the issues that are important for the the quality of the final products and their use in the climate modelling applications. It will not be a comprehensive compilation of all existing literature.

Upper tropospheric and stratospheric aerosols (either liquid or solid) are intensively studied for a number of reasons that can be roughly summarized as follows:

(1) they have an impact on the Earth radiative balance due to their optical properties;
(2) they play a crucial role in heterogeneous chemistry;
(3) they provide information on the emission of precursor species from which they originate.

From a practical point of view, the knowledge of the stratospheric aerosols extinction can provide a correction for satellite retrievals of tropospheric AOT, correction which can become significant during periods of high stratospheric aerosol loading following volcanic eruptions.

The AerGOM algorithm answers to this need of stratospheric aerosol information by providing the retrieved quantities:

- Particle extinction altitude profiles.
- Particle size distributions at the same altitudes, on a pre-defined logarithmic size grid. This is an intermediate product which is not distributed at this stage.
- Microphysical and radiative parameters derived from the particle size distribution: total particle number density, mean radius, distribution variance, surface area density, volume density, effective radius, asymmetry parameter and single scattering albedo.

Uncertainties are derived for each quantity.

Executive summary

The GOMOS instrument, which flew onboard ENVISAT and measured the atmospheric composition from 2002 to 2012, is based on stellar occultation. The use of the occultation techniques makes possible the retrieval of vertical profiles of abundance of the main atmospheric consituents, including concentration profiles of trace gases (ozone, NO_2, NO_3, etc.) and stratospheric aerosols extinction.

In the present work, these vertical profiles are retrieved using AerGOM, a retrieval algorithm developed for GOMOS to optimize the aerosol retrieval in the stratosphere and the upper troposphere. The retrieval of aerosols in the stratosphere and the upper tropospher are specifically the focus of this work, and we describe the methodology used to implement the aerosol extinction retrieval following two steps:

1. the spectral inversion providing the slant column densities of each gas species and the slant aerosol optical thickness at all wavelengths covered by the GOMOS spectral range.
2. The spatial inversion providing vertical profiles for each retrieved quantities from the integrated slant densities or optical thickness.

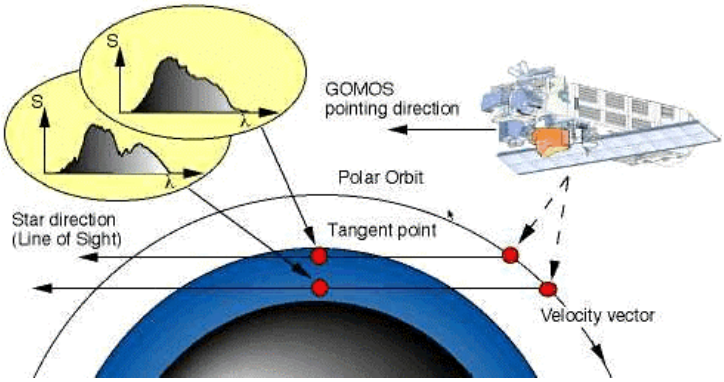
In a next step, the aerosol extinction spectra can be used to retrieve aerosol particle size distributions (PSD). This step is called “radial inversion”. From the knowledge of the particle size distribution, it is possible to derived in turn many microphysical and radiative properties of importance for the characterization and modelling of aerosols, such as their effective radius, surface area density, volume density, but also to estimate the Angström exponent, asymmetry factor and singls scattering density.
We model here the PSD using a sectional particle distribution function, what provides a versatile expression able to describe multimodal aerosol size distributions. A regularization scheme is used limit the number of degree of freedom and to constrain the inverse problem in an appropriate way. Making use of the assumption that aerosol particles are quasi-spherical, we the Mie theory to derived the radiative quantities of interest, while the microphysical properties are calculated from the momentums of the distribution function.

In order to provide adequate climate data records for modelling applications, Level 3 gridded products are producted from the AerGOM extinction profiles, with a high time resolution (5 days in the current version of the CCI-GOMOS products), and a three-dimensional spatial resolution (5° latitude x 60° longitude x 1 kilometer altitude in the current version).
The same kind of binning can be applied on all vertical profiles of microphysical and radiative aerosols properties. However, it was found that the accuracy of the retrieval in significantly improved if the radial inversion is performed on the Level 3 extinction data. Therefore, this approach is used from the present version of the CCI-GOMOS dataset (CCI version CCI_4.00, making use of the AerGOM version 3.00).

1. Instruments

The Envisat satellite was launched on 1 March 2002, and has been fully operational until the loss of communication on April 8th, 2012. On board, it carried a range of instruments designed to measure data with specific application in a wide diversity of Earth science studies. One of these instruments, GOMOS (Global Ozone Monitoring by Occultation of Stars; see [D1-D3], Bertaux et al., 1991, 2000, Kyrölä et al., 2004, Kyrölä and Blanot, 2007, 2012] is a UV/Visible/near-IR spectrometer that works in occultation mode: while orbiting the Earth, the instrument measures the transmission of light from stars that are setting below the Earth's horizon, as depicted in Figure 1. Since the starlight has to pass through the Earth's atmosphere, it is partly scattered or absorbed by atmospheric gases and particles. The measurements can therefore be used to retrieve gas concentration and aerosol extinction profiles. Using a scanning mirror and a star tracker, GOMOS continuously observes selected stars; during one orbit, several different occultations are measured. In this way, several hundreds of occultations with good global coverage can be measured per day. Measurements are taken both on the dark and Sun-illuminated side of the Earth, although in the latter case scattered sunlight represents an extra source of error for retrievals. During one orbit about 30 to 50 ocultations can be measured, demonstrating the potential of using stars as the light source: the measurement rate is much higher in comparison with the usual solar occultation technique, that is limited to the two occultations for sunrise and sunset. The price to pay is the smaller signal-to-noise ratio, since the Sun is a much brighter light source.

Figure 1 - The GOMOS principle of measurement by occultation of star. The spectrum of a star is first measured outside of the atmosphere, then through the atmosphere; the ratio of the two spectra is the atmospheric transmission in this spherical geometry, at altitude z.



While GOMOS was originally conceived as an instrument designed to measure highly accurate ozone profiles, a few other species can also be derived. Typically, the UV/Vis wavelength range combined with the sensitivity of the GOMOS spectrometers allows the retrieval of ozone, NO_2, air, NO_3, O_2 and aerosol extinction profiles. While ozone can be retrieved up to 100 km of altitude, the other species are usually only detectable from the upper troposphere to about 50 km. The actual altitude range differs for different species and depends on the abundance of the species and on the amount of light available; for aerosol/clouds in the upper troposphere/lower stratosphere (UTLS), profiles can be typically obtained between 10 and 30 km.

The spectrum of the starlight is measured by four spectrometers operating in a wavelength range from 250 to 950 nm. Additionally, GOMOS is equipped with two fast photometers whose measurements are used to correct for star scintillation and to retrieve high-resolution temperature profiles. More specifically, the GOMOS instrument consists of:

- Two spectrometers A1 and A2 (SPA1 and SPA2) covering the UV- Visible wavelength range (248-690 nm) with a spectral pixel width of 0.31 nm and a spectral resolution of 0.8 nm, with the purpose of measuring optical absorption by O_3, NO_2, NO_3, aerosol extinction and neutral density. A few other trace gases are also detectable (OClO, Na).
- A spectrometer B1 (SPB1) covering the longer wavelength range (755- 774 nm) with a spectral pixel width of 0.045 nm and a spectral resolution of 0.13 nm, with the purpose of measuring O_2 absorption.
- A spectrometer B2 (SPB2) covering the near-IR wavelength range (926-954 nm) with a spectral pixel width of 0.052 nm and a spectral resolution of 0.13 nm. It was included to measure water vapour absorption bands.
- Two fast photometers, sampling the blue (473-527 nm) and the red (646-698 nm) spectral domain at a sampling frequency of 1 kHz. Their main purpose is the removal of star scintillation in the spectrometer measurements.

GOMOS was primarily intended to deliver accurate altitude profiles of trace gas concentrations. Atmospheric particle populations pose a more challenging problem since the spectral shape is unknown, with the direct consequence that it is a priori impossible to find out which kind of particles are in the line of sight of the instrument. This is the reason why GOMOS delivers primarily one global common product, perhaps somewhat misleadingly dubbed “total aerosol extinction”. It should be understood that this data product embraces stratospheric aerosols, tropical subvisual cirrus clouds and Polar Stratospheric Clouds (PSCs) - or indeed even any unknown extinction phenomenon with a smooth spectral dependence. An analysis of the spectral dependence of the aerosol extinction coupled with altitude and temperature information is performed to infer the aerosol composition (sulfate aerosol, PSC, meteoretic dust; clouds should be added in the future) and to retrieve size information using the best possible assumption on the aerosol refraction properties, but the detection of the aerosol type is in some cases very uncertain. These aspects are discussed more into detail in Section 3.1.8.

2. Input and auxiliary data

The input data required for the retrieval by AerGOM are:

- GOM_EXT files (which contain the ECMWF analysis)
- GOM_TRA (optional: only needed if one wants to calculate the full covariance matrix)
- Database of gas absorption cross-sections
- Reference profiles (gas, aerosol extinction) for the first guess of iterative minimalization.

The GOM_EXT and GOM_TRA datasets are GOMOS files processed by ESA using the operational algorithm. We are using here the dataset from the last GOMOS reprocessing in the version 6.01.

Concerning the gas absorption cross-sections, an important work has been carried out during the last years to update the cross-section database with respect of the GOMOS cross-section database used in IPF. More recent, comprehensive and accurate spectra are available nowadays, improving the gas retrieval from GOMOS. The detail of the the gas absorption cross-section spectra used in AerGOM can be found in [Bingen et al., 2017].

Furthermore, for the radial inversion, look-up tables (LUT) of the extinction efficiency Q_{ext} as a function of temperature and composition are needed. These aspects are discussed in Sections 3.1.3 and 3.2.6.

Finally, for the purpose of the stratospheric AOD calculation in the framework of the Aerosol_cci, the tropopause height is determined using full ECMWF reanalysis data (containing temperature and potential vorticity).

3. Algorithms

3.1 General approach

3.1.1 The law of Beer-Lambert

Consider a ray of light that follows an optical path through a medium that absorbs and scatters light. The amount of absorption and scattering depends on the position inside the medium since the concentration of the absorbing/scattering species varies spatially. If we parameterize the optical path by a parameter s (a path coordinate), then the optical thickness along the path is given by:

(1) $$\tau_{tot}(\lambda) = \int_s \beta_{tot}(\mathbf{r}(s), \lambda) \mathrm{d}s$$

with β_{tot} the optical extinction coefficient, that depends on the spatial position r and the wavelength λ. The indication 'tot' was used to make clear that total optical extinction results from the separate contributions of the different species:

(2) $$\beta_{tot} = \beta_{air} + \beta_{O_3} + \beta_{NO_2} + \beta_{NO_3} + \ldots + \beta_{aero}$$

refering to the extinction by the neutral air, trace gases (ozone, NO_2, NO_3, …) and the remaining contribution from all kinds of particulate matters grouped under the name of "aerosols". The intensity of the light after travelling through the medium depends exponentially on the optical thickness, a behaviour that is described by the law of Beer-Lambert:

(3) $$I(\lambda) = I_0(\lambda) \exp\left(-\tau_{tot}(\lambda)\right)$$

Here, $I_0(\lambda)$ represents the unattenuated spectrum of the light source.

3.1.2 Optical extinction by gases

Optical extinction is the process where light from an incident optical beam is removed by gases and particles, through the combined process of absorption and scattering. Both processes are mathematically described by the use of a cross-section with the dimension of area (usually in units of cm^2).

3.1.2.1 Absorption by gases

Absorption cross-section spectra of gases are usually measured in the laboratory, and are often temperature dependent. For atmospheric applications, the cross-section spectra are typically measured at a number of temperatures that are representative of atmospheric conditions. The optical absorption coefficient is expressed as the product of the absorption cross-section and the gas number density N:

(4) $$\beta(\lambda, \mathbf{r}) = C_{abs}(\lambda, T(\mathbf{r})) \cdot N(\mathbf{r})$$

with T the (spatially varying) temperature.

Gas absorption cross-sections are provided by the operational GOMOS processor [Kyrölä and Blanot, 2007], and this cross-section database was used in the first version of AerGOM. However, it appeared that some features of this database, which was set up before ENVISAT's launch, were outdated. Therefore, the reference absorption cross-sections were revised for the gases retrieved by AerGOM, as well as their dependence in temperature, using the most recent published works. The general purpose of this revision was:

- To fill, whenever possible, gaps in the spectral and temperature dependence so that the whole GOMOS spectral range and the whole temperature range encountered can be covered.
- To strive for a maximal data consistency by avoiding combination of different data sources if not necessary.
- To give the preference to high-resolution absorption cross-section spectra.

The most rigourous description of the signal received by the instrument is obtained by using the high-resolution cross-section in Eqn. (4) to derive the optical thickness using Eqns. (1) and (2). The intensity of the light entered the instrument can then be calculated from the Beer-Lambert's law (3) and by convolution of the high resolution intensity by the point-spread function (PSF) of the instrument. However, this approach implying a convolution for each measurement in very demanding in computing resources. Therefore, we used so far the common approximation consisting in applying the convolution by the PSF on the absorption cross-sections, and not in the high-resolution input intensity. Doing so, only one convolution is required once and for all on the cross-sections before starting the retrieval, instead of several convolutions per measurement at each tangent altitude during the nonlinear optimization of the spectral retrieval. However, an option is available in the AerGOM code to perform the most rigourous approach if required.
Table 1, Table 2, and Table 3 give an overview of the revised absorption cross-section database for O_3, NO_2 and NO_3 respectively.
A detailed discussion of the choice of reference data is given in [Bingen et al., 2017], Appendix A.

Table 1 - Overview of the ozone cross-sections used in AerGOM after revision of the GOMOS cross-section database. T refers to the temperature, in K.

Species	Spectral range	Source/Data availability	Methodology
O_3	240-780 nm	[Serdyuchenko et al., 2014] Available at T=193, 203, ..., 293 K Spectral resolution: 0.02-0.24 nm	Linear interpolation in temperature

Table 2 - Overview of the NO_2 cross-sections used in AerGOM after revision of the GOMOS cross-section database. T refers to the temperature, in K.

Species	Spectral range	Source/Data availability	Methodology
NO_2	240-695 nm	[Vandaele et al., 2003] Simulated temperature dependence over 217-298.5 K based on [Vandaele et al., 2002] Spectral resolution = 0.01-0.11 nm	Use of the simulated temperature dependence.
	750-757. 8 nm	[Vandaele et al., 2003] Available at T=220, 240, 294 K Spectral resolution = .002-.003 nm	At each wavelength, determination of a linear temperature dependence by least square fit.
	757. 8-780 nm	[Vandaele et al., 1998] Available at 220 and 294 K Spectral resolution = 0.11-0.12 nm	At each wavelength, determination of a linear temperature dependence by a linear interpolation.
	694-780 nm	[Orphal et al., 2003] Available at T=294 K Spectral resolution = 0.03-0.04 nm	Use of data at T=294 K.

Table 3 - Overview of the NO_3 cross-sections used in AerGOM after revision of the GOMOS cross-section database. T refers to the temperature, in K.

Species	Spectral range	Source/Data availability	Methodology
NO_3	240-400 nm	[Sander, 1986]	Use of the value at 400 nm at the corresponding temperature.
	400-440 nm	[Sander, 1986] Available at 230 and 298 K Spectral sampling: 1 nm	Linear interpolation in temperature between 230 and 298 K. Below 230 K, use the data at this temperature.
	440-476 nm	[Yokelston et al., 1994] normalized to [Sander, 1986] at 298K Available at 220, 240, 260, 280 and 298 K Spectral resolution ~ 0.1 nm	Use of recommended temperature dependence when available. If not available (below T=220 K), use of the data at T=220 K.

	476-694 nm, except range 650-675 nm	[Orphal et al., 2003] Available at T=294 K Spectral resolution = 0.01-0.03 nm [Yokelston et al., 1994] Available at 220, 230, 260 and 280 K Spectral resolution ~ 0.1 nm	Use of [Orphal et al., 2003] available at 294 K with temperature dependence from [Yokelston et al., 1994]
	650-675 nm	[Orphal et al., 2003] Model of temperature dependence around the peak at 662 nm Spectral resolution ~ 0.03 nm	Use of the recommended temperature dependence.
	694-780 nm	[Orphal et al., 2003] Available at T=294 K Spectral resolution = 0.03-0.04 nm	Use of data at T=294 K.

3.1.2.2 Rayleigh scattering by the neutral density (air)
In a similar way, the extinction due to optical scattering by air is given by:

(5) $$\beta(\lambda,\mathbf{r}) = C_{sca}(\lambda)\cdot N(\mathbf{r})$$

A number of theoretical laws have been published in the past for the scattering cross-section. The following expression seems to be the most accurate one, and was taken from [Bodhaine et al., 1999]:

(6) $$C_{sca} = \frac{24\pi^3}{\lambda^4 N_{stp}^2}\left(\frac{n_{stp}(\lambda)^2-1}{n_{stp}(\lambda)^2+2}\right)^2\left(\frac{6+3\rho}{6-7\rho}\right)$$

With $n_{stp}(\lambda)$ the real part of the wavelength dependent air refractive index at standard temperature and pressure ($T_{stp}=288.15$ K, $P_{stp}=1013.25$ mb), $N_{stp}=2.546899\times10^{19}$ molecules cm^{-3} the air number density and ρ the depolarisation ratio that takes into account molecular anisotropy. The last term between brackets is known as the depolarisation term or the King factor:

(7) $$F_{air} = \frac{6+3\rho}{6-7\rho}$$

For air, it is commonly assumed to have a value of 1.06 [Lenoble, 1993]. The GOMOS Level II processor [Kyrölä and Blanot, 2007] also assumes this value, together with a slightly modified form of Eqn. (6). However, F_{air} depends on the wavelength and the actual composition of air, and this should be taken into account. A good overview of this subject was given by [Bodhaine et al., 1999]. First, we need the partial depolarization of nitrogen and oxygen as given by [Bates, 1984]:

(8) $$F(N_2,\lambda) = 1.034 + 3.17 \times 10^{-4} \lambda^{-2}$$

(9) $$F(O_2,\lambda) = 1.096 + 1.385 \times 10^{-3} \lambda^{-2} + 1.448 \times 10^{-4} \lambda^{-4}$$

Furthermore, Bates (1984) suggested to take F(Ar)=1, F(CO_2)=1.15 and to ignore other air constituents. Finally, the King factor for air can be calculated as a function of wavelength and CO_2 concentration as:

(10) $$F(air,\lambda) = \frac{C_{N_2} F(N_2,\lambda) + C_{O_2} F(O_2,\lambda) + C_{Ar} F(Ar) + C_{CO_2} F(CO_2)}{C_{N_2} + C_{O_2} + C_{Ar} + C_{CO2}}$$

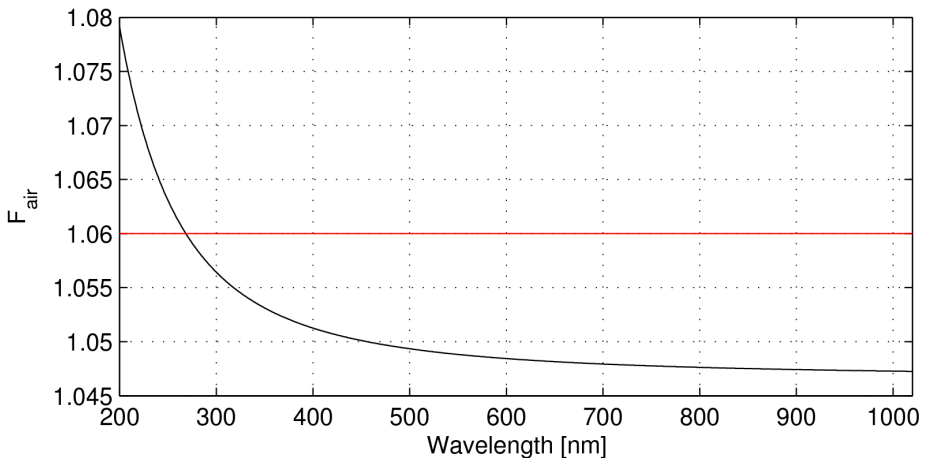
with the concentrations expressed in parts per volume by percent (e.g. use 0.036 for 360 ppm of CO_2). The concentrations are summarized in Table 4.

Table 4 - Volume mixing ratios of the major atmospheric gases (parts per volume by percent)

C_{N_2}	C_{O_2}	C_{Ar}	C_{CO_2}
78.084	20.946	0.934	0.036

Figure 2 shows the King factor calculated from the volume mixing ratios in Table 4. For example: F_{air} takes the value of 1.063 at λ =250 nm and 1.047 at λ =1 μm. The commonly used value (1.06) leads to an error in the Rayleigh cross-section of respectively 0.3 % and 1.2 %. *These are significant corrections if we want to retrieve relatively low aerosol extinction coefficients*.

Figure 2 - The King factor F_{air} as calculated by [Bodhaine et al., 1999]. The commonly used value of 1.06 is also indicated.



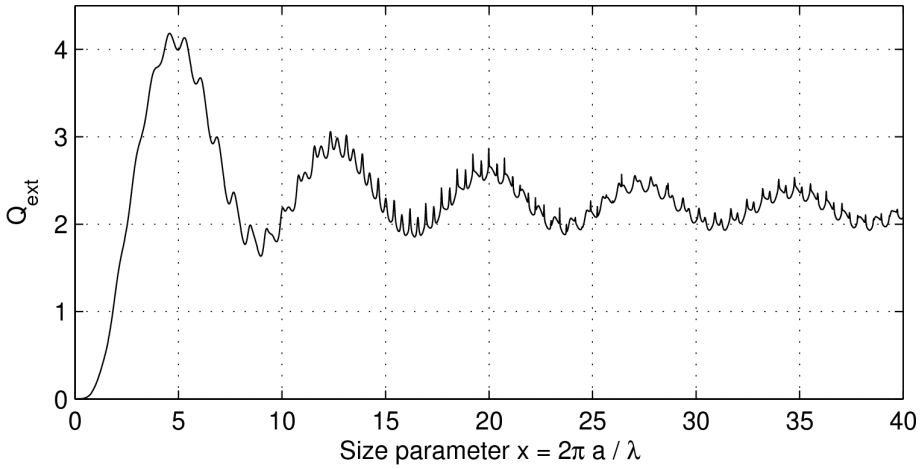
3.1.3 Optical extinction by atmospheric particles

3.1.3.1 Mie theory

Optical absorption and scattering by particles is a very complex process, determined mainly by the particle shape and size, the wavelength λ of the incoming light, the refractive index of the medium (n_m) and the particle (n_p). In the case of spherical particles, an exact theoretical solution was found by Gustav Mie. Excellent treatises can be found in [Bohren and Huffman, 1993] and in [van der Hulst, 1981]. Basically, the Maxwell equations are solved inside and outside the particle, and energy conservation is introduced by applying boundary conditions on the surface of the sphere. The theory is expressed in terms of the size parameter $x = ka = (2\pi a) / \lambda$ (*k* is the wavenumber, *a* the particle radius) and the relative refraction index $m = n_p / n_m$.

Finally, the cross-sections for absorption and scattering (of which the sum equals the extinction cross-section) can be calculated as a series expansion of special analytic functions. The obtained formulas do not give many direct physical insights, but excellent computer codes are available to calculate accurate numerical results, such as the FORTRAN routine BHMIE, that can be found in [Bohren and Huffman, 1993]. An example of the output of this code is given in Figure 3.

Figure 3 - Extinction efficiency for typical spherical sulphate aerosol particles (refractive index m=1.43) as function of the so-called size parameter x=2πa/λ. The refractive index of the medium (air) was assumed to be equal to 1.



3.1.3.2 Optical extinction by a particle population

While the optical extinction of a large number of gas molecules can be adequately described with a wavelength-dependent cross-section and the gas number concentration, this is not possible for particle populations. Two ozone molecules are identical, two aerosol droplets are not: they generally differ in size, chemical composition, state and even morphology. Even if we assume the idealised case of a population of spherical particles with identical chemical composition, an equation such as Eqn. (4) is not sufficient, and we need to use the cumulative aspects of all particles distributed in the particle size domain instead:

(11) $$\beta_{aero}(\lambda,\mathbf{r}) = \int_{a=0}^{+\infty} C_{ext}(\lambda,a) f_a(\mathbf{r},a)\mathrm{d}a$$

(12) $$= \int_{a=0}^{+\infty} \pi a^2 Q_{ext}(\lambda,a) f_a(\mathbf{r},a)\mathrm{d}a^1$$

with C_{ext} the *particle extinction cross-section* (including both scattering and absorption) and Q_{ext} the *extinction efficiency*, the latter indicating the ratio of the actual cross-section to the geometric cross-section of the particle. Both C_{ext} and Q_{ext} depend on wavelength λ and particle radius *a*. We denote the *particle size distribution* with $f_a(a)$, and it can also be written formally as:

(13) $$f_a(\mathbf{r},a) = \frac{\mathrm{d}N(\mathbf{r},a)}{\mathrm{d}a}$$

Expressed in this way, the meaning is clear: $f_a(a)$ is the fraction of particles per volume at atmospheric position **r** with a particle radius in the interval $[a,a+\mathrm{d}a]$. Typically, its units are $cm^{-3}\mu m^{-1}$.

3.1.3.3 Aerosol extinction modelisation in practical situations

Before the actual inversion of occultation measurements, it is difficult to know which kind of particles we are dealing with. Are they sulfate aerosols, PSC particles, or is it cirrus ice crystals that are causing the observed spectra? Since we need the chemical composition *a priori* to model the aerosol extinction (Mie theory depends on the refractive index of the particle material), it is not advisable to use the Mie theory directly in the retrieval process. It is usually preferred to use a simple spectral law for the aerosol extinction with a small number of parameters (that are to be fitted). For example, Angström found that graphs of the logarithm of measured aerosol extinction versus the logarithm of wavelength often showed (quasi-) linear behaviour:

(14) $$\log(\beta_{aero}(\lambda,\mathbf{r})) = c_0(\mathbf{r}) - c_1(\mathbf{r})\log(\lambda)$$

with fit coefficients c_0 and c_1 of course depending on the spatial location of the particles. Transformed, we recognize the so-called Angström law, which is a simple power law:

(15) $$\beta_{aero}(\lambda,\mathbf{r}) = \exp(c_0(\mathbf{r}))\lambda^{-c_1(\mathbf{r})}$$

Although this law is often used, it is not versatile enough to describe all particle population types. For example, researchers that prefer to use this law are often forced to make c_0 and c_1 wavelength dependent, an approach that is rather arbitrary.

In the current operational GOMOS Level 2 algorithm, a polynomial of wavelength (or any function of it) is assumed. In its simplest form:

(16) $$\beta_{aero}(\lambda,\mathbf{r}) = c_0(\mathbf{r}) + c_1(\mathbf{r})(\lambda - \lambda_{ref}) + c_2(\mathbf{r})(\lambda - \lambda_{ref})^2$$

with λ_{ref} some reference wavelength. The reason why such a spectral law is used is its versatility: a quadratic polynomial can assume a wide range of shapes, from a small-particle spectrum ($\beta \sim \lambda^{-4}$) through submicron-sized particles (spectra peaking in the visible wavelength range) to large particle spectra (β = constant).

In the past, the retrieval algorithms of other occultation instruments such as SAGE III [Thomason et al., 2001] and POAM III 0 were equipped with similar aerosol spectral laws, however often expressed as function of the natural logarithm of wavelength:

(17) $$\beta_{aero}(\lambda,\mathbf{r}) = c_0(\mathbf{r}) + c_1(\mathbf{r})\log(\lambda) + c_2(\mathbf{r})(\log(\lambda))^2$$

Spectral functions of inverse wavelength are also sometimes used. For example, a quadratic polynomial of inverse wavelength:

(18) $$\beta_{aero}(\lambda,\mathbf{r}) = c_0(\mathbf{r}) + c_1(\mathbf{r})\left(\frac{1}{\lambda}\right) + c_2(\mathbf{r})\left(\frac{1}{\lambda}\right)^2$$

The argument for the use of this last formula can probably be found in the physical theory for optical absorption/scattering by particles: theoretical cross-sections depend primarily on the size parameter $x = ka = 2\pi a / \lambda$ with *a* the particle radius. In this way, Eqn. (18) can be viewed as a truncated series expansion in the wave number *k*.

3.1.3.4 Aerosol spectral law: interpolation form

When inspecting the quadratic polynomial of Eqn. (16), it is clear that only coefficient c_0 has a physical meaning: it is the extinction coefficient at the reference wavelength λ_{ref}. There are two main reasons why the use of this formalism is far from optimal: (1) the three coefficients have a different unit and magnitude, giving rise to scaling problems during numerical inversion, and (2) it is not clear how to apply smoothing constraints (for example, is an altitude profile smoothing on c_2 in Eqn. (16) meaningful?). Difficulties are avoided when we restate the problem as an interpolation formula between a number of discrete extinction coefficients $\beta_{aero}(\lambda_i)$ that are to be retrieved. For the above quadratic spectral law this gives for example:

(19) $$\beta_{aero}(\lambda,\mathbf{r}) = \sum_{i=1}^{3} q_i(\lambda)\beta_{aero}\left(\lambda_i,\mathbf{r}\right)$$

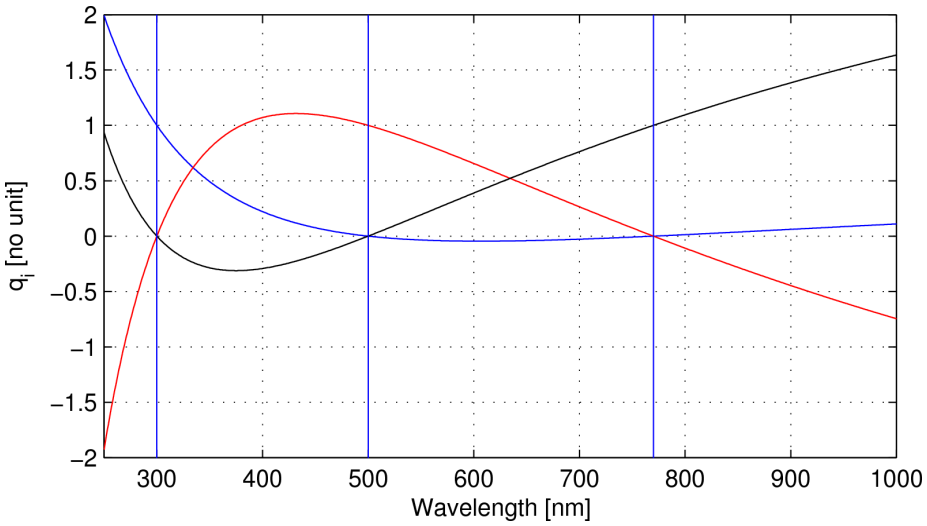
(20) with $$q_i(\lambda) = \frac{(\lambda-\lambda_j)(\lambda-\lambda_k)}{(\lambda_i-\lambda_j)(\lambda_i-\lambda_k)}$$

with λ_i, λ_j and λ_k three different wavelengths that have to be specified in advance. Of course this formalism can be extended. If we want to implement a quadratic polynomial of *inverse* wavelength, the spectral functions above need to be replaced by:

(21) $$q_i(\lambda) = \frac{(\lambda^{-1}-\lambda_j^{-1})(\lambda^{-1}-\lambda_k^{-1})}{(\lambda_i^{-1}-\lambda_j^{-1})(\lambda_i^{-1}-\lambda_k^{-1})}$$

These functions are shown in Figure 4.

Figure 4 - Aerosol spectral functions for a quadratic polynomial of inverse wavelength. The three wavelength parameters of the model are λ_1=300 nm, λ_2=500 nm, and λ_3=770 nm, and are indicated with vertical lines.



In general, a *m*-th degree polynomial of a function of wavelength $f(\lambda)$ can be written as:

(22) $$\beta_{aero}(\lambda,\mathbf{r}) = \sum_{i=1}^{m} q_i(\lambda)\beta_{aero}(\lambda_i,\mathbf{r})$$

(23) with $$q_i(\lambda) = \prod_{j\neq i}^{m} \frac{f(\lambda)-f(\lambda_j)}{f(\lambda_i)-f(\lambda_j)}$$

This formulation, which is a generalization of the traditional Lagrange polynomials, should allow enough flexibility to model realistic particle extinction spectra.

3.1.3.5 Aerosol spectral model: choice based on Mie theory

The actual choice of aerosol spectral law should be based on its ability to model realistic spectra for particle populations that are found in the atmosphere. We therefore simulated extinction spectra from actual particle size data derived from measurements that were performed by satellite instruments such as SAGE II, CLAES and POAM, field campaign results such as APE-THESEO[1] and many in situ and lidar instruments. Measurements of the following particle types were considered:

- Stratospheric sulphuric acid droplets
- Polar Stratospheric Clouds: NAT (Nitric Acid Trihydrate), STS (Supercooled Ternary Solution) and water ice clouds
- Cirrus and subvisual cirrus clouds
- Polar Mesospheric Clouds

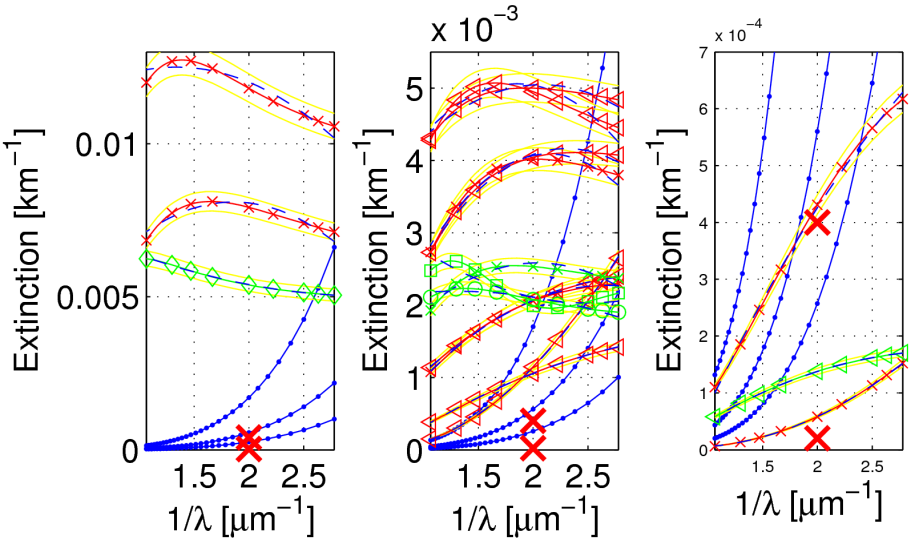
Starting from published values of microphysical parameters (typically lognormal parameters for total number density, mode radius and distribution width), we simulated extinction spectra with a Mie code (assuming spherical particles). The obtained spectra were fitted with a range of candidate spectral laws:

- Extinction was fitted with second and third order polynomials as function of λ, $1/\lambda$, $\log(\lambda)$
- The logarithm of extinction was fitted with second and third order polynomials as function of λ

After comparison of the fit quality, it was clear that the second order polynomial of inverse wavelength is a good versatile model for aerosol extinction spectra for the bulk of GOMOS measurements. Nevertheless, the AerGOM code is equipped with a number of other spectral laws to choose from, in order to check the associated retrieval quality. Examples of simulated spectra and fit are shown in Figure 5.

[1] Airborne Platform for Earth observation - contribution to the Third European Stratospheric Experiment on Ozone; see Stefanutti et al. [2004].

Figure 5 - Stratospheric sulphate aerosol extinction spectra for different aerosol size distributions. Red curves represent remote sensing measurements, in-situ measurements are shown in green. A 4% measurement uncertainty has been superimposed in yellow. Also shown are the fit with the second order polynomial of inverse wavelength (dashed blue) and Rayleigh scattering extinction at 18, 25 and 30 km altitude (dotted blue). Red crosses represent typical GOMOS aerosol extinction values at λ = 500 nm at 20 and 30 km.



3.1.4 A word on Mie calculations

As already mentioned before, in this work we are using the Mie theory for optical scattering and absorption by *spherical* particles. More specifically, the extinction efficiency in Eqn. (12) is calculated with a Mie code. This choice delivers exact results for spherical particles (stratospheric sulfate aerosols, STS). In the case of solid particles - whether they are crystalline (cirrus, NAT) or amorphous (NAT) - the results will not be exact and should be considered as representative for an equivalent spherical particle. It has to be noted that recent studies show that ice particles found in cirrus clouds in the tropical tropopause layer are predominantly small and quasi-spherical [Woods et al, 2017], validating the assumption made in our calculation of the extinction efficiency.

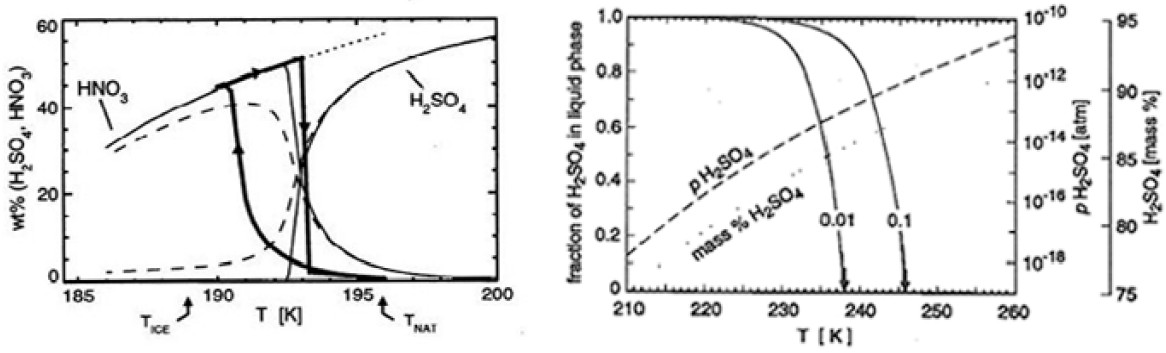
An important aspect in the evaluation of the extinction efficiency Q_{ext} is the particle refractive index. Temperature-induced changes in composition can lead to refractive index differences that in turn significantly alter optical extinction. Furthermore refractive indices in general depend on wavelength.

For every individual particle type it is therefore advisable to calculate the wavelength dependent refractive index at the considered temperature. This temperature is obtained from the ECMWF analysis profiles that are present in the GOM_EXT files.

For pure water ice this poses no problem: the complex refractive index can be interpolated easily from tabulated data that were published by [Warren, 1984]. Other particle types that are to be expected consist of binary and ternary solutions of sulfuric or nitric acid, of which the exact

composition depends on the temperature, as is shown on the plots (Figure 6) that were taken from [Meilinger et al., 1995] and [Carslaw et al., 1997]. It is these figures that will serve as the source for our estimates. From these weight percentages, the refractive index can be calculated with a code published by [Krieger et al., 2000] that is based on a generalized Lorentz-Lorenz equation for the refractive index. Finally, other aerosol types are expected such as meteoric smoke particles, which are likely to be the main aerosol type above 35 km altitude, and soot particles. So far, we do not consider this type of aerosol composition in the retrieval.

Figure 6 - Temperature dependence of the weight percentage of H_2SO_4 and HNO_3 in liquid aerosols. Left: behaviour at cold temperatures taken from [Meilinger et al., 1995]. Right: behaviour at higher stratospheric temperatures, as described by [Carslaw et al., 1997].



The various ways that we use to calculate the refractive index for commonly encountered particle types in GOMOS data are summarized in Table 5. The latest version (v. 3.00) of AerGOM provides a tool to infer the aerosol type from different criteria. This aspect in discussed in Section 3.2.6.3.

Table 5 - Types of particles to be expected in the GOMOS data, with characteristics. The methods used to estimate composition (from temperature) and refractive index are also indicated.

Type	State/Morphology Composition	Weight percentage	Refractive index
Background	Liquid/Spherical H_2O/H_2SO_4	[Carslaw et al., 1997]	[Krieger et al., 2000]
Volcanic	Liquid/Spherical H_2O/H_2SO_4	[Carslaw et al., 1997]	[Krieger et al., 2000]
Cirrus	Solid/Crystalline H_2O	/	[Krieger et al., 2000]
NAT PSC	Solid/Amorphous HNO_3/H_2O	[Meilinger et al., 1995]	[Krieger et al., 2000]
STS PSC	Liquid/Spherical H_2O/H_2SO_4 /HNO_3	[Meilinger et al., 1995]	[Krieger et al., 2000]
Ice PSC	Solid/Crystalline H_2O	/	[Warren, 1984]

In the AerGOM processor, the standard procedure is to evaluate the ECMWF-derived temperature at the tangent altitude under consideration, and then interpolate the necessary optical extinction cross-sections in numerical tables at different wavelengths and temperatures. For sulfate aerosols and PSCs, these tables are precalculated using the temperature-controlled weight percentages for HNO_3 and H_2SO_4 described above. When other aerosol types are identified, the processor is able to apply the associated optical Mie coefficients.

3.1.5 Assumptions on measurement geometry

3.1.5.1 The Earth model

For our purposes, the Earth can be accurately described by an ellipsoid, a surface of revolution of an ellipse with a semi-major axis *a* (the equatorial radius) and a semi-minor axis *b* (the polar radius). The World Geodetic System 1984 (WGS84; see [NIMA, 1984]) uses the values in Table 6.

Table 6 - WGS84 ellipsoid definition.

equatorial radius	*a*	6 378 137.0 m
polar radius	*b*	6 356 752.314 m
flattening	*f*=(*a*-*b*)/*a*	1/298.257

At a given latitude φ, the Earth surface can locally be considered as spherical, with a local equivalent radius R_{eq} that can be evaluated as:

(24) $$R_{eq}(\varphi) = \sqrt{\frac{(a^2\cos\phi)^2 + (b^2\sin\phi)^2}{(a\cos\phi)^2 + (b\sin\phi)^2}}$$

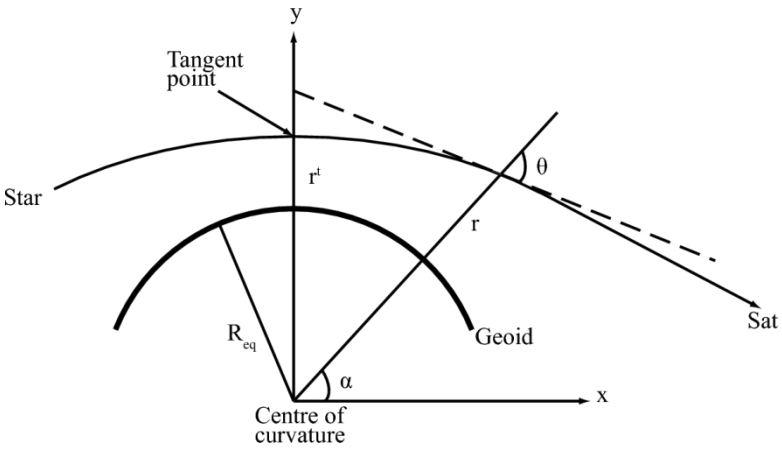
(25) $$= a\sqrt{\frac{\cos^2\phi + (1-f)^4\sin^2\phi}{1-(2f-f^2)\sin^2\phi}}$$

3.1.5.2 Homogeneous atmospheric shells

As with the surface of the Earth, the atmosphere can locally be considered as spherical in an excellent approximation. Due to a lack of information content in the GOMOS measurements it is not possible to retrieve gas concentrations and aerosol extinction coefficients at every atmospheric location in the line of sight of the instrument. This is the reason why we assume spherically homogeneous atmospheric layers (or shells), meaning that all physical quantities of interest (gas concentrations, refractive indices, aerosol extinction values, temperatures etc.) are considered to be constant within one spherical shell. One of the important consequences of optical theory is the fact that a ray of light in a spherically symmetric atmosphere with homogeneous shells propagates in a plane (containing the source, the detector and the center of the sphere). The propagation

problem is therefore essentially 2-dimensional, and can be described with polar coordinates (Figure 7): **r**=(r, θ).

Figure 7 - A spherically homogeneous atmosphere with polar coordinates. An incoming light ray is deflected downward due to refraction.



An important remark should be made here concerning the target species of the AerGOM processor. For relatively well-mixed particle populations (such as the background stratospheric aerosol layer) the assumption of homogeneous layers should approximate reality quite well. Even some cloud types (tropical subvisual cirrus) are known to be very thin cloud layers but with a horizontal extent of hundreds of kilometers. However, other particle populations will be more localised (volcanic plumes, certain Polar Stratospheric Clouds). For these cases, the retrieval should be considered as a *homogeneous layer equivalent* retrieval.

3.1.6 Atmospheric refraction

3.1.6.1 Refractive index of air

The corrected Edlén law [Edlén, 1966] is still widely used in atmospheric science; more specifically, it is implemented in the current GOMOS L1 processor [Kyrölä and Blanot, 2007]. However, a more accurate formulation was found by [Peck and Reeder, 1972] for air at standard pressure ($P_{stp} = 1013.25$ mb) and temperature ($T_{stp} = 288.15$ K), having 330 ppm of CO_2:

(26) $$(n_{stp} - 1) \times 10^8 = 8060.51 + \frac{2480990}{132.274 - \lambda^{-2}} + \frac{17455.7}{39.32957 - \lambda^{-2}}$$

with λ expressed in μm. The refractive index at a general temperature *T* and pressure *P* is found by:

(27) $$n - 1 = \frac{T_{stp}}{P_{stp}} \frac{P}{T} (n_{stp} - 1)$$

or, using the ideal gas law $P = Nk_BT$:

(28) $$n - 1 = \frac{N}{N_{stp}}(n_{stp} - 1)$$

with $N_{stp} = 2.546899 \times 10^{19}$ molecules cm^{-3}. In this way we can evaluate the refractive index everywhere if we know the local air density. The latter can be obtained from the ECMWF profiles available in the GOM_EXT files.

3.1.6.2 Refractive bending of optical paths

Due to the changing refractive index of air with altitude, the direction of incoming light is altered inside the atmosphere. Before doing any data retrieval, it is necessary to know the refracted optical paths, that can be calculated from the position of the spacecraft and the star. The point on the optical path that is closest to the Earth surface (at an altitude *h*) is called the *tangent point*. The paths and tangent points have already been calculated with the GOMOS Level 1 processor, and the associated tangent altitudes have been stored in the GOM_EXT data files. It is not necessary to repeat the optical path determination for this project; knowledge of the tangent altitude suffices (as will be shown later).

In radial coordinates, we have the following relation between the radial coordinate *r*, the angle between the path and the radial vector, and the path coordinate s (see Figure 7) [Peck and Reeder, 1972]:

(29) $$\frac{\mathrm{d}r}{\mathrm{d}s} = \cos\theta$$

Snell's law for refraction in circular symmetry reads:

(30) $$nr\sin\theta = r_g$$

with r_g a constant that is known as the *impact factor*. For our purposes, it is better to express the equation in terms of the refractive index and the radial coordinate at the tangent point, since tangent altitudes are given in the GOM_EXT files. At the tangent point, where $\theta = \pi/2$:

(31) $$nr\sin\theta = r_g = n(r^t)r^t$$

with r^t the distance of the tangent point to the Earth center, which can also be expressed with the *tangent altitude h*:

(32) $$r^t = R_{eq} + h$$

Finally, the infinitesimal line element along the optical path can be written, using Eqn. (29) and (31), as:

(33) $$\mathrm{d}s = \frac{1}{\cos\theta}\mathrm{d}r = \frac{1}{\sqrt{1-\sin^2\theta}}\mathrm{d}r = \frac{1}{\sqrt{1-\left(\frac{n(r^t)r^t}{n(r)r}\right)^2}}\mathrm{d}r$$

Notice that the line element depends on wavelength through the refractive index.

3.1.6.3 Other refractive effects

Since the atmospheric neutral air density decreases exponentially with altitude, the refractivity *n*-1 also exhibits this behaviour. Therefore, optical paths with lower tangent points experience stronger refractive bending then higher ones. In this way, parallel incident rays are refracted into a diverging beam, resulting in so-called *refractive dilution*: the light intensity decreases.

Furthermore, as we have seen, the air refractive index depends on wavelength. At short wavelengths, light experiences more downward bending since the refractive index is larger. *Chromatic refraction* therefore leads to an atmospheric prism effect: the spatial separation of rays of different color. For a multi-wavelength instrument such as GOMOS, this means that each spectral pixel measurement needs to be assigned to another tangent point.

Finally, the atmosphere exhibits fine-scale irregularities that are generated by internal gravity waves and turbulence. These irregularities manifest themselves as air density perturbations with characteristic scales from a few kilometers down to a dissipation scale (smaller than 1 meter). The associated refractive index perturbation causes the well-known *stellar scintillation*, which, if untreated, can render a transmittance measurement worthless.

All three effects are corrected for in an early stage of the current Level 2 GOMOS data processing. For all science details, see [Kyrölä and Blanot, 2007] or [Kyrölä and Blanot, 2012]. The scintillation component (fast refractive variation) is removed from the full transmittance with the aid of the GOMOS fast photometer data, while the dilution component (slow refractive perturbation) is calculated from an external air density profile. Finally, chromatic refraction is accounted for by an interpolating scheme where all transmittance measurements for all wavelengths are evaluated at the same tangent altitude, corresponding to a reference wavelength of $\lambda_{ref} = 500$ nm.

The finally obtained transmittance measurements can then be considered as determined only by atmospheric absorption and scattering by gases and particles/clouds. These are the data that have been stored in the GOM_EXT data files.

3.1.6.4 Residual scintillation

After the ENVISAT launch, an unpleasant feature was found in the data. After the above-mentioned scintillation removal, a remaining perturbation was still present in the measurements. Analysis showed that this component was associated with isotropic scintillation: while the operational algorithm is perfectly capable to remove anisotropic scintillation (caused by atmospheric fluctuations in homogeneous horizontal density layers), the code breaks down when isotropic scintillation is present. Caused by random density fluctuations, it is impossible to find a one-to-one correspondence between the fast photometer signals and the photometer values at different wavelengths. The random nature of these fluctuations forced the construction of a statistical method to estimate the amplitude of these perturbations. This so-called *Full Covariance Matrix* (FCM) method produces an error covariance matrix that is added to the transmission covariance matrix prior to the Level 2 retrieval chain. The perturbations do not necessarily disappear in the obtained retrievals, but are characterized as statistically insignifiant. See [Kyrölä and Blanot, 2012] for more details.

For AerGOM, it was decided to incorpolate this FCM method unaltered into the processor. The method (that can be optionnally switched on or off) needs extra information from the Level 1 transmission files.

3.1.7 Atmospheric weight functions

As already said, homogeneous spherical layers imply propagation of light in the plane formed by the star, the satellite and the center of the Earth. The problem becomes twodimensional and can be described with polar coordinates (r,α). Eqn. (1) therefore reduces to:

(34) $$\tau_{tot}(\lambda) = \int_{star}^{sat} \beta_{tot}(r(s), \alpha(s), \lambda) \mathrm{d}s$$

However, the assumption of spherical homogeneous layers also imply that the angular dependence can be dropped:

(35) $$\tau_{tot}(\lambda) = \int_{star}^{sat} \beta_{tot}(r(s), \lambda) \mathrm{d}s$$

Our problem has become one-dimensional. We can now transform it to a purely radial problem by using Eqn. (33):

(36)
$$\tau_{tot}(\lambda, r^t) = \int_{star}^{sat} \frac{\beta_{tot}(r,\lambda)}{\sqrt{1-\left(\frac{n(r^t)r^t}{n(r)r}\right)^2}} \mathrm{d}r$$

(37) or:
$$\tau_{tot}(\lambda, r^t) = \int_{star}^{sat} \beta_{tot}(r,\lambda) \frac{n(r)r}{\sqrt{(n(r)r)^2 - (n(r^t)r^t)^2}} \mathrm{d}r$$

The integral can be cut off outside the Earth atmosphere $r > r_{top}$ where $\beta_{tot} = 0$. Furthermore, since the integrand is symmetric with respect to the tangent point r^t, we can write the integral as:

(38)
$$\tau_{tot}(\lambda, r^t) = 2\int_{r^t}^{r_{top}} \beta_{tot}(r,\lambda) \frac{n(r)r}{\sqrt{(n(r)r)^2 - (n(r^t)r^t)^2}} \mathrm{d}r$$

The slant path optical thickness is thus the integral of the total extinction altitude profile, but weighted with a special function that depends on refractive bending and the associated tangent altitude. Once again: this weight function depends on wavelength through the refractive index.

3.1.8 Size information on a population of aerosol particles

As was already described in Section 3.1.3, the optical extinction by a particle population can be described using Eqn. (12) based on the extinction cross-section $C_{ext} = \pi a^2 Q_{ext}$ and a particle size distribution (PSD) $f_a(a)$, where a is the particle radius.

Conversely, the spectral dependance of the extinction contains information that can be used to retrieve size information by inversion of Eqn. (12). Hence, while the spectral and spatial inversion lead to the desired quantities for gases (local gas concentration altitude profiles), one more inversion is necessary to transform the obtained aerosol extinction spectra at each local altitude to PSD. This step is sometimes called the *radial inversion*. The purpose of the radial inversion is the retrieval of the particle size distribution $f_a(a)$ starting from the known (measurement-derived) extinction spectrum $\beta_{aero}(\lambda)$.

In practice, remote sounders provide extinction measurements at a restricted number of spectral channels that limits the information available to retrieve the PSD, and the number of degrees of freedom that can be effectively retrieved. Therefore, it is useful to constrain the problem in order to reduce the ill-posedness of the inversion problem. A frequently used method consists in using a prescribed analytical form for the particle size distribution. A very common choice of function for the particle size distribution is the lognormal function defined as

(39)
$$f(a) = \frac{1}{\sqrt{2\pi a}} \exp\left(\frac{-\log^2(a/a_m)}{2\log^2(s)}\right)$$

The mode radius ρ and mode width σ of this distribution are given respectively by the median radius $\rho = a_m$ (expressed in μm) and the quantity $\sigma = \log(s)$ (dimensionless), with *s* the

geometrical standard deviation *s*. This function has the advantage of ensuring the positive character of the particle size distribution and is only defined for positive radii r. In the case of a homogeneous population of aerosol particles, it provides a reasonable estimate of the real distribution by retrieving only three parameters: the mode parameters ρ and σ (respectively called the median radius and the mode width), and the particle number density *N* expressed in number of particles per cm^3. When particles of different origins coexist (e.g. aged background aerosols and fresh volcanic particles), the PSD might be preferably expressed as a sum of particle modes, each of them characterized by a median radius ρ_i, a mode width σ_i and particle number density N_i.

Even if the prescription of an analytical function for the PSD reduces the number of measurements needed to retrieve size information, the inversion of Eqn. (12) remains an ill-conditioned problem and the quality of the retrieved size parameters depends on the information content of the set of extinctions: the larger the spectrum, the more stable and reliable the solution. The use of a constraint to regularize the inversion problem is often needed to provide a solution with the necessary stability King et al., 1978].

Another way to solve the radial inversion problem is to discretize the PSD using a set of n_a fixed particle classes, the i^{th} of them centred on a particle radius a_i. The PSD is then retrieved as a discrete function f=[f_1, …, f_{na}] for the corresponding size bins with centres [a_1, …, a_{na}]. In this case also, the ill-posedness of the problem is alleviated by using some regularization constraint able to increase the stability of the solution.

The second approach was preferred for the retrieval of size information quantities using AerGOM, and a dedicated algorithm was developed for this purpose. The methodology used is discussed in Chapter 3.2.6.

3.2 AerGOM Implementation

3.2.1 From satellite measurements to vertical abundance profiles

The first step in the chain of GOMOS inversions involves the retrieval of local gas concentration profiles *N*(*r*) and aerosol extinction profiles $\beta_{aero}(r,\lambda_i)$ from the GOMOS transmittance measurements $\Im$. From the discussion in Section 3.1, the general equation that connects these quantities can be derived from Eqns. (1)-(5), (19), and (33), giving:

$$\Im(r^t,\lambda)=\frac{I(r^t,\lambda)}{I_0(r^t,\lambda)}=\exp\left[-\int_{r'}^{r^{top}} g(r',r)(C_{air}(\lambda)N_{air}(r)+\ldots+q_3(\lambda)\beta_{aero}(r,\lambda_3))\mathrm{d}r\right] \quad (40)$$

with $g(r^t,r)$ the weights that take into account the contributions from different atmospheric layers (see Eqn. (33)). The inversion of Eqn. (40) to obtain estimates of $N_{air}(r)$, …, $\beta_{aero}(r,\lambda_3)$ can in principle be tackled in one step. However, this so-called global one-step inversion is computationally very demanding since the problem has to be solved for all altitudes and all tangent wavelengths at once. Faster processing requires us to split up the inversion in two steps: *spectral inversion*, followed by *spatial inversion*.

3.2.2 Main assumptions

Our spectral inversion actually resembles the operational GOMOS Level 2 (IPF) method only in principle: at every tangent altitude separately, the transmittance spectrum is inverted to the slant path integrated column density and aerosol optical thickness using a nonlinear least-squares algorithm (Levenberg-Marquardt). However, in IPF, NO_2 and NO_3 are retrieved separately with a DOAS method, a quadratic polynomial of wavelength is assumed for the aerosol spectral law and is implemented in a peculiar way, and only spectrometer A is used (SPB1 and SPB2 are exclusively used for the retrieval of O_2 and H_2O). The AerGOM spectral inversion method on the other hand has the following features:

- No DOAS is applied, all gases and aerosols are retrieved simultaneously with a Levenberg-Marquardt code. It is nevertheless possible to remove the Rayleigh scattering contribution using ECMWF data before the inversion, to avoid retrieval correlations with the spectrally similar aerosols.

- A quadratic polynomial of inverse wavelength is chosen as standard aerosol spectral shape. However, the code is equipped with a number of spectral laws to choose from if necessary. All laws are parameterized by extinction coefficients (as explained in Sections 3.1.3.4-3.1.3.5).

- The code extends the spectral coverage by using the SPB1 spectrometer, which is of crucial importance for the retrieval of size information (See Section 3.2.5.5).

During spatial inversion, slant path integrated gas column densities and particle optical thickness *altitude profiles* are inverted to local gas concentration profiles and particle extinction profiles. The operational GOMOS processor does this *separately* for each gas and aerosol spectral coefficient, and assumes therefore (wrongly) that the retrievals of two different species after the spectral inversion are uncorrelated. The inversion is linear (it is a simple matrix equation), and a Tikhonov altitude smoothing constraint is applied. The AerGOM *full* spatial inversion module on the other hand has the following features:

- All tangent integrated quantities are inverted together to local quantities. This problem has larger dimensions, but has also the major advantage that the covariance between species resulting from the spectral inversion can be taken into account.

- An altitude regularization is implemented using a first-derivative linear operator. A special profile scaling is applied that resolves problems related to the large dynamic range of profiles in altitude and between species.

3.2.3 The forward model

3.2.3.1 Discretization

The transmittance spectra in the GOM_EXT files are delivered at an array of discrete tangent altitudes h_i (and thus discrete radial values $r^{t,i}$) for $i = 1 \ldots n_k$, where $i = 1$ corresponds to the first

layer at the top of the atmosphere, and $i = n_h$ to the layer close to the ground. We will now make our homogeneous layer assumption more specific. We assume that all gas densities, aerosol extinction coefficients, refractive indices etc. are constant within a layer that is centered on each GOMOS tangent altitude r_i^t . So we have for the total extinction:

(41) $$\beta_{tot}(r,\lambda) = \beta_{tot}(r_i^t,\lambda) \quad \text{for} \quad \frac{r_i^t + r_{i+1}^t}{2}{}^2 \le r < \frac{r_{i-1}^t + r_i^t}{2}$$

The slant path optical thickness (Eqn.(38)) at tangent altitude r_i^t can now be discretized as follows:

(42) $$\tau_{\text{tot}}(\lambda,r_i^t) = \beta_{\text{tot}}(r_i^t,\lambda) \int_{r=\frac{r_i^t+r_{i-1}^t}{2}}^{r_i^t} 2\frac{n(r_i^t)r}{\sqrt{(n(r_i^t)r)^2 - (n(r_i^t)r_i^t)^2}}\,\mathrm{d}r$$

(43) $$+ \sum_{j=1}^{i-1} \beta_{\text{tot}}(r_j^t,\lambda) \int_{r=\frac{r_j^t+r_{j+1}^t}{2}}^{\frac{r_j^t+r_{j-1}^t}{2}} 2\frac{n(r_j^t)r}{\sqrt{(n(r_j^t)r)^2 - (n(r_i^t)r_i^t)^2}}\,\mathrm{d}r$$

The first term contains a singularity at $r = r_i^t$. However, the problem can be solved by a substitution, that also allows explicit calculation of the second integral:

(44) $$x = \frac{\sqrt{(n(r_j^t).r)^2 - (n(r_i^t).r_i^t)^2}}{n(r_j^t)}$$

Then:

(45) $$\mathrm{d}x = \frac{n(r_j^t)r}{\sqrt{(n(r_j^t)r)^2 - (n(r_i^t)r_i^t)^2}}\,\mathrm{d}r$$

and both integrals therefore have the form:

(46) $$G = 2\int \mathrm{d}x = x_{\mathrm{up}} - x_{\mathrm{low}}$$

(47) or: $$\tau_{\mathrm{tot}}(\lambda, r_i^t) = \sum_{j=i}^{n_h} G(r_i^t, r_j^t)\beta_{\mathrm{tot}}(r_j^t, \lambda)$$

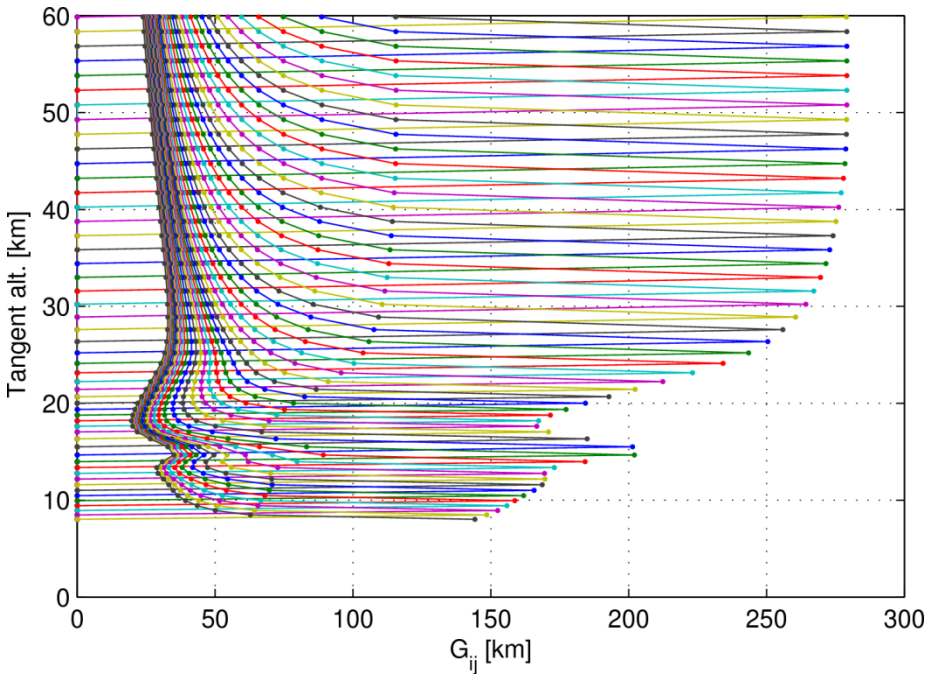
(48) with: $$G(r_i^t, r_j^t) = 2\left(\frac{\sqrt{\left[n(r_j^t)\frac{r_j^t + r_{j-1}^t}{2}\right]^2 - \left[n(r_i^t).r_i^t\right]^2}}{n(r_j^t)} - \frac{\sqrt{\left[n(r_j^t)\frac{r_j^t + r_{j+1}^t}{2}\right]^2 - \left[n(r_i^t).r_i^t\right]^2}}{n(r_j^t)} \right)$$

(49) and: $$G(r_i^t, r_i^t) = 2\left(\frac{\sqrt{\left[n(r_i^t)\frac{r_i^t + r_{i-1}^t}{2}\right]^2 - \left[n(r_i^t).r_i^t\right]^2}}{n(r_i^t)} \right)$$

and of course $G(r_i^t, r_j^t) = 0$ for $r_i^t > r_j^t$ (i.e. tangent altitudes r_j^t below r_i^t do not contribute to the slant optical thickness at tangent altitude r_i^t. The geometric weight function (dimension: length) $G(r_i^t, r_j^t)$ actually represents the path length traversed through the *j*-th atmospheric layer for an optical path with tangent point at $r = r_i^t$.

All weight functions can thus be evaluated when the tangent altitudes and the refractive indices are known. The entire model is specified by Eqns. (47) and (48), and (49). In theory these weight functions should be evaluated for each wavelength since refractive index depends on wavelength. This is not necessary; the GOM_EXT data have been corrected for chromatic refraction with an interpolation scheme: all transmittance spectra are associated with one tangent altitude and its corresponding optical path, which is calculated at a reference wavelength (to be found in the GOM_EXT files). It is this reference wavelength that we will use to evaluate the weight functions. Figure 8 shows a few actual GOMOS examples, representative for the lower atmosphere.

Figure 8 - Examples of GOMOS geometric weight functions for the lower atmosphere. Notice how the functions values are smaller at lower altitudes, since the tangent points of subsequent measurements are closer together due to atmospheric refraction, leading to thinner layers.



3.2.3.2 Contribution of the different species

We have seen that total extinction coefficient equals the sum of the different species extinction coefficients (Eqn. (2)). Therefore, Eqn. (47) becomes:

(50)
$$\tau_{tot}(\lambda, r_i^t) = \sum_{j=i}^{n_h} G(r_i^t, r_j^t)\left\{\beta_{air}(r_j^t, \lambda) + \beta_{O_3}(r_j^t, \lambda) + \ldots + \beta_{aero}(r_j^t, \lambda)\right\}$$

$$= \sum_{j=i}^{n_h} G(r_i^t, r_j^t)\left\{C_{air}(\lambda) N_{air}(r_j^t)\right\}$$

$$+ \sum_{j=i}^{n_h} G(r_i^t, r_j^t)\left\{C_{O_3}(r_j^t, \lambda) N_{O_3}(r_j^t)\right\} + \ldots$$

$$+ \sum_{j=i}^{n_h} G(r_i^t, r_j^t)\left\{q_1(\lambda)\beta_{aero}(r_j^t, \lambda_1)\right\} + \ldots$$

$$+ \sum_{j=i}^{n_h} G(r_i^t, r_j^t)\left\{q_3(\lambda)\beta_{aero}(r_j^t, \lambda_3)\right\}$$

where we have assumed an aerosol spectral law with 3 spectral functions q_1, q_2 and q_3. As is implicitely indicated in the equation, the air cross-section does not depend on the position (it is independent of temperature). The same is true for the aerosol spectral functions q_i. Gas absorption cross-sections do however depend on temperature. From now on we will make the assumption that the cross-section is constant along the line of sight and equals the cross-section at the tangent point. This assumption makes sense because most of the absorption takes usually place around the tangent point. So, for instance, for ozone we have $C_{O_3}(r_j^t, \lambda) \approx C_{O_3}(r_i^t, \lambda)$, and the ozone term becomes:

(51)
$$\tau_{O_3}(\lambda, r_i^t) = C_{O_3}(r_i^t, \lambda) \sum_{j=i}^{n_h} G(r_i^t, r_j^t) N_{O_3}(r_j^t)$$

(52)
$$= C_{O_3}(r_i^t, \lambda) \aleph_{O_3}(r_i^t)$$

with $\aleph_{O_3}(r_i^t)$ the slant path integrated column density for ozone. On the other hand, the slant path aerosol optical thickness is given by:

(53)
$$\tau_{aero}(\lambda, r_i^t) = q_1(\lambda) \sum_{j=i}^{n_h} G(r_i^t, r_j^t)\beta_{aero}(r_j^t, \lambda_1) + \ldots$$

$$+ q_3(\lambda) \sum_{j=i}^{n_h} G(r_i^t, r_j^t)\beta_{aero}(r_j^t, \lambda_3)$$

(54)
$$\tau_{aero}(\lambda, r_i^t) = q_1(\lambda)\tau_{aero}(\lambda_1, r_i^t) + \ldots + q_3(\lambda)\tau_{aero}(\lambda_3, r_i^t)$$

which shows that the slant path aerosol optical thickness is described by the same spectral law as the local extinction (but with different parameters).

To summarize, the slant path optical thickness is given by:

(55) $$\tau_{tot}(\lambda, r_i^t) = C_{air}(r_i^t, \lambda)\aleph_{air}(r_i^t) + C_{O_3}(r_i^t, \lambda)\aleph_{O3}(r_i^t) + \ldots + q_3(\lambda)\tau_{aero}(\lambda_3, r_i^t)$$

3.2.3.3 Spectral/spatial separation

What we actually have accomplished is a separation of the full inversion problem in two easier subproblems. The first step is called the *spectral inversion*, and it involves inversion of Eqn. (55) for all wavelengths and at every tangent altitude separately to slant path integrated gas column densities and aerosol optical thickness. The second step involves equations such as

(56) $$\aleph_{O3}(r_i^t) = \sum_{j=i}^{n_h} G(r_i^t, r_j^t)\mathrm{N}_{O3}(r_j^t)$$

(57) and $$\tau_{aero}(\lambda_3, r_i^t) = \sum_{j=i}^{n_h} G(r_i^t, r_j^t)\beta_{aero}(r_j^t, \lambda_3)$$

and consists of inverting the slant path integrated gas column densities and aerosol optical thickness at all tangent altitudes to local gas concentration and aerosol extinction profiles. This spatial deconvolution is appropriately called *spatial inversion*.

3.2.4 Spectral inversion

3.2.4.1 The spectral matrix

At every tangent altitude r_i^t, we can now construct a matrix $\mathbf{S}^{(i)}$ with dimensions $(n_\lambda \times n_c)$, with cross-sections and aerosol spectral functions:

(58) $$\mathrm{S}^{(i)} = \begin{bmatrix} C_{air}(\lambda_1) & C_{O3}^{(i)}(\lambda_1) & C_{NO2}^{(i)}(\lambda_1) & C_{NO3}^{(i)}(\lambda_1) & q_1(\lambda_1) & q_2(\lambda_1) & q_3(\lambda_1) \\ \vdots & & & & & & \\ C_{air}(\lambda_{n\lambda}) & C_{O3}^{(i)}(\lambda_{n\lambda}) & C_{NO2}^{(i)}(\lambda_{n\lambda}) & C_{NO3}^{(i)}(\lambda_{n\lambda}) & q_1(\lambda_{n\lambda}) & q_2(\lambda_{n\lambda}) & q3(\lambda_{n\lambda}) \end{bmatrix}$$

with n_λ the number of spectral pixels and n_c the number of species to be retrieved (every aerosol spectral function counts as one species). The Rayleigh cross-section C_{air} is calculated with Eqn. (6). The gas cross-sections are interpolated at the GOMOS wavelengths and the tangent point temperature (using 2D interpolation) in the lookup table of cross-sections as function of wavelength and temperature (See Sections 3.1.2.1 and 3.2.3.2 for a discussion about the reference cross-sections). And the functions $q_i(\lambda)$ are the aerosol spectral functions of Eqn. (23). Eqn. (55) can now be written as:

(59)
$$\underset{(n_\lambda \times 1)}{\boldsymbol{\tau}_{tot}^{(i)}} = \underset{(n_\lambda \times n_c)}{\mathbf{S}^{(i)}} \cdot \underset{(n_c \times 1)}{\aleph^{(i)}}$$

where $\aleph^{(i)}$ represents the vector with slant path integrated quantities for all species at the *i*-th tangent altitude, the outcomes of the spectral inversion. From Eqns. (3) and (40), the actual spectral model for the measured transmittance at tangent altitude r_i^t reads:

(60)
$$\Im_{\text{model}}^{(i)} = \exp\left(-\boldsymbol{\tau}_{tot}^{(i)}\right) = \exp\left(-\mathbf{S}^{(i)} \cdot \aleph^{(i)}\right)$$

3.2.4.2 Spectral inversion: the cost function

Finding a solution $\aleph^{(i)}$ at every tangent altitude for the nonlinear Eqn. (60) involves the minimization of a cost function:

(61)
$$M = \left(\Im^{(i)} - \exp\left[-\mathbf{S}^{(i)} \cdot \aleph^{(i)}\right]\right)^T \mathbf{S}_{\Im(\mathbf{i})}^{-1} \left(\Im^{(i)} - \exp\left[-\mathbf{S}^{(i)} - \aleph^{(i)}\right]\right)$$

with $\Im^{(i)}$ and $\mathbf{S}_{\Im(\mathbf{i})}$ respectively the measured transmittance vector and the associated covariance matrix. The latter is diagonal if we assume independent spectral measurements. However, when a Full Covariance Matrix calculation in carried out, it has off-diagonal elements.

3.2.4.3 Levenberg-Marquardt minimization

As already said, Eqn. (60) is nonlinear. We therefore need to use a nonlinear least-squares algorithm, and the Levenberg-Marquardt method [Press et al., 1992] is still the most robust one.

Minimization of the cost function *M* is equivalent to solving, for each tangent altitude r_i^t, the nonlinear equation:

(62)
$$\mathbf{S}_{\Im(\mathbf{i})}^{-1/2} \Im^{(i)} = \mathbf{S}_{\Im(\mathbf{i})}^{-1/2} \exp\left(\mathbf{S}^{(i)} \cdot \aleph^{(i)}\right)$$

The difference between the left hand and right hand side of the equation gives us the so-called *objective function* (of dimension (n_λ x 1))

(63)
$$\mathbf{F} = \mathbf{S}_{\Im^{(\mathbf{i})}}^{-1/2} \left(\Im^{(i)} - \exp\left(-\mathbf{S}^{(i)} \cdot \aleph^{(i)}\right)\right)$$

The numerical values of this function are implicitely squared and summed by the Levenberg-Marquardt (LM) algorithm in order to get the merit function *M*. Furthermore, the *Jacobian* of this function is also passed to the LM algorithm. We have:

(64) $$\mathbf{J}_{mn} = \frac{\partial F_m}{\partial \aleph_n^{(i)}} = \sum_k \left[\mathbf{S}_{\Im^{(i)}}{}^{-1/2} \right]_{mk} \mathbf{S}_{kn}^{(i)} \Im_k^{(i)}$$

With m=1… n_λ , n=1… n_c , and k=1… n_λ , or in matrix form:

(65) $$\mathbf{J} = \mathbf{S}_{\Im^{(i)}}{}^{-1/2} \mathrm{diag}(\Im^{(i)}) \mathbf{S}^{(i)}$$

At the final iteration, the LM algorithm returns the solution at the minimum, and the Jacobian at this point. The solution covariance matrix is given by:

(66) $$\mathbf{S}_{\aleph^{(i)},\mathbf{sol}} = (\mathbf{J}^T \mathbf{J})^{-1}$$

This operation can cause difficulties when the matrix between brackets is ill-conditioned. It is better to calculate the Cholesky factor **Q** first:

(67) $$\mathbf{Q} = chol\left(\mathbf{J}^T \mathbf{J}\right)$$

(68) and thus: $$\mathbf{Q}^T \mathbf{Q} = \mathbf{J}^T \mathbf{J}$$

(69) Then: $$\mathbf{S}_{\aleph(\mathbf{i}),\mathbf{sol}} = (\mathbf{Q}^T \mathbf{Q})^{-1} = \mathbf{Q}^{-1} \mathbf{Q}^{-T}$$

It is numerically much better to invert this 'square root' **Q** of $\mathbf{J}^T\mathbf{J}$.

It has to be noted that the Cholesky decomposition is only possible for Hermitian positive-definite matrixes. This condition is in principle fulfilled by the covariance matrix, but in real cases or poor measurements (e.g. in case of dim stars) where significant noise may lead to a badly conditioned problem, it may happen that the positive-definite character of the full covariance matrix is not fullfilled. In such a case, the use of the Cholesky decomposition fails and the retrieval cannot be performed successfully.

3.2.4.4 Rayleigh scattering correction

The cross-section for Rayleigh scattering ($C_{air} \approx \lambda^{-4}$) is a smooth function that potentially can interfere with the aerosol spectral functions $q_i(\lambda)$. As a matter of fact, when only very small particles are present, their optical extinction spectrum is identical to the one from the neutral air density, with the result that the spectral matrix $\mathbf{S}^{(i)}$ becomes ill-conditioned. This potential danger should be taken into account.

We can avoid problems by removing the air scattering contribution prior to spectral inversion. This can be done with an external air density profile, and in our case the ECMWF profile in the GOM_EXT data files can be used. If we denote the slant path integrated air density as $\aleph_{air}^0$, then the associated transmittance at wavelength λ is calculated as:

(70) $$\Im^0_{air}(\lambda) = \exp\left(-C_{air}(\lambda)\aleph^0_{air}\right)$$

The transmittance, corrected for Rayleigh scattering, is then:

(71) $$\Im_{corr}(\lambda) = \frac{\Im_{meas}(\lambda)}{\Im^0_{air}(\lambda)}$$

and the associated measurement variance:

(72) $$\mathbf{S}_{\Im corr}(i,j) = \frac{\mathbf{S}_{\Im}(i,j)}{\Im^0_{air}(i).\Im^0_{air}(j)}$$

Following this procedure, the spectral inversion is carried out as described before, however without the Rayleigh cross-section in the spectral matrix. The algorithm will be equipped with the possibility to choose if we want to correct for air or not.

3.2.5 Full spatial inversion

3.2.5.1 The model implementation

At the *i*-th tangent altitude, the slant path integrated vector can be expressed as follows. Let us use the following notation for the row vector with path lengths:

(73) $$\mathbf{G_{(i)}} = \begin{bmatrix} G_{i1} & G_{i2} & \cdots & G_{i(m-1)} & G_{im} \end{bmatrix}$$

with $m = 1,\ldots n_h$, and $G_{ij} = G(r_i^t, r_j^t)$ the optical path length as defined before (Eqn. (48) and (49)). The slant path ozone column (Eqn.(56)) can now be written as:

(74) $$\aleph^{(i)}_{O3} = \mathbf{G_{(i)}}\mathbf{N_{O3}}$$

with $\mathbf{N_{O3}}$ the local ozone concentration profile: $\mathbf{N_{O3}} = [n_{O3}(r_1^t)\cdots n_{O3}(r_{n_h}^t)]^T$. In the same way, a slant path aerosol optical thickness (for example at λ_3; Eqn. (57)) is given by:

(75) $$\tau^{(i)}_{aero}(\lambda_3) = \mathbf{G_{(i)}}\beta_{aero,3}$$

with $\beta_{aero,3}$ the aerosol extinction profile at λ_3.

The entire slant path integrated vector (with all constituents) at tangent point r_i^t can be expressed with the use of a block-diagonal matrix:

$$\aleph^{(i)} = \begin{bmatrix} G_{(i)} & 0 & \cdots & 0 & 0 \\ 0 & G_{(i)} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & G_{(i)} & 0 \\ 0 & 0 & \cdots & 0 & G_{(i)} \end{bmatrix} \cdot \begin{bmatrix} N_{air} \\ N_{o3} \\ \vdots \\ \beta_{aero,2} \\ \beta_{aero,3} \end{bmatrix}$$

(76)

The vector to be retrieved contains the altitude profiles of all local gas densities and aerosol extinction coefficients.

Stacking all tangent altitudes together, we obtain the *full spatial model*:

$$\aleph^{tot} = \begin{bmatrix} \aleph^{(1)} \\ \vdots \\ \aleph^{(n_h)} \end{bmatrix} = \begin{bmatrix} G_{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & G_{(1)} \\ \vdots & \ddots & \vdots \\ G_{(n_h)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & G_{(n_h)} \end{bmatrix} \cdot \begin{bmatrix} N_{air} \\ N_{o3} \\ \vdots \\ \beta_{aero,2} \\ \beta_{aero,3} \end{bmatrix}$$

(77)

or, in short notation:

$$\underset{(n_c.n_h \times 1)}{\aleph^{tot}} = \underset{(n_c.n_h \times n_c.n_h)}{\mathbf{G}^{\mathbf{tot}}} \cdot \underset{(n_c.n_h \times 1)}{\mathbf{N}^{\mathbf{tot}}} \quad (78)$$

The total covariance matrix for the slant path integrated densities and aerosol extinction is of course given by the block-diagonal matrix:

$$\mathbf{S}_{\aleph^{\mathbf{tot}}} = \begin{bmatrix} \mathbf{S}_{\aleph(\mathbf{1})} & & \\ & \ddots & \\ & & \mathbf{S}_{\aleph(n_h)} \end{bmatrix} \quad (79)$$

with $\mathbf{S}_{\aleph(\mathbf{i})}$ the covariance matrix of the solution of the spectral inversion at the *i*-th tangent altitude. So all covariances in between species are taken into account in the model.

3.2.5.2 Spatial inversion: optimal choice of the cost function

We will now invert the forward model (Eqn. (78)) using a least-squares approach. Two main issues have to be addressed:

- The problem should be well-conditioned to give a stable and useful solution. Therefore, one has to take into account the fact that atmospheric gas concentration and aerosol extinction profile values span several orders of magnitude, what can be resolved by applying an appropriate scaling. A natural choice for such scaling is to weight the quantities by their associated uncertainty which are available for the slant integrated quantities, and can be estimated for the retrieved quantities by a simple linear least-square inversion.

- The vertical profiles obtained as solution should render vertical structures associated with atmospheric patterns, but the influence of the experimental noise should be minimized. This can be realized by regularizing the inversion problem using a suitable value of the regularization parameters.

As a first step, we solve Eqn.(78) using a linear least-squares approach *without regularization*. This is perfectly possible since the matrix $\mathbf{G}^{\mathbf{tot}}$ is well-conditioned: it is square, and the kernels are strongly peaked (see Figure 8). The least-squares solution

(80) $$\mathbf{N}_{\mathbf{LS}}^{\mathbf{tot}} = \left[\mathbf{G}^{\mathbf{tot}}\right]^{-1} . \aleph^{tot}$$

has as associated covariance matrix:

(81) $$\mathbf{S}_{N\mathbf{tot}} = \left(\left(\mathbf{G}^{\mathbf{tot}}\right)^{T} . \mathbf{S}_{\aleph\mathbf{tot}}^{-\mathbf{1}} . \mathbf{G}^{\mathbf{tot}}\right)^{-1}$$

that we will use to scale the regularized problem. Formally, we can express the weighting by the uncertainty explicitly by defining a diagonal matrix **D** with elements equal to the least-square standard deviations:

(82) $$\mathbf{D}_{ii} = \sqrt{\mathbf{S}_{N\mathbf{tot},ii}}$$

Then, the least-square covariance matrix can be written as the decomposition:

(83) $$\mathbf{S}_{N\mathbf{tot}} = \mathbf{D}.\mathbf{R}.\mathbf{D}$$

With **R** the correlation matrix.

The second step consists in *regularizing the inversion*. The regularization is achieved with a classical Tikhonov altitude smooting, using a first derivative operator (see Appendix A) in order to minimize the first derivative in a suitable way to avoid spurious instabilities in the vertical profile. Since we are dealing with a full spatial inversion, we need a composed[2] first derivative operator $\mathbf{L}^{\mathbf{tot}}$. Applying a scaling by the least-squares error estimates discussed above using the associated diagonal matrix **D**, we define the scaled regularization operator as:

[2] $\mathbf{L}^{\mathbf{tot}}$ is a block-diagonal matrix, with one first-derivative operator $\mathbf{L}^{(\mathbf{j})}$ per species on the diagonal. Furthermore, the operator $\mathbf{L}^{(\mathbf{j})}$ associated with the j-th species is multiplied with its own regularization parameter μ_j, so that the strength of the regularization can be tuned separately for each species.

(84) $$\mathbf{L}_{\mathbf{S}}^{\mathbf{tot}} = \mathbf{L}^{\mathbf{tot}}.\mathbf{D}^{-1}$$

The cost function to be minimized therefore can be written as the sum of the least-squares term and the regularization term:

(85) $$M = \left[\aleph^{\mathbf{tot}} - \mathbf{G}^{\mathbf{tot}}.\mathbf{N}^{\mathbf{tot}}\right]^{T} \mathbf{S}_{\aleph\mathbf{tot}}^{-1}\left[\aleph^{\mathbf{tot}} - \mathbf{G}^{\mathbf{tot}}.\mathbf{N}^{\mathbf{tot}}\right]$$

(86) $$+ \left[\mathbf{N}^{\mathbf{tot}}\right]^{T}.\left(\mathbf{L}_{\mathbf{s}}^{\mathbf{tot}}\right)^{T}.\mathbf{L}_{\mathbf{s}}^{\mathbf{tot}}.\mathbf{N}^{\mathbf{tot}}$$

3.2.5.3 Minimization: linear inversion

Since the problem is linear, the solution can be found in one step and is actually the solution of the matrix equation:

(87) $$\begin{bmatrix} \mathbf{S}_{\aleph\mathbf{tot}}^{-1/2}\aleph^{\mathbf{tot}} \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{\aleph\mathbf{tot}}^{-1/2}\mathbf{G}^{\mathbf{tot}} \\ \mathbf{L}_{\mathbf{s}}^{\mathbf{tot}} \end{bmatrix} \mathbf{N}^{\mathbf{tot}}$$

Written out, the regularized solution is given by

(88) $$\mathbf{N}_{\mathrm{sol}}^{\mathbf{tot}} = \left(\mathbf{G}^{\mathbf{tot},T}.\mathbf{S}_{\aleph\mathbf{tot}}^{-1}.\mathbf{G}^{\mathbf{tot}} + \mathbf{L}_{\mathbf{s}}^{\mathbf{tot},\mathrm{T}}.\mathbf{L}_{\mathbf{s}}^{\mathbf{tot}}\right)^{-1}.\mathbf{G}^{\mathbf{tot},\mathrm{T}}.\mathbf{S}_{\aleph\mathbf{tot}}^{-1}\aleph^{\mathbf{tot}}$$

with covariance matrix:

(89) $$\mathbf{S}_{\mathbf{Ntot},\mathrm{sol}} = \left(\mathbf{G}^{\mathbf{tot},\mathrm{T}}.\mathbf{S}_{\aleph\mathbf{tot}}^{-1}.\aleph^{\mathbf{tot}} + \mathbf{L}_{\mathbf{s}}^{\mathbf{tot},\mathrm{T}}.\mathbf{L}_{\mathbf{s}}^{\mathbf{tot}}\right)^{-1}$$

Using Eqs (81), (83) and (84), the solution covariance can also be written as

(90) $$\mathbf{S}_{\mathbf{Ntot},\mathrm{sol}} = \mathbf{D}.\left(\mathbf{R}^{-1} + \mathbf{L}_{\mathbf{s}}^{\mathbf{tot},\mathrm{T}}.\mathbf{L}_{\mathbf{s}}^{\mathbf{tot}}\right)^{-1}.\mathbf{D}$$

Comparing this equation with Eqn. (83), we see that our regularization method works directly on the correlation matrix, while leaving the global structure of the standard deviations (the matrix D) unchanged.

3.2.5.4 Averaging kernels and vertical resolution

The averaging kernels for the solution of the full spatial model can be written as

(91) $$\mathcal{A} = \mathcal{G}.\mathcal{K}$$

where the gain matrix $\mathcal{G}$ expressing the sensitivity of the retrieved quantity to the measurement can be derived from Eqn. (88) :

(92) $$\mathcal{G} = \left(\mathbf{G}^{\mathbf{tot},T}.\mathbf{S}_{\aleph\mathbf{tot}}{}^{-1}.\mathbf{G}^{\mathbf{tot}} + \mathbf{L}_{\mathbf{s}}^{\mathbf{tot},\mathrm{T}}.\mathbf{L}_{\mathbf{s}}^{\mathbf{tot}}\right)^{-1}.\mathbf{G}^{\mathbf{tot},\mathrm{T}}.\mathbf{S}_{\aleph\mathbf{tot}}{}^{-1}$$

and $\mathcal{K}$ is the forward model, corresponding here to the full spatial inversion model given by Eqn. (78). Consequently, the averaging kernels is given by

(93) $$\mathcal{A} = \left(\mathbf{G}^{\mathbf{tot},T}.\mathbf{S}_{\aleph\mathbf{tot}}{}^{-1}.\mathbf{G}^{\mathbf{tot}} + \mathbf{L}_{\mathbf{s}}^{\mathbf{tot},\mathrm{T}}.\mathbf{L}_{\mathbf{s}}^{\mathbf{tot}}\right)^{-1}.\left(\mathbf{G}^{\mathbf{tot},\mathrm{T}}.\mathbf{S}_{\aleph\mathbf{tot}}{}^{-1}.\mathbf{G}^{\mathbf{tot}}\right)$$

Following Eqns. (81) and (89), the averaging kernel can be expressed in terms of the covariance matrices:

(94) $$\mathcal{A} = \mathbf{S}_{\mathbf{Ntot},\mathrm{sol}} \cdot \mathbf{S}_{N\mathbf{tot}}{}^{-1}$$

In the most general case, this $(n_c.n_h \times n_c.n_h)$ matrix, which expresses the sensitivity of the retrieval to the true vertical gas and aerosol profiles can show dependences between every pair of species at their respective altitudes. In practice, we can expect that dependences are small between contributions from remote heights and different species, so that a good approximation for the averaging kernels, saving much computer resources, is obtained by keeping only $(n_h \times n_h)$ blocks along the diagonal of $\mathcal{A}$, reading:

(95) $$\mathcal{A}_{approx} = \begin{bmatrix} \mathcal{A}_{air} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathcal{A}_{O3} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathcal{A}_{aero,\,\lambda 1} & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathcal{A}_{aero,\,\lambda 3} \end{bmatrix}$$

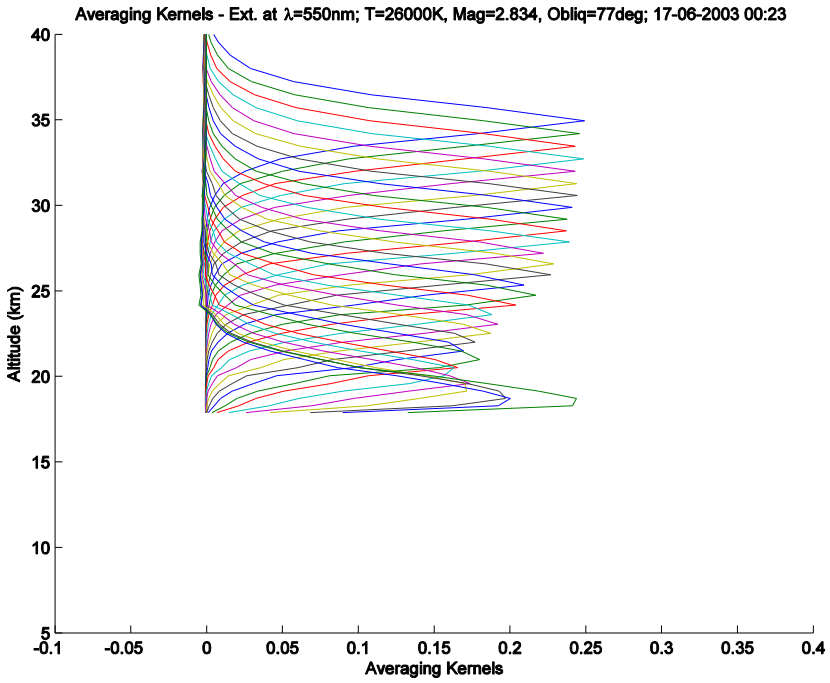
The averaging kernel is used to express the vertical resolution of the local vertical profiles. Different metrics can be used. Rodgers et al. (2000) proposes and discusses several choices, such as the width of the averaging kernel associated to the considered species and altitude, or the "Backus-Gilbert spread" used by Sofieva et al. (2004) in the framework of GOMOS. Here, we choose the simple approach of Purser and Huang (1992), discussed in [Rodgers et al, 2000]. In the simple particular case where each averaging kernel is mostly positive and has an area close to 1, the resolution can be estimated as the inverse of its maximum value. Scaling this quantity in the actual grid of the profile gives for the estimate of the vertical resolution:

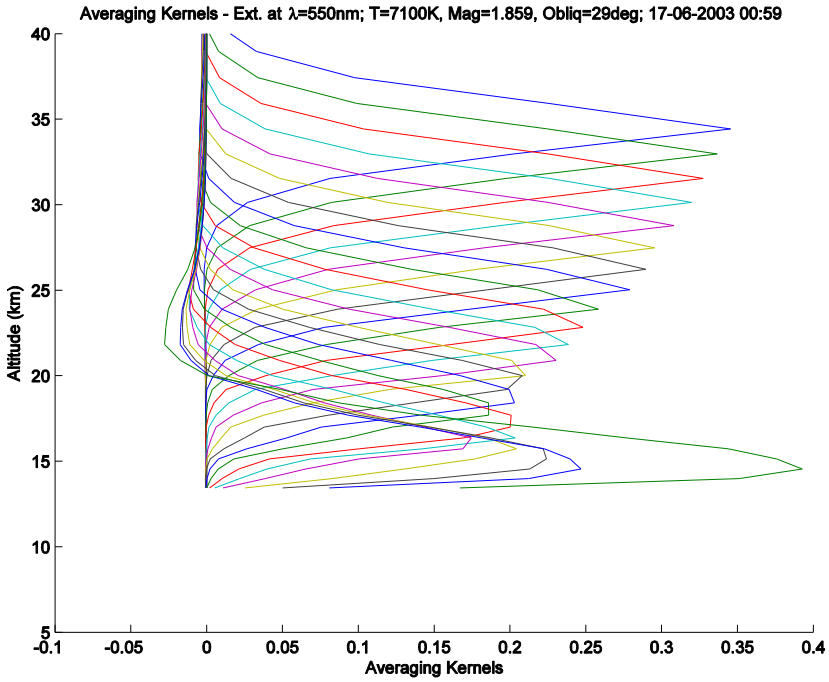
(96) $$\delta z_i = \mathcal{A}_{kk}{}^{-1} \cdot \frac{(z_{i+1} - z_{i-1})}{2} = \mathcal{A}_{kk}{}^{-1} \cdot \Delta z_i$$

where k is the index of $\mathcal{A}$ corresponding to the to ith altitude level of the vertical profile for the jth species. δz_i is an expression of the *resolving length*, providing a simple and computationally cheap characterization of the vertical resolution for each species at each altitude level. Notice from Eqn.

(93) that for a retrieval without regularization, the averaging kernels $\mathcal{A}$ equals the identity matrix, and δz_i equals the sampling resolution Δz_i.

Figure 9 - Averaging kernels for two occultations on 17 June 2003. Above: Star temperature=26000K, star magnitude=2.834, Obliquity=77°. Below: Star temperature=7100K, star magnitude=1.859, Obliquity=29°. When the obliquity is high, the star setting lasts much longer, and the vertical resolution is improved. On the other hand, the regularization decreases the vertical resolution. In this case, the regularization parameter is the same for both retrievals.





3.2.5.5 Extinction-derived quantities: Angström coefficient determination

The Angström exponent (AE) was originally introduced as a wavelength-independent constant in a power law to describe wavelength-dependent extinction (or optical depth) of light by aerosols (see [Angström, 1929]) :

(97) $$\beta(\lambda) = K \cdot \lambda^{-AC}$$

with *K*, a constant.

For the stratospheric aerosols, the AE is computed for both the extinction and the AOD by computing the slope of ln(b(l)) (or ln(AOD(l))) vs ln(l) using an error-weighted least-square linear fit method. In the current CCI-GOMOS version CCI_4.00, the spectral window for the fit is between 355 and 756 nm. The AE is considered determined only if enough (>4) extinction/AOD data points are available. Some data points might not be taken into account if they are negative (the retrieval allows for negative extinction values) which could potentially lead to biases in the AE estimation. The associated error is also provided.

It must be noted that a L1-norm fitting method would be more robust, but would not provide an error estimate on the AE, which is why the least-square approach was adopted.

3.2.6 Radial inversion:From aerosol extinction spectra to particle size distributions

3.2.6.1 Introduction

While the spectral and spatial inversions lead to the desired quantities for gases (local gas concentration altitude profiles), one more inversion is necessary to transform the obtained aerosol

extinction spectra at each local altitude to particle size distributions. This step is sometimes called the *radial inversion*. As was already described in Section 3.1.3.2, the optical extinction by a particle population at a location **r** can be calculated by integrating the contribution of all particles in this sampling characterized by a size distribution $f_a(a)$. The optical extinction can then be expressed as the sum of the extinction cross-section $C_{ext} = \pi a^2 Q_{ext}$ weighted by the PSD $f_a(a)$, giving at a location **r**:

(98) $$\beta_{aero}(\lambda, \mathbf{r}) = \int_{a=0}^{+\infty} \pi a^2 Q_{ext}(\lambda, a) f_a(\mathbf{r}, a) \mathrm{d}a$$

The purpose of the radial inversion is the retrieval of the PSD $f_a(a)$ starting from the known (measurement-derived) extinction spectrum $\beta_{aero}(\lambda)$.

3.2.6.2 General approach of radial inversion

As discussed in Section 3.1.8, two methods are commonly used to do the inversion. As a first method, the particle size distribution (PSD) $f_a(a)$ is modelised as an analytic function with a limited number of parameters, that are subsequently retrieved. Typically a lognormal distribution is assumed, parameterized by the total number density *N*, the mode radius a_m (the median of the distribution) and the distribution width *s* (the geometric standard deviation of the distribution); see e.g. [Vanhellemont and Fussen, 2002]. The reason for this choice is mainly that observed PSDs cover several orders of magnitude in particle size *a*, a behaviour that is well captured by a lognormal function (which is actually a normal distribution on a logarithmic scale). However, a number of problems are associated with this approach. Often, actual atmospheric particle populations have multiple modes ('fine', 'coarse') while a lognormal function is unimodal by definition. Equally important is the difficulty of the model error, which is unknown; standard error estimation alone will lead to small error bars for particle sizes where no information is to be expected.

Both problems are resolved with a second approach, in which the PSD is simply treated as an array of discretised bins (see e.g. [Twomey, 1975; King et al., 1978]). Of course the number of unknowns increases strongly, and the resulting matrix equation needs to be regularized with smoothing constraints since it is notoriously ill-conditionned for the simple reason that the kernels in the optical model are very smooth. But the retrieved PSD can basically assume any shape it wants, and the resulting error bars reflect the information content of the retrieval.

For the radial inversion of GOMOS data, we will use this second method. We developed a specific version for AerGOM, with its own particular features:

- We will define our PSD that is to be retrieved on a logarithmic particle size grid. If not, the number of bins will be too large, considering the fact that aerosol PSDs typically span several orders of magnitude. Furthermore, there is also a physical reason for this: the particle size a is by definition a positive quantity, so it can be modelled as $a \approx e^u$.

- The PSD f_a is by definition a positive quantity (number of particles per unit of volume per unit of particle size), so it is better to model it as $f_a \approx e^b$ and retrieve the new unknown b instead.

- Of course regularization will be necessary. We have chosen a classical approach using first derivative linear operator.

3.2.6.3 A pre-requisite : the aerosol composition

Retrieving particle size distribution from the aerosol extinction requires that we know the refractive index and its dependence in wavelength and in the local atmospheric conditions (mostly the temperature). This implies in turn that the aerosol composition is known. In the case of GOMOS, we cannot rely on available measurements to easily infer the composition, and assumptions have to be made using the parameters at our disposal. For instance:

- The *temperature* allows the exclusion of some kind of particles. In particular, icy particles requires temperature values below some threshold.

- Similarly, the *latitude* value can be used to exclude some particle types: PSCs are not expected at low latitudes, and the presence of low temperatures far from the polar regions is rather an indication of the presence of cirrus clouds.

- *Altitude* provides also a diagnostic tool for aerosol type: in the stratosphere, sulfate, which is the main source of aerosols in the middle atmosphere, is no longer liquid above ~35 km, making the presence of liquid sulfate aerosol very unlikely above this height.

- The *retrieved PSD* is an important a posteriori criterion to identify the aerosol type. Volcanic sulfate aerosols are known to have an averaged particle radius not larger than ~0.4 µm outside of volcanically active periods; and even in volcanic plumes, the averaged particle radius is expectd to be smaller than 1 µm. On the contrary, PSCs have much larger typical sizes (radius ~ 1 µm), while 1 µm looks like a lower limit for cirrus clouds particles of which the size can reach hundreds or even thousands of µm [Pruppacher and Klett, 1997].

The temperature and latitude values are used in AerGOM to discriminate liquid sulfate aerosols, PSCs and cirrus clouds, and to choose refractive index consequently for the PSD retrieval. From AerGOM version 3.0, an aerosol type flag was introduced, and all four parameters cited above (including the retrieved PSD) are used to provide a more accurate determination of the particle type. However, these parameters are not sufficient to provide an exhaustive discrimination of the aerosol type:

- The retrieved PSD provides in some cases outlying values of the mean particle size making the discrimination uncertain. For instance, a mean particle radius of about 0.9 µm is really large to be attributed to volcanic aerosols, but somewhat small to be attributed to cirrus clouds.

- Other aerosol types, like soot particles are also found in the stratosphere. Their typical size range overlaps with the size range of other aerosol particles types, impeding an unambiguous type detection.

Hence, assumptions have to be made, which are, at this stage, necessarily a limiting factor for the performance of the particle type detection. In AerGOM, version 3.0 (currently used), the assumptions chosen to provide the most likely local aerosol type, are the following:

Below 35 km altitude:

- If the temperature is lower than 197°K, the aerosol forms polar stratospheric clouds for latitudes higher than 45°.

- At other temperatures, aerosols are assumed to be liquid sulfate aerosol particles. At this stage, we do not dispose of effective criteria to discriminate clouds.

Above 35 km altitude, all available sulfuric acid is in vapour phase, and aerosols consist of meteoric smoke particles (typical radius range: 1.5 nm). Such small particle size range implies that the extinction spectral dependence follows, in the GOMOS spectral range, a decreasing function in λ^{-4} typical for the Rayleigh regime. If the spectral behavious does not follow the Rayleigh regime, it is assumed that aerosols consist of soot (black carbon). The composition of meteoric smoke particles may be olivine (Mg_2xFe_2-2xSiO4) or pyroxene (Mg_xFe_1-xSiO3), with x varying between 0 and 1 [Saunders et al., 2012]. At this time, we do not consider the case of meteoric smoke particle, and no PSD is retrieved so far.

3.2.6.4 Mathematical model development

3.2.6.4.1 Step 1: transformation to logarithmic grid

The particle radius a (with $a > 0$!) can be written as:

(99)
$$a = a_0 \exp\left(\frac{u - u_0}{\sigma_u}\right)$$

The constant a_0 determines the origin for our new u-grid: when $a=a_0$, $u=u_0$. Furthermore, after discretization (see below) σ_u determines the *increment* of the a-grid; if Δu equals the *linear* spacing of the u-grid, then:

(100)
$$\frac{a_{i+1}}{a_i} = \frac{a_0 \exp((u_{i+1} - u_0)/\sigma_u)}{a_0 \exp((u_i - u_0)/\sigma_u)} = \frac{\exp((u_i + \Delta u - u_0)/\sigma_u)}{\exp((u_i - u)_0)/\sigma_u} = \exp(\Delta u/\sigma_u)$$

For example, if we choose $\sigma_u = \Delta u$, then $a_{i+1}/a_i = e$.

Of course, we need a transformation of the PSD as well. First, notice the definition of a PSD: $f_a(a)\,da$ is the number of particles per cm^3 with a radius in the interval [a,a+da]. If we change to the new u-grid, this number should remain the same ! So:

(101) $$f_a(a)\mathrm{d}a = f_u(u)\mathrm{d}u$$

and therefore:

(102) $$f_u(u) = f_a(a)\frac{da}{du} = \frac{f_a(a)a}{\sigma_u}$$

Our physical extinction model, transformed to the u-grid, finally reads:

(103) $$\beta_{aero}(\lambda) = \int_{u=-\infty}^{+\infty} \pi a_0^2 e^{2(u-u_0)/\sigma_u} Q_{ext}(\lambda, a = a_0 e^{(u-u_0)/\sigma_u}) f_u(u)\mathrm{d}u$$

3.2.6.4.2 Step 2: transformation to logarithmic PSD

A (continuous) positivity constraint can be imposed on the PSD by modelling it as an exponential of a new function b:

(104) $$f_u(u) = f_{u0}\exp\left(\frac{b(u)-b_0}{\sigma_b}\right)$$

Once again, the constants have a specific meaning: f_{u0}, b_0 and σ_b play similar roles as the ones for the u-grid, a_0, u_0 and σ_u. Our extinction model (Eqn.(103)) reads:

(105) $$\beta_{aero}(\lambda) = \pi a_0^2 f_{u0} \int_{u=-\infty}^{+\infty} e^{2(u-u_0)/\sigma_u} Q_{ext}(\lambda, a = a_0 e^{(u-u_0)/\sigma_u}) e^{(b(u)-b_0)/\sigma_b}\mathrm{d}u$$

3.2.6.4.3 Step 3: Discretization of integral

Discretisation means that we now will split up the u-domain in equally-spaced discrete intervals (or 'bins') [u_i-Δu/2 , u_i+Δu/2] with centers u_i and width Δu. Of course, u_0 (for i=0) is the number that was defined before. In this way, every discrete point can be calculated as:

(106) $$u_i = u_0 + i\Delta u$$

with the index i running from -∞ to +∞.

The extinction model (Eqn. (103)) can now be written as a discrete sum of integrals over the bins:

(107) $$\beta(\lambda) = \sum_{i=-\infty}^{+\infty} \pi a_0^2 f_{u0} \int_{u_i-\Delta u/2}^{u_i+\Delta u/2} e^{2(u-u_0)/\sigma_u} Q_{ext}(\lambda, a = a_0 e^{(u-u_0)/\sigma_u}) e^{(b(u)-b_0)/\sigma_b} \mathrm{d}u$$

Now we make the (usual) assumption that $b(u)$ is in good approximation constant within the interval $[u_i-\Delta u/2, u_i+\Delta u/2]$. So we can evaluate $b(u)$ at the value u_i and bring it outside the integral:

(108) $$\beta(\lambda) = \sum_{i=-\infty}^{+\infty} \pi a_0^2 f_{u0} e^{(b(u_i)-b_0)/\sigma_b} \int_{u_i-\Delta u/2}^{u_i+\Delta u/2} e^{2(u-u_0)/\sigma_u} . Q_{ext}(\lambda, a = a_0 e^{(u-u_0)/\sigma_u}) \mathrm{d}u$$

The integral has to be evaluated numerically with Q_{ext} provided by a Mie code, and taking into account the refractive index of the particle type under consideration, as described in Section 3.1.4. We use the following notation for the result:

(109) $$\tilde{Q}(\lambda, u_i) = \int_{u_i-\Delta u/2}^{u_i+\Delta u/2} e^{2(u\_u_0)/\sigma_u} Q_{ext}(\lambda, a = a_0 e^{(u\_u_0)/\sigma_u}) \mathrm{d}u$$

And so our extinction model becomes:

(110) $$\beta(\lambda) = \sum_{i=-\infty}^{+\infty} \pi\, a_0^2 f_{u0} e^{(b(u_i)-b_0)/\sigma_b} \tilde{Q}(\lambda, u_i)$$

3.2.6.4.4 Step 4: Construction of matrix equation

In this next step we consider discrete wavelengths λ_i, change indices and introduce new notation:

(111) $$\beta(\lambda_i) \rightarrow \beta_i$$

(112) $$\tilde{Q}(\lambda_i, u_j) \rightarrow \tilde{Q}_{ij}$$

(113) $$b(u_j) \rightarrow b_j$$

and introduce the matrix **K** with coefficients:

(114) $$K_{ij} = \pi a_0^2 f_{u0} e^{-b_0/\sigma_b} \tilde{Q}_{ij}$$

then our extinction model (Eqn. (110)) can be elegantly written as the following *nonlinear* matrix equation:

(115) $$\beta = \mathbf{K} \cdot e^{\mathbf{b}/\sigma_b}$$

with β of course a column vector containing the retrieved aerosol extinction spectrum and **b** the column vector with the quantities to be retrieved.

3.2.6.5 Retrieval procedure

3.2.6.5.1 The cost function

The retrieval of the vector **b** involves the minimization of a quadratic form called the cost (or merit) function, that consists of two terms: the usual least-squares term that quantifies the distance between measurement and theory, and a second term that represents regularization (in the case of Tikhonov-Twomey smoothing) or a priori information (in optimal estimation). This second term is necessary since our model matrix **K** is notoriously ill-conditionned. In the case of optimal estimation for example [Rodgers, 2000], our merit function to be minimized is:

$$M = \left[\beta - \mathbf{K}\cdot e^{\mathbf{b}/\sigma_b}\right]^T \mathbf{S}_\beta^{-1}\left[\beta - \mathbf{K}\cdot e^{\mathbf{b}/\sigma_b}\right] + \left[\mathbf{b}-\mathbf{b}_\mathbf{a}\right]^T \mathbf{S}_\mathbf{a}^{-1}\left[\mathbf{b}-\mathbf{b}_\mathbf{a}\right] \quad (116)$$

with $\mathbf{b_a}$ and $\mathbf{S_a}$ an a priori solution and its covariance matrix respectively.

3.2.6.5.2 Regularization

While the optimal estimation approach generally gives good results for the retrieval of species, it will pose problems in our case since aerosol PSDs have a large natural variability, and it is therefore difficult to define an a priori estimate. It is thus better not to include external a priori data, but let the measurements speak for themselves. This is why PSD retrievals are usually performed with Tikhonov regularization, which we also do here. The approach is to limit small-scale oscillations within certain bounds. Let **L** be a *first-difference operator* (See details in appendix A). In Eqn. (116), we then put:

$$\mathbf{b}-\mathbf{b}_\mathbf{a} = (\mathbf{I}-\mathbf{D})\mathbf{b} \overset{def}{=} \mathbf{L}\mathbf{b} \quad (117)$$

We furthermore assume the a priori covariance matrix $\mathbf{S_a}$ to be diagonal, with a constant value (the variance) on this diagonal:

$$\mathbf{S}_\mathbf{a} = \sigma_a^2\mathbf{I} \quad (118)$$

Then the merit function reads:

$$M = \left[\beta - \mathbf{K}\cdot e^{\mathbf{b}/\sigma_b}\right]^T \mathbf{S}_\beta^{-1}\left[\beta - \mathbf{K}\cdot e^{\mathbf{b}/\sigma_b}\right] + \mu^2\mathbf{b}^T\mathbf{L}^T\mathbf{L}\mathbf{b} \quad (119)$$

with $\mu^2 = 1/\sigma_a^2$ the regularization parameter that tunes the smoothness of the solution.

Notice from the definition of the quantity **b** (Eqn. (104)) that an absolute deviation on $b(u)$ corresponds to a *relative* deviation on $f_u(u)$. This is an extremely interesting property from a retrieval point of view, since the choice of regularization parameter is now very intuitive. For example, if we allow oscillations of 100 % on f_u ($\sigma_{f_u}/f_u = 1$), then we need to specify a value of $\sigma_a = \sigma_b$, or $\mu = 1/\sigma_b$.

3.2.6.5.3 Minimization with Levenberg-Marquardt

Minimization of the merit function is equivalent to solving the nonlinear equation:

(120) $$\begin{bmatrix} \mathbf{S}_\beta^{-1/2}\beta \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{S}_\beta^{-1/2}\mathbf{K}\cdot e^{(\mathbf{b}/\sigma_b)} \\ \mu\mathbf{L} \end{bmatrix}\mathbf{b}$$

The minimization is performed by a Levenberg-Marquardt algorithm, which needs the so-called objective function as input:

(121) $$\mathbf{F} = \begin{bmatrix} \mathbf{S}_\beta^{-1/2}(\beta - \mathbf{K}\cdot e^{(\mathbf{b}/\sigma_b)})\mathbf{b} \\ -\mu\mathbf{L}\mathbf{b} \end{bmatrix}$$

The numerical values of this function are implicitely squared and summed by the LM algorithm in order to get the merit function *M*. Furthermore, the *Jacobian* of this function is also passed to the LM algorithm.

For the upper part of the objective function:

(122) $$\mathbf{J}_{ij} = \frac{\partial F_i}{\partial b_j} = -\frac{1}{\sigma_b}\sum_k \left[\mathbf{S}_\beta^{-1/2}\right]_{ik}\mathbf{K}_{kj}e^{(b_j/\sigma_b)}$$

For the lower part:

(123) $$\mathbf{J}_{ij} = -\mu\mathbf{L}_{ij}$$

and thus, the Jacobian matrix reads:

(124) $$\mathbf{J} = -\begin{bmatrix} \frac{1}{\sigma_b}\mathbf{S}_\beta^{-1/2}\mathbf{K}\cdot \mathrm{diag}(e^{(\mathbf{b}/\sigma_b)}) \\ \mu\mathbf{L} \end{bmatrix}$$

At the final iteration, the LM algorithm returns the solution at the minimum, and the Jacobian at this point. The solution covariance matrix is given by:

(125) $$\mathbf{S}_\mathbf{b}^{\mathbf{sol}} = (\mathbf{J}^T\mathbf{J})^{-1}$$

This operation can be troublesome, if the matrix between brackets is ill-conditioned. It is better to calculate the Cholesky factor **Q** first:

(126) $$\mathbf{Q} = \mathrm{chol}\left(\mathbf{J}^T\mathbf{J}\right)$$

(127) and thus: $$\mathbf{Q}^T\mathbf{Q} = \mathbf{J}^T\mathbf{J}$$

(128) then: $$\mathbf{S}_{\mathbf{b}}^{\mathbf{sol}} = (\mathbf{Q}^T\mathbf{Q})^{-1} = \mathbf{Q}^{-1}\mathbf{Q}^{-T}$$

It is numerically much better to invert this 'square root' **Q** of $\mathbf{J}^T\mathbf{J}$. However, as mentioned in Section 3.2.4.3, the Cholesky decomposition is only possible for hermitian positive-definite matrixes. This condition is in principle fulfilled by the covariance matrix, but for real data in the case of very noisy measurements, it can happen that the positive-definite character of the full covariance matrix is not fullfilled. In that case, the use of Cholesky decomposition fails and the retrieval cannot be performed successfully.

3.2.6.6 Final transformations

Finally we have to convert the solution $\mathbf{b}^{\mathbf{sol}}$ and covariance matrix $\mathbf{S}^{\mathbf{sol,b}}$ back to the desired form. First we convert back to the PSD on the *u*-grid, using Eqn. (104):

(129) $$\mathbf{f}_{\mathbf{u}}^{\mathbf{sol}} = f_{u0}e^{(\mathbf{b}^{\mathbf{sol}} - b_0)/\sigma_b}$$

(130) $$\mathbf{S}_{\mathbf{fu},ij}^{\mathbf{sol}} = \frac{1}{\sigma_b^2}\left[\mathbf{f}_{\mathbf{u},i}^{\mathbf{sol}} \cdot \mathbf{S}_{\mathbf{b},ij}^{\mathbf{sol}} \cdot \mathbf{f}_{\mathbf{u},j}^{\mathbf{sol}}\right]$$

This last expression is not correct in general; it is an approximation that is only valid when the errors on the components of the vector $\mathbf{b}^{\mathbf{sol}}$ are much smaller than the parameter σ_b. Large errors do not transform properly with this formula, and this should be taken into account while doing data analysis of the retrievals[3]. It is also advisable to store the original data $\mathbf{b}^{\mathbf{sol}}$ and $\mathbf{S}^{\mathbf{sol,b}}$ in the final product files.

Finally, we transform back to the PSD $\mathbf{f}^{\mathbf{sol,a}}$ with Eqn. (102). For the *i*-th component:

(131) $$\mathbf{f}_{\mathbf{a,i}}^{\mathbf{sol}} = (\sigma_u / a_i) \cdot \mathbf{f}_{\mathbf{u,i}}^{\mathbf{sol}}$$

(132) and $$\mathbf{S}_{\mathbf{fa,ij}}^{\mathbf{sol}} = (\sigma_u / a_i) \cdot \mathbf{S}_{\mathbf{fu,ij}}^{\mathbf{sol}} \cdot (\sigma_u / a_j)$$

3.2.6.7 PSD-derived quantities

A PSD contains basically all particle-related information that can be obtained from a remote sensing instrument such as GOMOS. Nevertheless, it is sometimes useful to investigate data that are derived from this PSD. It is for example sometimes more interesting to study the total particle number density of the PSD, or the average particle radius etc. In a context of comparison with

[3] For example, error-weighted means of ensembles of retrievals will contain a bias if we take the mean of the transformed quantity f_u. It is better to do statistics on the original quantity $\mathbf{b}^{\mathbf{sol}}$.

aerosol properties provided by climate models, the availability of such derived quantities is also an added value to assess the performance of both the retrieval algorithm and the climate model. In our methodology, these quantities should be calculated using the transformations that we introduced for the retrieval. Also, they can be calculated in a fairly straightforward way with the use of *moments*.

3.2.6.7.1 Moments of the PSD

The *n*-th moment of the PSD $f_a(a)$ is given by:

(133) $$M^{(n)} = \int_{a=0}^{+\infty} a^n f_a(a)\,\mathrm{d}a = \int_{u=-\infty}^{+\infty} a_0^n e^{n(u-u_0)/\sigma_u} f_u(u)\,\mathrm{d}u$$

As before, we discretize the integral and we assume f_u to be constant within one *u*-bin:

(134) $$\boldsymbol{M^{(n)}} = \sum_{i=-\infty}^{+\infty} a_0^n f_u(u_i) \int_{u_i-\Delta u/2}^{u_i+\Delta u/2} e^{n(u-u0)/\sigma_u}\,\mathrm{d}u$$

(135) $$= \sum_{i=-\infty}^{+\infty} a_0^n e^{n(u_i-u0)/\sigma_u} f_u(u_i) \frac{2\sigma_u}{n} \sinh\left(\frac{n\Delta u}{2\sigma_u}\right)$$

which can also be written as function of the particle radius a_i:

(136) $$\boldsymbol{M^{(n)}} = \sum_{i=-\infty}^{+\infty} a_i^n f_u(u_i) \frac{2\sigma_u}{n} \sinh\left(\frac{n\Delta u}{2\sigma_u}\right)$$

(137) $$= \left[\frac{\sinh\left(\frac{n\Delta u}{2\sigma_u}\right)}{\frac{n\Delta u}{2\sigma_u}}\right] \Delta u \sum_{i=-\infty}^{+\infty} a_i^n f_u(u_i)$$

These results can be put in vector notation if we define a column vector **q** with coefficients:

(138) $$q_i = \left[\frac{\sinh\left(\frac{n\Delta u}{2\sigma_u}\right)}{\frac{n\Delta u}{2\sigma_u}}\right] \Delta u a_i^n$$

Then the *n*-th moment equals:

(139) $$M^{(n)} = \mathbf{q}^{\mathrm{T}}\mathbf{f}_{\mathbf{u,sol}}$$

and has an associated variance:

(140) $$S_{M^{(n)}} = \mathbf{q}^{\mathrm{T}}\mathbf{S}_{\mathbf{fu,sol}}\mathbf{q}$$

3.2.6.7.2 Microphysical quantities derived from the PSD

With the use of a few calculated moments, we can obtain useful scalar quantities that characterize the retrieved PSD. The associated variances are calculated with the standard error propagation calculations.

3.2.6.7.2.1 Total particle number density [cm^{-3}]

The total number of particles per volume of air is simply the integrated PSD:

(141) $$N_{tot} = \int_{a=0}^{+\infty} f_a(a)\mathrm{d}a = M^{(0)}$$

(142) with variance: $$S_{N_{tot}} = S_{M^{(0)}}$$

3.2.6.7.2.2 Mean radius [um]

The average radius of the particle distribution is calculated as follows:

(143) $$a_m = \left(\int_{a=0}^{+\infty} a f_a(a)\mathrm{d}a\right) \Big/ \left(\int_{a=0}^{+\infty} f_a(a)\mathrm{d}a\right) = M^{(1)} / M^{(0)}$$

(144) with variance: $$S_{a_m} = \left[\frac{M^{(1)}}{M^{(0)2}}\right]^2 S_{M^{(0)}} + \left[\frac{1}{M^{(0)2}}\right] S_{M^{(1)}}$$

3.2.6.7.2.3 PSD variance [um^2]

The average squared difference between the particle radius and the mean radius equals:

(145) $$\sigma^2 = \left(\int_{a=0}^{+\infty} (a - a_m)^2 f_a(a)\mathrm{d}a\right) \Big/ \left(\int_{a=0}^{+\infty} f_a(a)\mathrm{d}a\right)$$

which, after some algebraic manipulation, leads to:

(146)
$$\sigma^2 = \frac{\int_{a=0}^{+\infty} a^2 f_a(a)\,\mathrm{d}a}{\int_{a=0}^{+\infty} f_a(a)\,\mathrm{d}a} - \left[\frac{\int_{a=0}^{+\infty} a f_a(a)\,\mathrm{d}a}{\int_{a=0}^{+\infty} f_a(a)\,\mathrm{d}a}\right]^2 = \frac{M^{(2)}}{M^{(0)}} = \left(\frac{M^{(1)}}{M^{(0)}}\right)^2$$

(147) with variance:
$$S_{\sigma^2} = \left[\frac{M^{(2)}}{M^{(0)2}}\right]^2 S_{M^{(0)}} + \left[\frac{1}{M^{(0)2}}\right] S_{M^{(2)}} + 4 a_m{}^2\, S_{a_m}$$

3.2.6.7.2.4 Surface area density [um^2 cm^{-3}]

The total surface area of the (spherical) aerosol particles per volume can be calculated by integration of the individual particle surface areas within the distribution:

(148)
$$S = \int_{a=0}^{+\infty} 4\pi a^2 f_a(a)\,\mathrm{d}a = 4\pi M^{(2)}$$

(149) with variance:
$$S_S = (4\pi)^2\, S_{M^{(2)}}$$

3.2.6.7.2.5 Volume density [um^3 cm^{-3}]

The total volume of the particles per volume of air that contains these particles equals the integral of all individual volumes:

(150)
$$V = \int_{a=0}^{+\infty} \frac{4}{3}\pi a^3 f_a(a)\,\mathrm{d}a = \frac{4}{3}\pi M^{(3)}$$

(151) with variance:
$$S_V = (\frac{4}{3}\pi)^2\, S_{M^{(3)}}$$

3.2.6.7.2.6 Effective radius [um]

The radius of a sphere with volume *v* and surface area *s* equals *a*=3*V*/*S*. In a similar way, for a PSD one can define an effective radius from the volume density and the surface area density:

(152)
$$a_{eff} = 3\frac{V}{S} = \frac{M^{(3)}}{M^{(2)}}$$

(153) with variance:
$$S_{a_{eff}} = \left[\frac{M^{(3)}}{M^{(2)2}}\right]^2 S_{M^{(2)}} + \left[\frac{1}{M^{(2)2}}\right] S_{M^{(3)}}$$

All these quantities are very easy to calculate numerically from a retrieved PSD $f_u(u)$ with Eqns. (139) and (140).

3.2.6.7.3 Partial moments

All derived quantities above were expressed as functions of the moments $M^{(n)}$ calculated in Eqn. (137). Notice that the summation runs from $-\infty$ to $+\infty$. For actual retrievals, the summation will of course run from the first to the last element of the retrieval grid.

However, it is sometimes useful to calculate moments within an even more limited range. Such an occasion occurs when we want to compare the quantities N_{tot}, a_m, σ^2, *S*, *V*, and a_{eff} with the ones derived from other experiments (for example optical counters) that have a different sensitivity with respect to particle size. See [Bingen et al., 2004] for some good examples of the associated difficulties. It is then better to perform the comparison in a radius range where both instruments are able to detect particles. This means that we have to use the *partial PSD moments*:

$$M^{(n)}_{[a_{\min},a_{\max}]} = \left[\frac{\sinh\left(\frac{n\Delta u}{2\sigma_u}\right)}{\frac{n\Delta u}{2\sigma_u}}\right]\Delta u \sum_{i=i_{\min}}^{i_{\max}} a_i^n f_u(u_i) \quad (154)$$

with $a_{\min} = a_0 \exp(u_{\min}/u_0)$, $u_{\min} = u_{i_{\min}} = u_0 + i_{\min}\Delta u$ and identical relations for $a_{\max}$.

3.2.6.7.4 Radiative properties derived from the PSD

Using the solution $\mathbf{b}^{\mathbf{sol}}$ found for the PSD (see Section 3.2.6.6) and the extinction efficiency Q_{ext} computed by the Mie code, we can rebuilt the corresponding extinction coefficient β_{ext} at any wavelength λ using Eqn. (110), giving with (111), (112), and (113):

$$\beta_{ext}^{rebuilt}(\lambda) = \sum_{i=-\infty}^{+\infty} \pi a_0^2 f_{u0} e^{(\mathbf{b}^{\mathbf{sol}}{}_i - b_0)/\sigma_b} \cdot \tilde{Q}_{ext,i}(\lambda) \quad (155)$$

This can be useful in particular for diagnostic purposes.

3.2.6.7.4.1 Scattering coefficient and absorption coefficient [cm^{-1}]

Similarly, using the inferred aerosol composition (see Section 3.2.6.3), we can compute the corresponding scattering or absorbing efficiency $Q_{scat}(\lambda, a_i = a_0.e^{(u_i-u_0)/\sigma_u}) = Q_{scat,i}(\lambda)$ and $Q_{abs}(\lambda, a_i = a_0.e^{(u_i-u_0)/\sigma_u}) = Q_{abs,i}(\lambda)$, and rebuild the scattering coefficient β_{scat} and absorbing coefficient β_{abs} corresponding to this aerosol composition and to the retrieved PSD:

(156) $$\beta_{scat}(\lambda) = \sum_{i=-\infty}^{+\infty} \pi a_0^2 f_{u0} e^{(\mathbf{b}^{\mathbf{sol}}{}_i - b_0)/\sigma_b} \cdot \widetilde{Q}_{scat,i}(\lambda)$$

(157) $$\beta_{abs}(\lambda) = \sum_{i=-\infty}^{+\infty} \pi a_0^2 f_{u0} e^{(\mathbf{b}^{\mathbf{sol}}{}_i - b_0)/\sigma_b} \cdot \widetilde{Q}_{abs,i}(\lambda)$$

3.2.6.7.4.2 Asymmetry factor [dimensionless]

Furthermore, we can calculate an estimate of the asymmetry parameter $\overline{\cos\theta}$ which quantifies the averaged spread of the light scattered by the particle population. The Mie scattering theory provides an analytical expression of the product $\overline{\cos\theta} \cdot Q_{scat}$ for a given wavelength λ and particle radius *a* [Bohren and Hofman, 1993, van de Hulst, 1981], so that the overall expression of the asymmetry parameter for the whole PSD, $\overline{\cos\theta}_{PSD}$ can be computed for the retrieved PSD using:

(158) $$\overline{\cos\theta}_{PSD}(\lambda) \cdot \beta_{scat}(\lambda) = \sum_{i=-\infty}^{+\infty} \pi a_0^2 f_{u0} e^{(\mathbf{b}^{\mathbf{sol}}{}_i - b_0)/\sigma_b} \cdot \left[\overline{\cos\theta} \cdot Q_{scat}\right]_i(\lambda)$$

3.2.6.7.4.3 Single-scattering albedo [dimensionless]

In a similar way, the single-scattering albedo (SSA) $\overline{\omega_0}$ quantifying the proportion of light scattered by particles at a wavelength λ, is expressed for a single particle of radius *a* by

(159) $$1 - \overline{\omega_0}(\lambda, a) \cdot = \frac{Q_{abs}(\lambda, a)}{Q_{ext}(\lambda, a)}$$

which implies:

(160) $$Q_{ext}(\lambda, a) - Q_{abs}(\lambda, a) = Q_{scat}(\lambda, a) = Q_{ext}(\lambda, a) \cdot \overline{\omega_0}(\lambda, a)$$

(161) i.e.: $$1 - \overline{\omega_0}(\lambda, a) \cdot = \frac{Q_{abs}(\lambda, a)}{Q_{ext}(\lambda, a)}$$

This relation provides an estimate of the overall SSA for the retrieved PSD, $\overline{\omega_{0,PSD}}$ through

(162) $$\beta_{scat}(\lambda) = \beta_{ext}(\lambda) \cdot \overline{\omega_{0,PSD}} = \sum_{i=-\infty}^{+\infty} \pi a_0^2 f_{u0} e^{(\mathbf{b}^{\mathbf{sol}}{}_i - b_0)/\sigma_b} \cdot \overline{\omega_{0,i}} \cdot \widetilde{Q}_{ext,i}(\lambda)$$

So that:

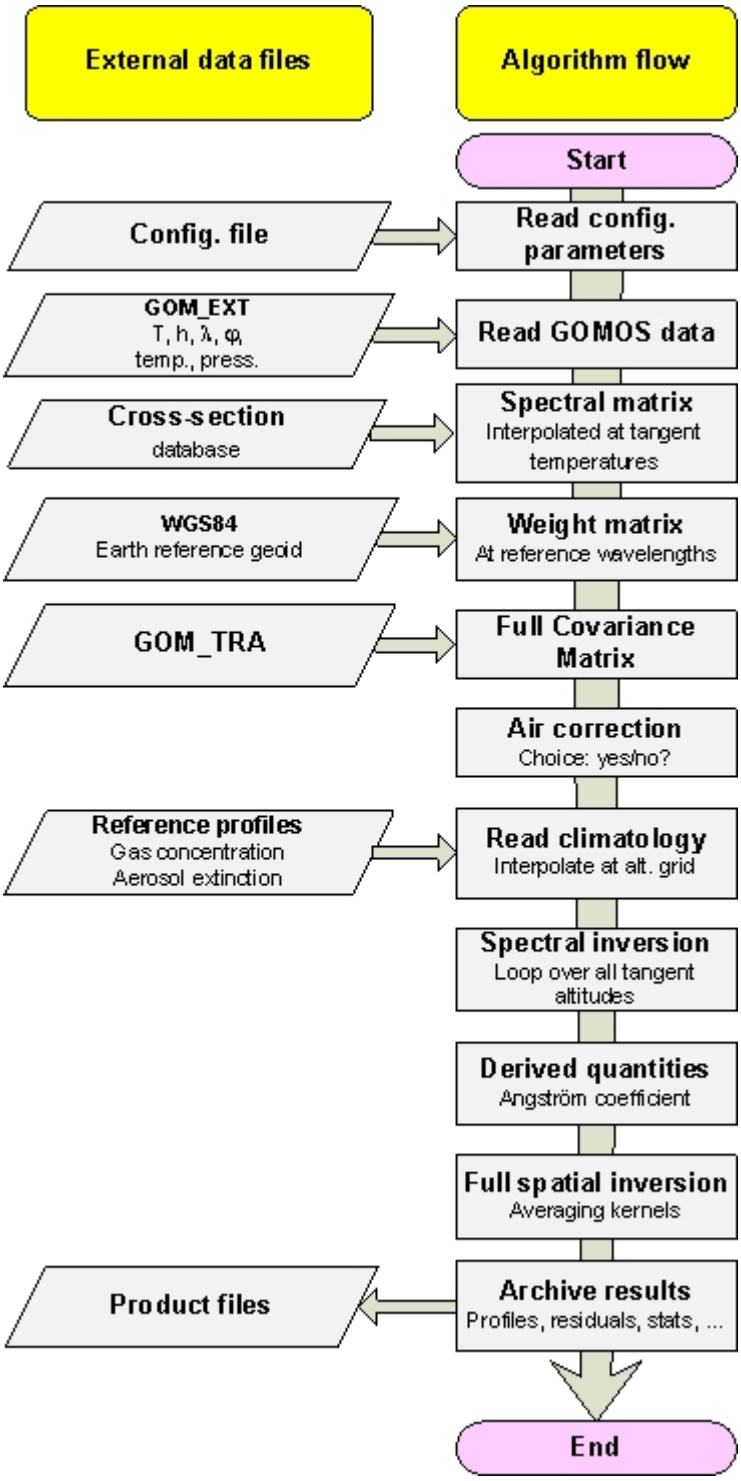
(163) $$\overline{\omega_{0,PSD}} = \left(\sum_{i=-\infty}^{+\infty} \pi a_0^2 f_{u0} e^{(\mathbf{b}^{\mathbf{sol}}{}_i - b_0)/\sigma_b} \cdot \overline{\omega_{0,i}} \cdot \widetilde{Q}_{ext,i}(\lambda) \right) / \beta_{ext}(\lambda)$$

3.2.7 Flowcharts

3.2.7.1 Spectral/full spatial inversion

A flowchart for the spectral/full spatial algorithm is shown in Figure 10.

Figure 10 - Flow chart representing the general structure of the algorithm. The left column represents external data files, the right column the algorithm flow.

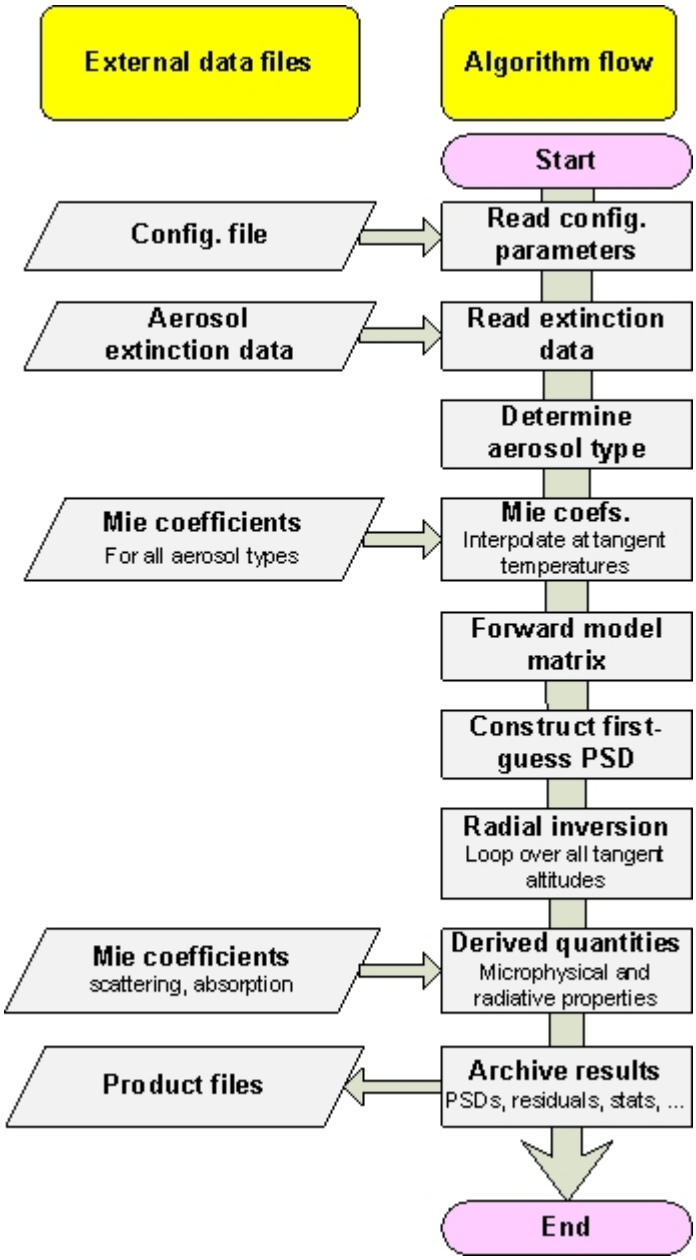


The code starts by reading configuration parameters (Levenberg-Marquardt iteration parameters, spectral pixel selections, a switch to select air correction or not, a switch to select the aerosol spectral model, etc.). Then, all necessary GOMOS data are read from the GOM_EXT file (transmittance measurements and associated variances, tangent altitude, longitude, latitude, temperature and pressure profiles etc.). The spectral matrix (Eqn. (58)) is constructed at all tangent altitudes by linearly interpolating within tabulated values at different temperatures, which are stored in a cross-section database. Then the weight matrix (Eqn. (48) and (49)) is calculated at the reference wavelength which can be found in the GOM_EXT files. Depending on the optional settings of the retrieval, the full covariance matrix with modelisation errors included may be calculated, and the transmittance spectra may be corrected for air using the meteorological density profile. Then, reference profiles for all species are interpolated from stored profiles, to be used as first guess in the spectral inversion, which is performed subsequently, for each tangent altitude separately. The full spatial inversion is then performed in one step. Finally, retrievals, chi-squared statistics, residuals, iteration statistics, averaging kernels, etc. are stored in an output file.

3.2.7.2 Flowchard of the radial inversion

A graphical representation of the radial inversion algorithm is shown in Figure 11.
The code starts by reading configuration parameters (Levenberg-Marquardt iteration settings, particle radius grid parameters etc.), and then reads the GOMOS aerosol extinction data at all considered wavelengths and tangent altitudes. The aerosol type is determined as a function of the temperature, altitude and geolocation values, and/or the extinction spectral dependence (See discussion in Section 3.2.6.3). Then the Mie coefficients (Eqn. (109)) are interpolated at the tangent temperatures using precalculated tables for different temperatures and species. The forward model matrix is calculated (Eqn. (115)). A reference PSD is read to be used as a first guess in the following Levenberg-Marquardt minimization. One PSD for each tangent altitude is retrieved. Finally, PSD-derived quantities (microphysical and radiative parameters) can be calculated: tangent altitude profiles for total particle number density, effective radius etc. and the outputs are written to a file.

Figure 12 - Flow chart representing the general structure of the radial inversion algorithm. The left column represents external data files, the right column the algorithm flow.



3.3 Development of Level-3 gridded products

3.3.1 General approach

The aim of this stratospheric activity of Aerosol_cci is to provide, in particular to the Climate Modeling Community, time series of stratospheric aerosol parameters relevant for the characterization of the target aerosols properties in the models. A detailed description of this activity, including the algorithm development, validation aspects, and the use of the Level-3 products in Chemistry-Climate Modelling, is the object of a paper [Bingen et al., 2017].

To achieve the best quality, and in view of the performance dependence of AerGOM measurements on star properties (temperature and magnitude) and measuring conditions (e.g. related to obliquity and solar zenith angle), the solution with the greatest ease of use for climate applications is to provide gridded products resolving latitude, longitude, altitude, and time.
The relevance of the final aerosol products depends on several aspects which were addressed in the development of the CCI-AerGOM aerosol records:

- *Aerosol types*: The retrieved aerosols can have different origins: liquid sulfate aerosols, PSCs and cirrus clouds, but also potential aerosol modes with another composition, such as meteoric smoke particles, soot, etc. Based on criteria discussed in Section 3.2.6.3, separate fields are provided in the CCI-AerGOM record for the aerosol compounds which could be identified, especially for the PSD and related properties. For aerosol extinction, a set of fields including the total extinction and the extinction for each individual aerosol fraction are provided.
- *Grid resolution*: The choice of grid resolution must be made in a coherent way for the three spatial dimensions and time in order to adequately render dynamical spatial patterns such as the propagation of volcanic plumes. This coherence requirement was investigated and led to a better discrimination of volcanic signatures in the extinction time series.
- *Event selection*: The variable measurement quality depending on star properties and measuring conditions requires an appropriate selection of the events before processing of the time series. The choice of event selection criteria is discussed taking into account all critical aspects of the Level 3 processing.

The present chapter describes all these aspects and issues addressed during the derivation of the CCI/GOMOS Level 3 products from the AerGOM Level 2 vertical profiles, and the methodology chosen to process the gridded aerosol products.

3.3.2 Level-3 size-related parameters

The CCI, version 3.00 dataset is the first CCI-GOMOS dataset including PSD-derived products. In this version, the PSD was derived using the AerGOM radial inversion algorithm (See Section 3.2.6) applied on each sets of AerGOM extinction profiles at 355, 440, 470, 550, and 750 nm (Level 2). A set of derived products (total number density, effective radius, surface area density, volume density, asymmetry factor, and single scattering albedo) was then applied to the retrieved Level 2 PSD, and then binned to produce Level-3 CCI-GOMOS products.
The investigation of this dataset showed in many cases poor performances, due, for a large part, to the unrealistic PSD derived from occultation measurements using stars with poor properties (cold and dim stars). In that case, the SNR is particularly reduced and the retrieved extinction profiles are of bad quality. As a consequence, the additionnal inversion of the set of extinction coefficients amplifies the noise even more, and the resulting PSD estimates present a very large uncertainty, making the result useless.
In order to address this issue, the order of the processing was inverted in the present CCI-GOMOS version (CCI_4.00). Consequently, the set of extinction coefficients, now provided as an averaging of all Level 2 extinction profiles in the corresponding grid cell, is of (much) better quality, and better suited for the radial inversion. The PSD is then retrieved directy from the Level-3 gridded extinction profiles, and is used to derive all gridded PSD-related products for the CCI-GOMOS dataset.

In the same way, the Angström exponent is, in the current CCI_4.00 version, retrieved from the binned extinction profiles, using a linear regression applied to the extinction coefficients at 355, 440, 470, 550, and 750 nm at each altitude separately.

3.3.3 Analysis of the dataset

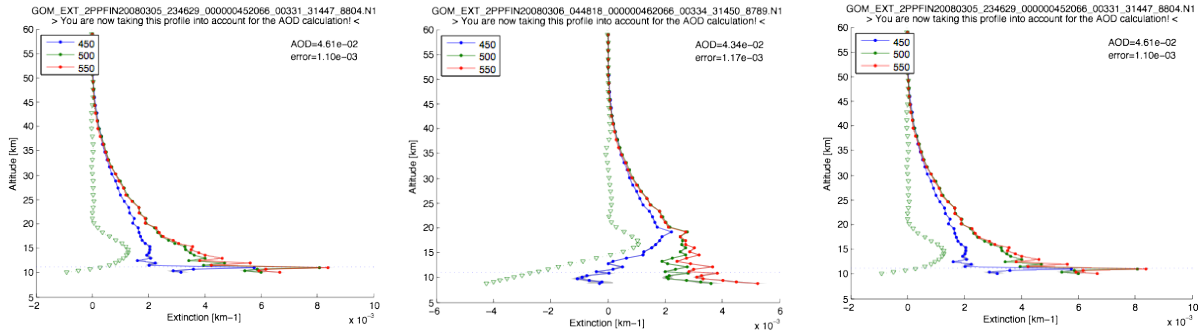
3.3.3.1 PSC identification

As explained in Sections 3.1.4 and 3.2.6.3, the 'detection' of PSC within the AerGOM data is done by comparing the theoretical Nitric Acid Trihydrate (NAT, i.e. PSC type I particles) condensation temperature and the GOMOS-associated ECMWF temperature profile. If $T_{ECMWF}(p) < T_{NAT}(p)$ between 15 and 25 km, a PSC is assumed to be present. Whenever one or more PSC are detected during an occultation, the profile is flagged as containing a PSC.
The NAT condensation temperature analytic formula is taken from [Hanson and Mauersberger, 1988]. The mixing ratios of HNO_3 and H_2O in the polar stratosphere are assumed to be 9.5 ppb and 5 ppm respectively.

3.3.3.2 Anomalous profiles identification

Initial comparisons of operational aerosol profiles (IPF) and AerGOM inverted stratospheric aerosol profiles showed that, under certain conditions, the AerGOM profiles are unrealistic, with large extinction value, especially obvious at altitudes between 30 and 45 km. Figure 13 shows 3 examples of such profiles compared with the operational retrieval.

Figure 13 - Examples of anomalous stratospheric aerosol profiles with large extinction values between 30 and ~45 km at 450, 500 and 550nm. The dotted green curve (triangles) shows the IPF retrieval result at 500nm for the same profile for comparison.



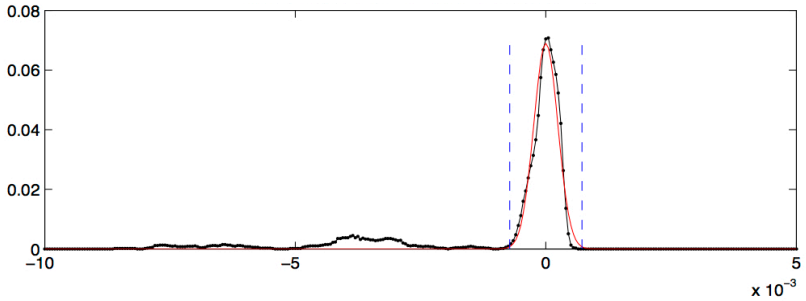
In order to identify these anomalous profiles, a technique has been devised which is exploiting the high extinction values of these profiles at high altitude.
The technique needs a collection of profiles, typically upwards of 200. For each profile, the upper stratospheric aerosol optical depth, AOD*, is calculated as follows

$$AOD^* = \int_{30km}^{45km} \beta\left(z, \lambda = 500nm\right) dz \qquad (164)$$

The distribution of the AOD* values is then calculated, and a Gaussian curve (with median u and standard deviation s) is fitted centered on the AOD* median, which is very close to zero. All profiles

with AOD* values outside the interval [u-5s, u+5s] are identified as anomalous profiles and discarded (see Figure 14).

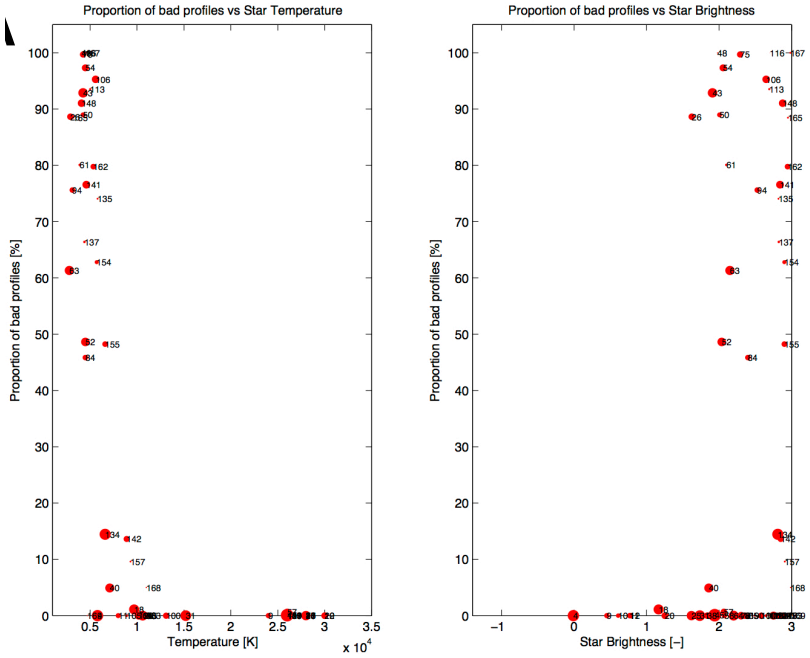
Figure 14 - Example of the frequency distribution of AOD* (black). The red solid curve shows the Gaussian fit and the blue dotted lines delimit the range of acceptable values.



By looking at the proportion of anomalous profiles per star, it became clear that the problems mostly arise for occultations performed with cold (Tstar < 5 000 K) and dim stars (Mag ≥ 2), as can be seen in Figure 15.

The cause of this behaviour seems to be due to the low signal-to-noise ratio (SNR) of the transmittance at shorter wavelengths, which can lead to an erroneous attribution of the extinction to the different trace gases and the aerosols, and to a poor retrieval of the O_3, NO_2, NO_3, and aerosol profiles. The problem was solved by changing the first-guess for the Levenberg-Marquardt optimization scheme used in the spectral inversion (see Section 3.2.4.3), and by choosing estimates closer to the final solution to avoid convergence to unrealistic profiles.

Figure 15 - Proportion of anomalous profiles for a given star temperature (left) and magnitude (right).



3.3.3.3 Tropopause Height

In this work, the stratospheric AOD is calculated as:

$$AOD_{strat}(\lambda) = \int_{TH+2km}^{50km} \beta(z,\lambda)\,dz \qquad (165)$$

Therefore, the tropopause height (TH) is essential to calculate the stratospheric contribution to the AOD. There exist different ways to determine the tropopause level. The WMO defines the tropopause as "the lowest level at which the lapse rate decreases to 2 °C/km or less, provided that the average lapse rate between this level and all higher levels within 2 km does not exceed 2 °C/km." This is actually the definition of the thermal tropopause. For the purpose of our work, we use a slightly improved technique to determine the tropopause level, as explained in the appendix of [Hoinka et al., 1998].

At high latitudes however, it may happen that no thermal gradient is present. In this context, we must rely on dynamical properties of the transition between troposphere and stratosphere. For this work, the dynamical tropopause is taken as the level at which the potential vorticity is ±2.5 potential vorticity units.

The global tropopause is taken as a combination of both definition. For latitudes polewards of 30°, the dynamical tropoause is assumed to better represent the transition between both atmospheric layers. In the tropics (|lat| < 20°), the thermal definition of the tropopause is assumed. In the transition region between the two regimes, both criteria are used and weighted with the distance to the regime boundaries.

Global tropopause maps are produced with a 1°x1° resolution using ECMWF reanalysis data for each day at 12.00. The tropopause level for a given occultation is then interpolated in space and time from these maps.

3.3.4 Implementation

This section summarizes the main aspects of the implementation of the Level 3 processing, and details the issues and decisions made for the development of these gridded products.

3.3.4.1 Event selection criteria

The quality of the aerosol retrieval depends critically on the quality of the signal entering the GOMOS instrument. The primary factors influencing the measurement quality are the star magnitude, star temperature and the solar zenith angle.

3.3.4.1.1 Star magnitude

The star magnitude characterizes the bright/dim character of the star. Dim stars provide a noisy light signal, while bright stars ensure better the signal-to-noise ratios.

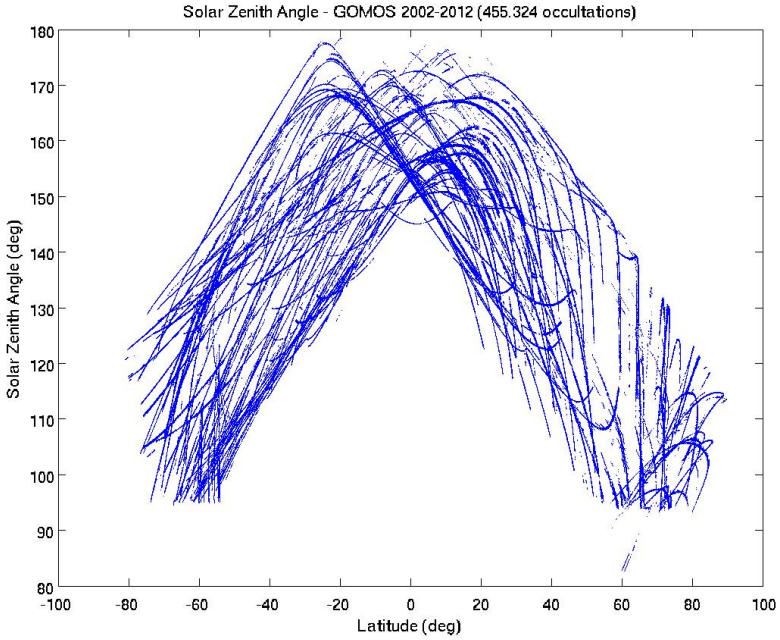
3.3.4.1.2 Star temperature

The star temperature, determining the spectral range in which the star is primarily emitting, influence the ill-posedness of the spectral inversion problem. Cool stars with temperatures ~2000K to 3000 K emit mainly toward the infrared spectrum, and the emittance in the lower part of the GOMOS spectral range (< ~400 nm) is very low. On the contrary, hot stars emittance (T > ~7500K) is relatively stable over the whole GOMOS spectral range. The lower limit acceptable for the star temperature depends on how well other retrieval issues can be solved.

3.3.4.1.3 Solar zenith angle

The solar zenith angle (SZA) determines the amount of scattered light entering the aperture of the instrument. From an SZA of 90°, the instrument does not receive direct light from the sun at all. However, stray light may affect the SNR, and it is recommended to reject all values of the SZA lower than a critical value higher than 90°. In the official ESA product [D2], a SZA of 97° is used to define bright limb conditions, and values of 110° and 120° are used as lower limit to define pure twilight conditions ans straylight conditions respectively. In the first version of the CCI-GOMOS version, a lower limit of 130° was used as selection criterion on the SZA.

Figure 16 - Latitudinal distribution of the SZA during the whole GOMOS mission. Each point represents one occultation, indicated by its latitude and corresponding SZA.



On the other hand, the range of SZA values which depends on the geolocation of the satellite, is not equally distributed over the latitude range, as shown in Figure 16. Hence, putting a too strong selection criterion on the SZA is a problem for the development of global aerosol records, and a trade-off must be found between maximal coverage and maximal data quality. If a sufficient number of events is available, the statistical significance increases and the selection criterion on the SZA can be smoothed. Table 7 gives an overview of the selection criteria used in the most prominent versions of the CCI-GOMOS aerosol dataset.

3.3.4.2 Grid definition

Although monthly zonal means are very commonly used for stratospheric aerosol climatologies to constrain models, it appeared during the project that such a choice of grid, even when using a good latitude resolution, only offers a poor description of the aerosol evolution to chemistry-climate models. Typical motions of volcanic plumes reaching the upper troposphere and lower stratosphere (UTLS) evolve quite fast over a few days immediately after the eruption, before diluting over typical time scales of weeks. By considering monthly averaged values, the dynamics of the plume evolution is lost most of the time. Volcanic signatures may go unnoticed, and much information about the

microphysical evolution of aerosols which is potentially important for atmospheric modelling, is lost. The same issue concerns potentially other transient phenomena, like local/regional dynamical patterns controlling exchanges at the UTLS, e.g. the Asian summer monsoon.
Further, in terms of horizontal displacement of air masses, longitude, latitude and time are linked through the typical velocity of the atmospheric circulation, and the dominance of the zonal direction imposed by the Earth's rotation. An optimal choice of the grid resolution in longitude, latitude and time must be made in order to describe the typical atmospheric motions in a coherent way.

Table 7 - Main specifications of the most prominent versions of the CCI-GOMOS aerosol dataset.

CCI-GOMOS and related (AerGOM) version	Selection criteria	Spatial resolution	Temporal resolution	Extinction wavelength provided (nm)
2.14 (AerGOM 2.0)	Star magnitude < 2.6 Dark limb only No SAA profiles Star temperature > 5000 K	2.5° latitude 10° longitude 1 km vertical	1 month	550
2.19 (AerGOM 2.0)	Star magnitude < 2.6 SZA > 105° No SAA profiles Star temperature > 5000 K	5° latitude 60° longitude 1 km vertical	5 days	550
3.00 (AerGOM_3.0)	Star magnitude < 3 SZA > 100° No SAA profiles Star temperature ≥ 2800 K	5° latitude 60° longitude 1 km vertical	5 days	355, 440, 470, 550, 750
4.00 (AerGOM_3.0)	Star magnitude < 3 SZA > 100° No SAA profiles Star temperature ≥ 2800 K	5° latitude 60° longitude 1 km vertical	5 days	355, 440, 470, 550, 750

Finally, the choice of grid resolution obviously depends on the amount of available measurements, and a trade-off must be made between statistically representativeness of the measurement sampling and description of fine structures in the dynamical aerosol patterns.
The evolution of the grid resolution for the various CCI-GOMOS versions (see Table 7) reflects the growing awareness of the importance of this choice on the quality and usefulness of the derived time series. The very first grid choice is the one used in version 2.14, where a fine latitude and longitude grid was chosen to get an accurate description of aerosol patterns, and a large time interval aimed at providing a good sampling. This choice shows actually a poor coherence in the description of the spatio-temporal evolution pattern. From version 2.19 and onwards, the grid was adapted to a 5-day time resolution, which provides a good description of typical volcanic plume events. This choice was also made to match the time grid used by the MIPAS team, making simultaneous use of both datasets easy. The latitude/longitude resolution was then made after

observation of typical displacement of volcanic plumes during 5-day intervals following an eruption. The coarse resolution of the longitudinal grid with respect to the latitudinal grid reflects how the effect of prevailing wind direction on atmospheric motions was taken into account. The smoothing of event selection criteria described in Section 3.3.3.1 led to a high number of events per bin, making the statistical significance more favourable.

3.3.4.3 Choice of a metric

The metric used for the binning and characterization of the uncertainty was chosen in order to address the following issues:

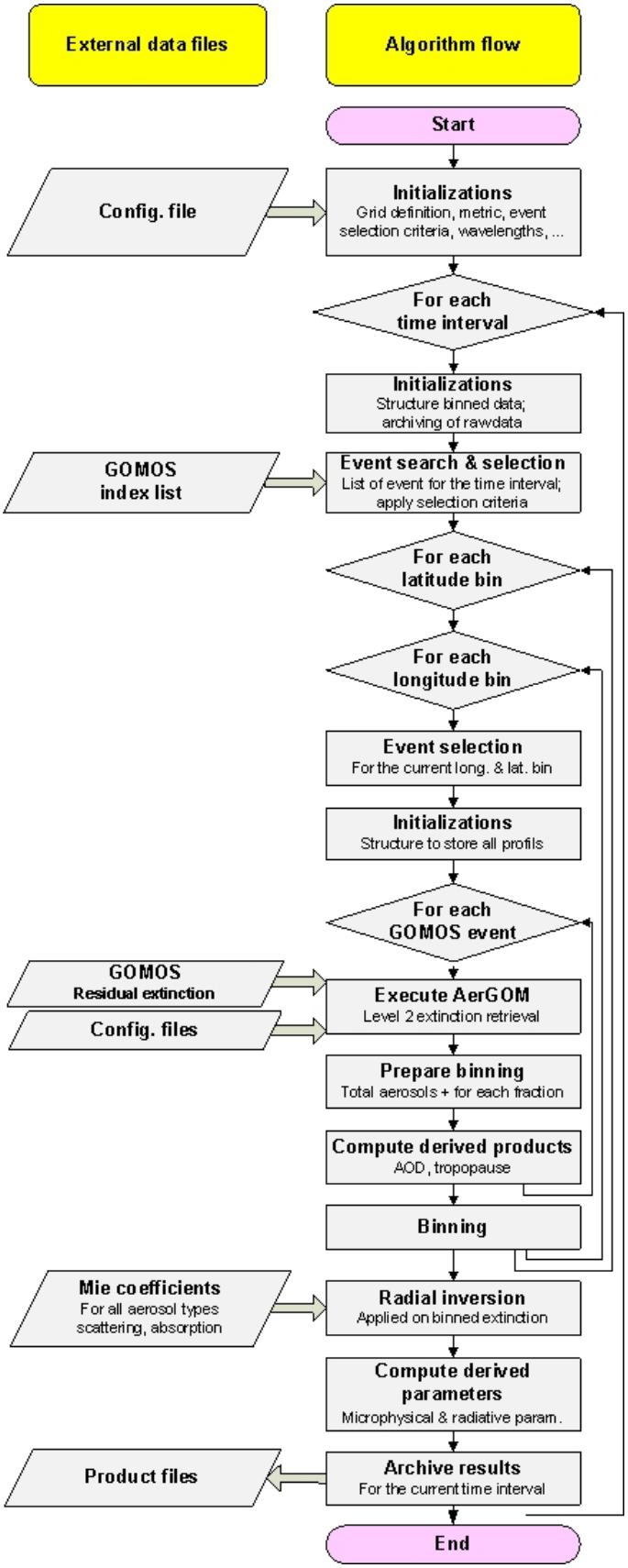
- The bin value should not be sensitive to outliers.
- The choice of metric should be suitable for bins with very few events.
- The weight of a profile in the averaging should take into account the varying profile quality.

Version 2.14 and 2.19 addressed these issues by using interquartile means for the computation of the binned mean values, i.e. the unweighted averaging of the part of the sampling between percentiles 25 and 75. This quantity has in principle the property to include information from a significant part of the sampling without suffering of the influence of outliers. The uncertainty was calculated using the interquartile mean of the uncertainty sampling, which is meant to be more realistic than the error of the mean, because it excludes more efficiently outliers corresponding to occultation with a poor SNR.
In order to address the limited coverage obtained in versions 2.14 and 2.19 by still filtering outliers, the metric was revised for version 3.00: binned values are computed as the weighted mean of all sampling values included between percentiles 10 and 90. The computation of the bin uncertainty remains unchanged (i.e. using interquartile means of the set of uncertainties).
The current version CCI_4.00 uses the same metric as version 3.00.

3.3.5 Flowchart

Figure 17 shows the flowchart of the full CCI-GOMOS processing, including the retrieval of Level 2 AerGOM data, their use to compute the Level 3 gridded dataset, and the radial inversion and computation of derived microphysical and radiative properties. Details of the Level 2 retrieval are illustrated in Figure 5 5 and Figure 6 6 and are not reproduced here.
After the necessary initializations (grid definition, choice of the metric, event selection criteria as described above in this section; choice of the wavelengths at which radiative parameters have to be provided, etc.), the processing is started producing one output file per time interval.
The processing of each time interval occurs as follows: after initializing structures for archiving the binned field and the raw Level 2 AerGOM profiles, an inventory is made of all available GOMOS measurements based on a GOMOS index list, and event selection criteria on the star parameters and SZA are applied on this inventory as discussed in Section 3.3.4.1. The resulting event list constitutes the sampling used for the processing of the CCI-GOMOS records. Based on criteria 3.2.6.3, the composition is inferred for each altitude of all Level 2, and additionnal samplings are determined for each aerosol fraction (only sulfate aerosols and PSC are considered in the current version CCI_4.00).

Figure 17 - Flow chart representing the general structure of the complete CCI-AerGOM processing.

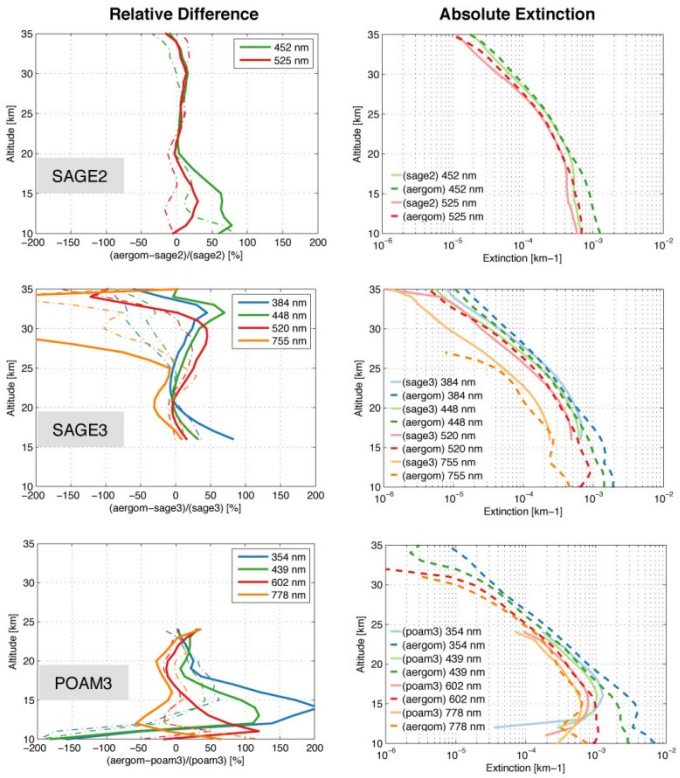


The binning process can start. After identification of all events belonging to the considered latitude and longitude bin, AerGOM is applied to retrieve all Level 2 vertical extinction profiles of interest from the GOMOS residual extinction (operational ESA product). Several profile-derived quantities are then computed, including the tropopause level, the stratospheric aerosol optical depth calculated by integration of the vertical extinction profile from 2 kilometers above the tropopause to the highest available height.
These quantities are binned using the metric described in Section 3.3.4.3.
From the Level 3 extinction profiles, the radial inversion is processed, and a suite of microphysical and radiative parameters are derived, including the total number density, , the surface aread density, volume density and effective radius, the asymmetry parameter, the single scattering albedo, and the Angström exponent.
The option exists to compute monthly zonal means from the binned values, although this product is not delivered in the CCI-GOMOS dataset.
Finally, all resulting time series are archived and stored in netCDF files.

3.4 Expected accuracy

Several validation activities have been performed for both AerGOM and the CCI-GOMOS dataset. Extended intercomparisons have been performed between AerGOM and a set of sattellite datasets. This work is the subject of a publication [Robert et al., 2016], to which we refer the reader.

Figure 18 - Intercomparison between AerGOM and different satellite dataset: SAGE 2 (upper panel), SAGE 3 (central panel), POAM 3 (lower panel). For each case, the left panel shows the relative difference for each dataset, and the right panel shows the absolute aerosol extinction vertical profiles for each common wavelength.



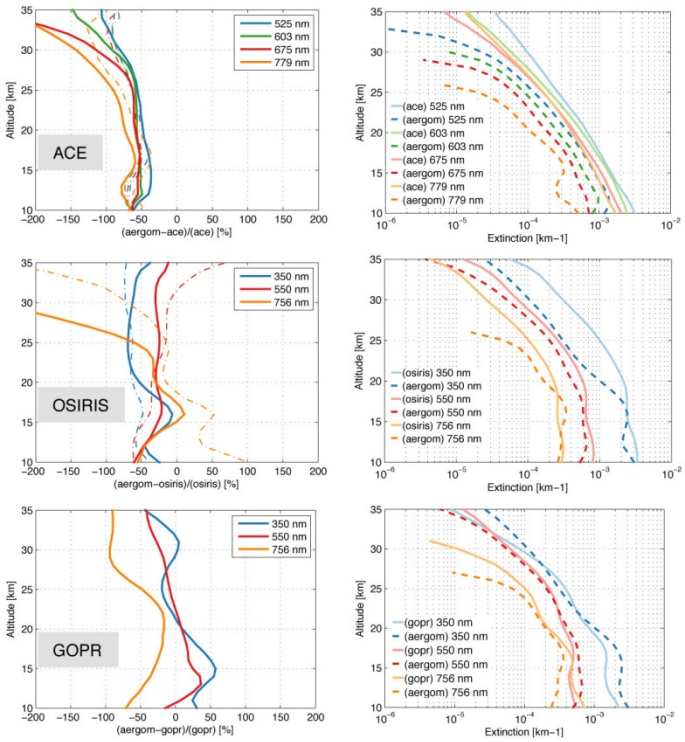
Several validation activities have been performed for both AerGOM and the CCI-GOMOS dataset.

Extended intercomparisons have been performed between AerGOM and a set of sattellite datasets. This work is the subject of a publication [Robert et al., 2016], to which we refer the reader.

Figure 18 and Figure 19 show an overview of the intercomparison work between AerGOM and datasets from SAGE II, SAGE III, MAESTRO, OSIRIS, and IPF, the operational GOMOS retrieval algorithm. A very good agreement is found between AerGOM and SAGE II in the 20-30 km range, which differs from the poorer agreement with SAGE III while both data overall agree within a few percents [Damadeo et al., 2013]. This apparent contradictory behaviour is likely to be due to coverage aspects, colocations with SAGE III being solely found at soutern mid-latitudes, while SAGE II ensures a near-global coverage. Also noticeable is the presence of a strong bias at 750 nm above 25 km height, with SAGE III and OSIRIS. However, it has to be pointed out that this problem has been improved after the revision of the absorption cross-sections for the retrieved trace gases.

Figure 19 - Same as Several validation activities have been performed for both AerGOM and the CCI-GOMOS dataset.
Extended intercomparisons have been performed between AerGOM and a set of sattellite datasets. This work is the subject of a publication [Robert et al., 2016], to which we refer the reader.

Figure 18, for the satellite experiments ACE (upper panel), OSIRIS (central panel) and IPF (lower panel).



Further, intercomparisons were made between different lidar time series and both AerGOM (Level-2 data) and GOMOS-CCI (Level-3 data). The lidar stations used for these intercomparisons are Garmisch Partenkirchen, Germany (47.5°N, 11.1°E), Mauna Loa, Hawaii, USA (19.5°N, 155.6°W and

Dumont d'Urville, Antarctica (66.7°S, 140.0°E). This work is reported in [Bingen et al., 2017]. Figure 20 shows an example of such intercomparison for an AerGOM profile with a lidar measurement above Mauna Loa. In this case, the aerosol type is identified as liquid sulfate aerosol. The agreement between both profiles is within the error bars, but it has still to be noted that such a comparison is not straightforward because the measured quantities are different in both cases: for lidar measurements, the backscatter is measured and has to be converted to extinction using an appropriate conversion factor. The value of this extinction-to-backscatter ratio (E/B) can differ from station to station and is function of the aerosol type. In the present case, the common value of 50 was used for E/B, except in the case of Garmisch-Partenkirchen for which recommended values were provided.

Figure 20 - Example of comparison between GOMOS Level 2 (red) and LIDAR profiles (blue) from Mauna Loa observatory. Shown are: the original lidar data (blue fine line); the smoothed lidar profile interpolated to the GOMOS grid (blue solid line); and the lidar data convoluted with GOMOS averaging kernel (black). The black error bars on the lidar data show the variations when using E/B of 30 sr and 70 sr instead of 50 sr.

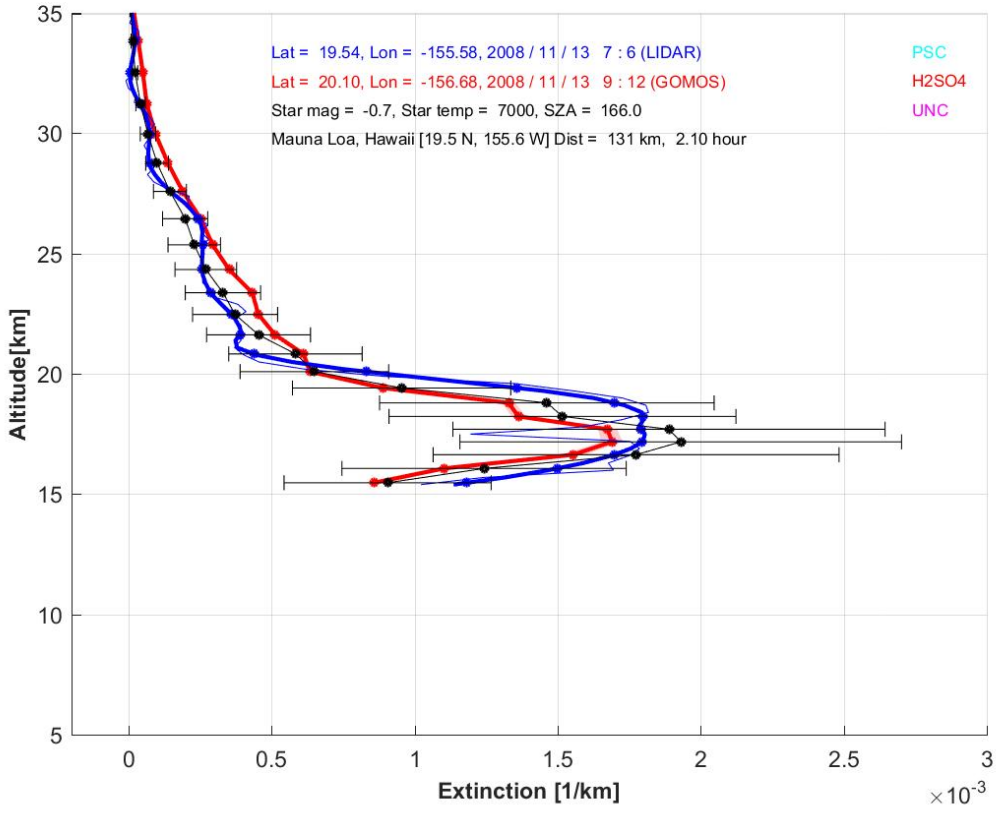


Figure 21 and Figure 22 show a summary of the intercomparisons with the four time series, respectively for the AerGOM and the CCI-GOMOS datasets. More details and discussion about the methodology and the results can be found in [Bingen et al., 2017].

Figure 21 - Comparison between GOMOS Level 2 (blue) and lidar profiles (black) for Garmisch Partenkirchen (upper left panels), Mauna Loa observatory (upper right panels) and Dumont D'Urville identified as background aerosol (lower left panels), and as observations affected by PSC influence (lower right panels). Shown are the median profiles for the entire period (plus standard deviations), the median and inter-quartile range of the relative differences (%) between the GOMOS and the lidar data. The numbers of co-located datasets, which have been found in the respective altitude bins, are given on the right y-axis.

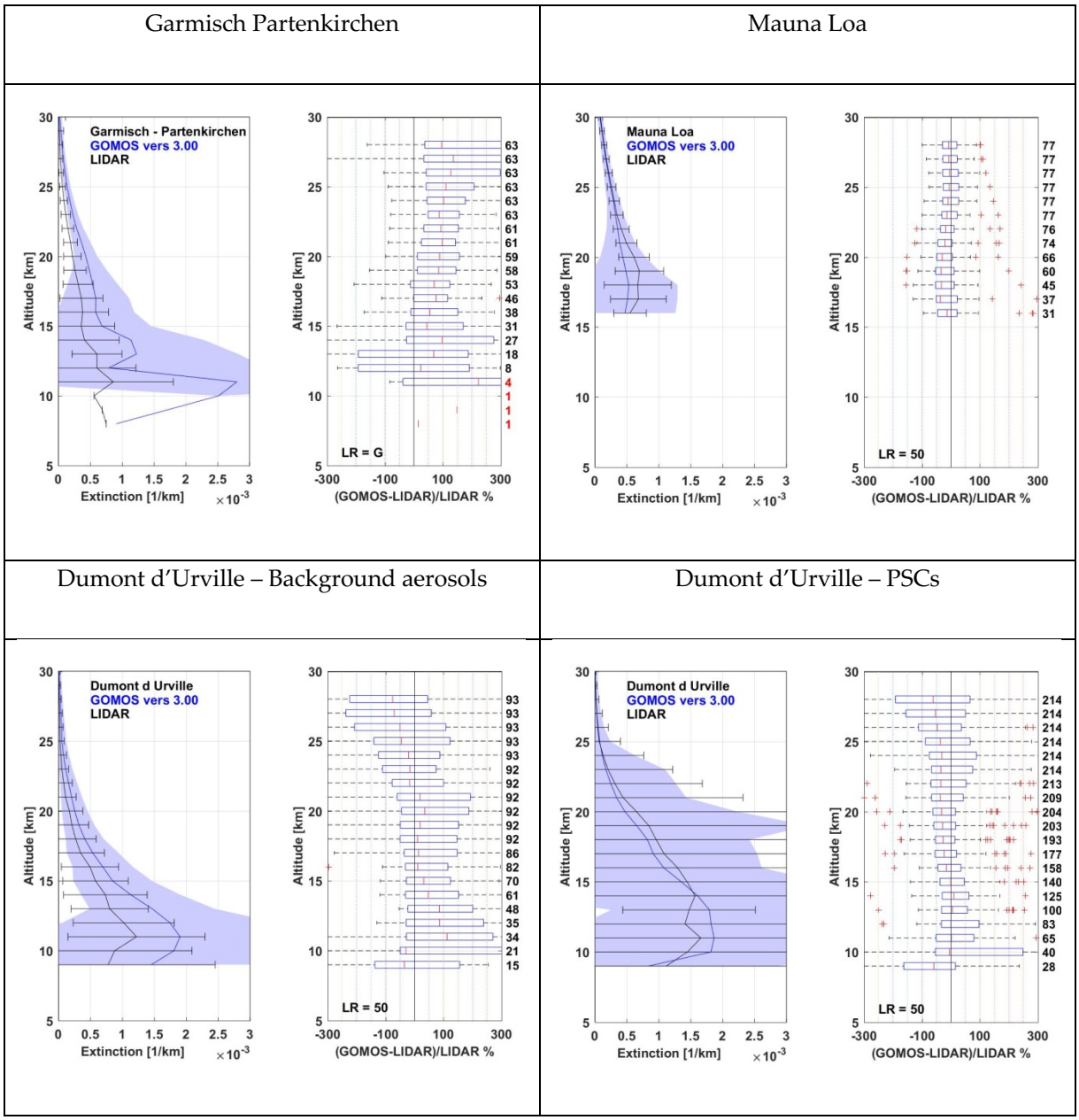
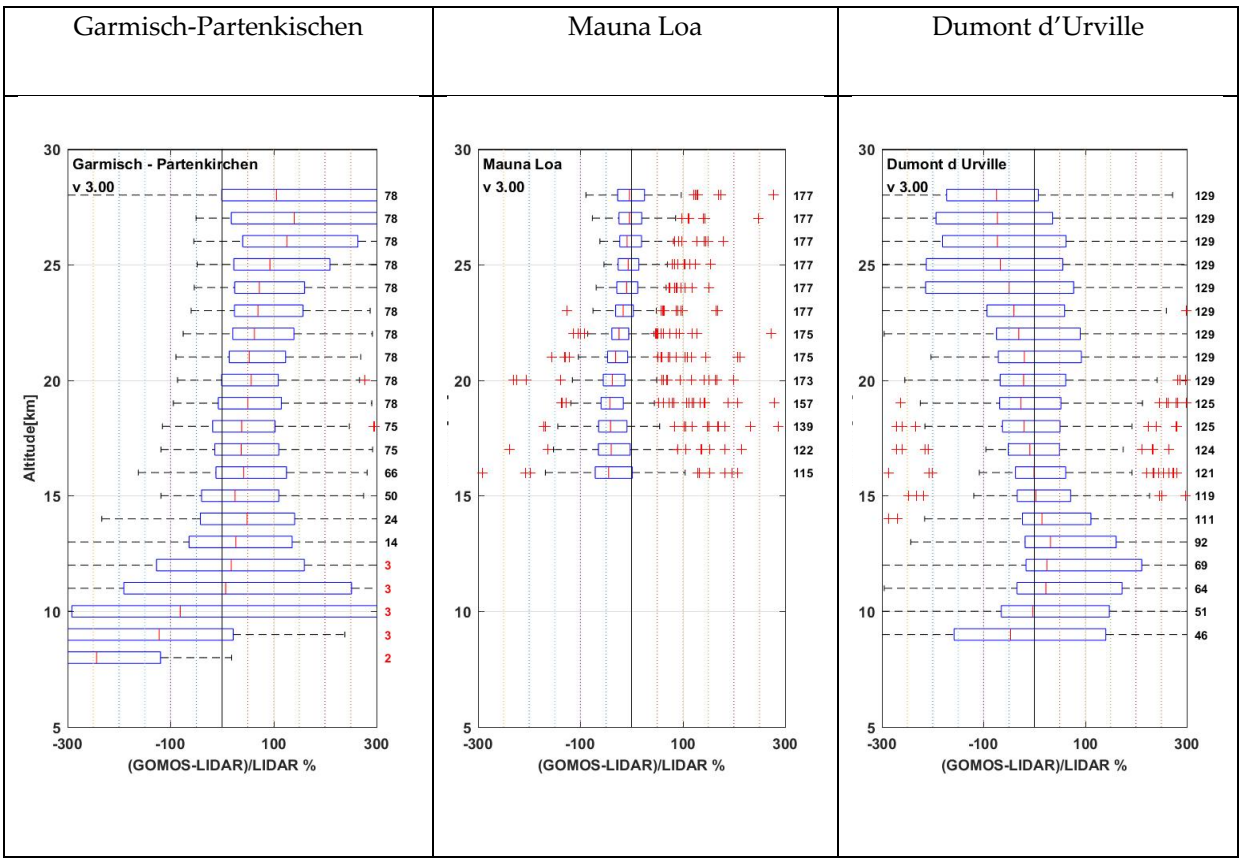


Figure 22 - Statistics for comparison between GOMOS L3 version 3.00 and NDACC lidar profiles for Garmisch-Partenkirchen (left), Mauna Loa observatory (middle) and Dumont D'Urville (right). The whisker plots show median, inter-quartile range and outliers for the relative differences (%) between the GOMOS and the lidar data.



4. Output data

4.1 Aerosol data products

The output of the algorithm consists of:

- Gridded total aerosol extinction profiles. In the current version CCI_4.00, we use the following bin size: 5-day by 60° longitude by 5° latitude by 1 km height, and extinctions are provided at 355, 440, 470, 550, and 750 nm.
- Gridded total stratospheric AOD at 355, 440, 470, 550, and 750 nm. As for the total aerosol extinction, we use 5-day mean at a fixed grid of 60° longitude by 5° latitude and stratospheric AOD are provided at 355, 440, 470, 550, and 750 nm.
- Gridded Angström exponent for the total extinction and AOD (same grid resolution as the previous associated products). In the current version CCI_4.00, the Angström exponent is computed from 355 to 750 nm.
- Tropopause height.

In the version CCI_4.00, the following fields are also provided specifically for the liquid sulfate aerosol fraction and for the PSC fraction:

- Gridded aerosol extinction profiles, provided at 355, 440, 470, 550, and 750 nm on the same grid as the total aerosol extinction.
- Gridded Angström exponent for the extinction (same grid resolution as the previous related products), computed from 355 to 750 nm.
- Gridded microphysical aerosol fields derived from the PSD: Total number density (=total number of particles per unit of air volume), effective radius, surface area density and volume density. These fields are provided on the same grid as the total aerosol extinction, but not higher than 35 km height. It has to be noted that sulfate aerosols are not expected at higher altitude.
- Gridded radiative aerosol fields derived from the PSD: asymmetry parameter, and single scattering albedo, provided at 355, 440, 470, 550, and 750 nm. These fields are provided on the same grid as the total aerosol extinction, but not higher than 35 km height.

All these products are delivered with an estimate of their uncertainty, and are provided over the whole ENVISAT period (April 2002-April 2012).

Latitude, longitude, altitude used for the total extinction field and altitude used for the PSD-derived quantities are specified.

As diagnostic fields, the number of events per bin is provided for the total extinction and stratospheric AOD, the liquid sulfate extinction and PSC extinction.

It must be noted that PSD-derived parameters are very recent. They have not been validated adequately yet, and should be considered as a scientific demonstration product.

4.2 Error budget estimates

Quantifying the aerosol extinction retrieval error is challenging. A detailed description can be found in [Tamminen et al, 2010]. Generally, the random error on a profile is determined by two contributions:

(1) The uncorrected residual scintillation component.

The error estimation of the GOMOS data product version 5 (used as source files for the first versions of the AerGOM retrieval) did not take the uncorrected residual scintillation component into account, so that retrieval errors are likely underestimated. However, the current version (v. 6.01) of these source files has been corrected with regards to this particular problem.

(2) The measurement noise which changes from one stellar source to another due to star magnitude and temperature differences;

The star magnitude basically affects the signal-to-noise ratio, with brighter stars being better than dim stars. The star temperature, on the other hand, determines the main spectral emission range, with hot stars emitting more in the UV and colder ones in the visible and near-infrared domain.

An evaluation of the AerGOM retrieval data has been carried out to look at the relative uncertainty of the extinction profiles at various wavelengths for occultations performed with different types of star. A summary of the results can be seen in Figure 23 and the detail of the analysis can be found in [Robert et al., 2016] and [Bingen et al., 2017]. From the figure, one can observe that the mean uncertainty at 550 nm and between 20-30 km vary as follows, for different star types:

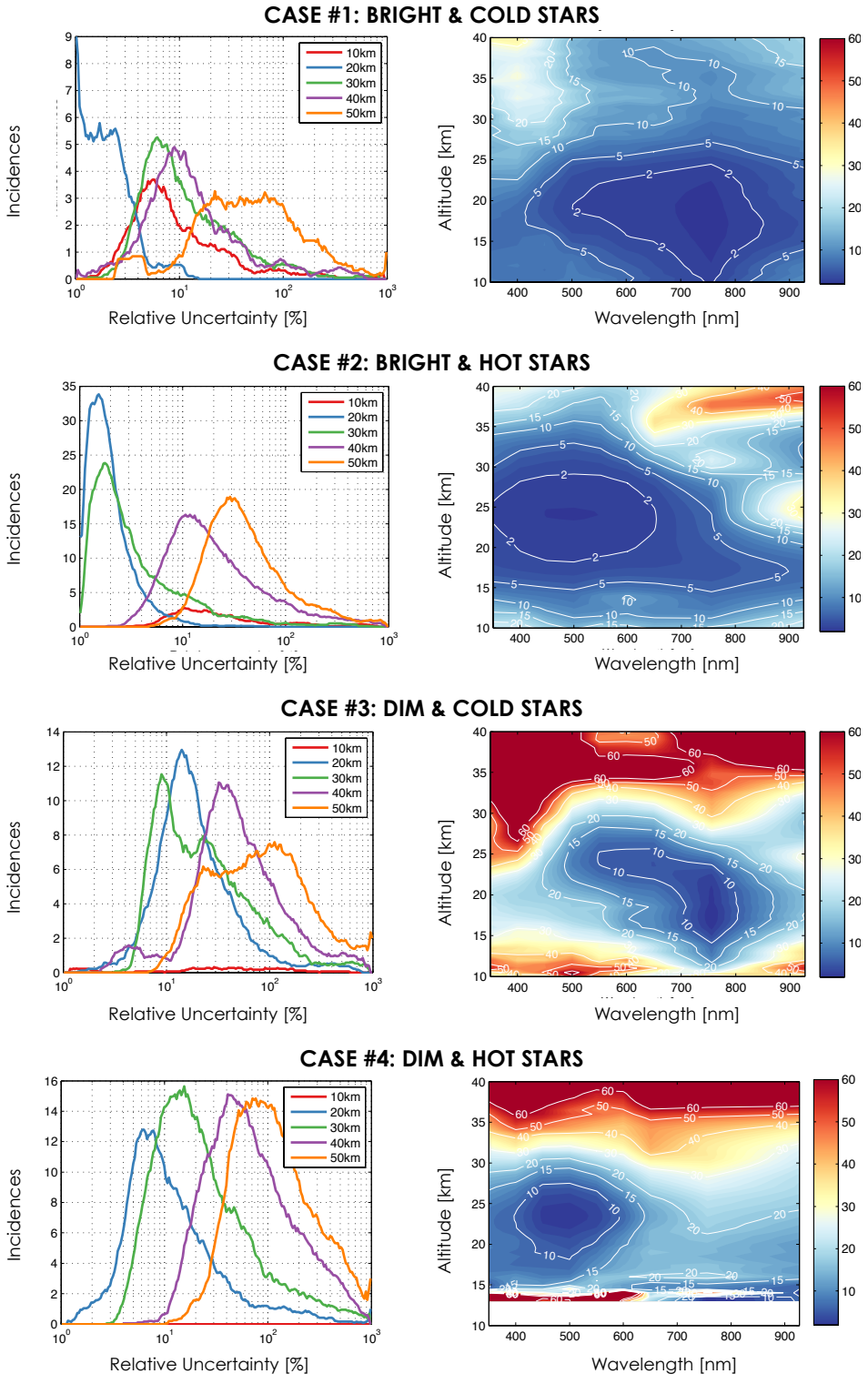
- Bright-Cold stars: 2-10%
- Bright-Hot stars: 2-5%
- Dim-Cold stars: 20-25 %
- Dim-Hot stars: 15-25 %

This analysis confirms, for the case of AerGOM, the conclusions of [Tamminen et al., 2010] that the star magnitude has indeed the largest impact on the uncertainty. Moreover, this preliminary work also confirms the influence of the star temperature on the uncertainty of the retrieval: it can be clearly seen that the minimum relative uncertainty in the contour plots changes between 750 nm and 500 nm for cold and hot stars, respectively.

In terms of systematic errors, it should be noted that possible sources are (1) A wrong aerosol spectral model, and (2) An imperfect ECMWF air density profile.

Both sources of systematic errors have been estimated in [Tamminen et al, 2010].

Figure 23 - Left: AerGOM's extinction relative uncertainty distribution for λ=550nm at various altitudes under different measurement conditions. Right: Median relative uncertainty of the extinction as a function of the wavelength and altitude.



Appendix A: The first-difference operator

When inverse problems need to be regularized, one typically tries to minimize a high-frequency component in the retrieval vector. Very often this component is estimated with the first derivative of the vector, although higher derivatives are also used. Given a function $f(x)$ the first derivative is given by:

$$f'(x) = \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \lim_{\mathrm{d}x \to 0} \frac{f(x+\mathrm{d}x) - f(x)}{\mathrm{d}x}$$

In numerical problems we work with discretized functions (vectors or arrays) **f** with components f_i for *I* = 1 … *n*. Here, the first derivative can be approximated by

$$f'(x_i) \approx \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}$$

Or in matrix notation:

$$\mathbf{f'} = \mathbf{Lf}$$

With **L** the (*n*-1) x *n* first-derivative operator

$$\mathbf{L} = \begin{bmatrix} \frac{-1}{x_2 - x_1} & \frac{1}{x_2 - x_1} & 0 & \dots & 0 & 0 & 0 \\ 0 & \frac{-1}{x_3 - x_2} & \frac{1}{x_3 - x_2} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \frac{-1}{x_{n-1} - x_{n-2}} & \frac{1}{x_{n-1} - x_{n-2}} & 0 \\ 0 & 0 & 0 & \dots & 0 & \frac{-1}{x_n - x_{n-1}} & \frac{1}{x_n - x_{n-1}} \end{bmatrix}$$

If the grid is regular (uniform spacing), then this matrix reduces to the well-known first-difference operator

$$\mathbf{L} = \frac{1}{\Delta x} \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}$$

In some cases it is preferable to have a square matrix. In that case we can always add an extra row of zeros to the upper or lower edge of the matrix.

References

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

J. L. Bertaux, G. Megie, T. Widemann, E. Chassefière, R. Pellinen, E. Kyrölä,S. Korpela, and P. Simon. Monitoring of ozone trend by stellar occultations: The GOMOS instrument. Advances in Space Research, 11 (3):3237–3242, 1991.

J.L. Bertaux, E. Kyrölä, and T. Wehr. Stellar occultation technique for atmospheric ozone monitoring: GOMOS on Envisat. Earth observation quarterly, (67): 17–20, 2000.

C. Bingen, D. Fussen, and F. Vanhellemont. A global climatology of stratospheric aerosol size distribution parameters derived from SAGE II data over the period 1984-2000: 2. Reference data. Journal of Geophysical Research, 109, D06202, doi:10.1029/2003JD003511, 2004.

C. Bingen, C. E. Robert, K. Stebel, C. Brühl, J. Schallock, F. Vanhellemont, N. Mateshvili, M. Höpfner, T. Trickl, et al. Stratospheric aerosol data records for the Climate Change Initiative: development, validation and application to Chemistry-Climate Modelling, under review for Remote Sensing of Environment, 2017.

B. A. Bodhaine, N. B. Wood, E. G. Dutton, and J. R. Slusser. On Rayleigh optical depth calculations. Journal of Atmospheric and Oceanic Technology, 16:1854–1861, 1999.

C.F. Bohren and D.R. Huffman. Absorption and Scattering of Light by Small Particles. John Wiley & Sons, Inc., New York, USA, 1993.

K. S. Carslaw, T. Peter, and S. L. Clegg. Modeling the composition of liquid stratospheric aerosols. Reviews of Geophysics, 35 (2):125–154, 1997.

R. P. Damadeo, J. M. Zawodny, L. W. Thomason, N. Lyer, SAGE version 7.0 algorithm: application to SAGE II, Atmos. Meas. Tech., 6, 3539-3561, 2013.

B. Edlén. The refractive index of air. Metrologia, 2:71–80, 1966.

D. Hanson, D. and K. Mauersberger, Laboratory studies of the nitric acid trihydrate: Implications for the south polar stratosphere. Geophysical research letters 15, 855–858, 1988.K. P. Hoinka, Statistics of the global tropopause pressure. Monthly Weather Review 126, 3303–3325, 1998.

M. D. King, D. M. Byrne, B. M. Herman, and J.A. Reagan. Aerosol size distributions obtained by inversion of spectral optical depth measurements. Journal of the Atmospheric Sciences, 35:2153–2167, 1978.

U. K. Krieger, J. C. Mössinger, B. Luo, U. Weers, and T. Peter. Measurement of the refractive indices of H2SO4-HNO3- H2O solutions to stratospheric temperatures. Applied Optics, 39(21): 3691–3703, 2000.

E. Kyrölä, J. Tamminen, G. Leppelmeier, V. Sofieva, S. Hassinen, J. Bertaux, A. Hauchecorne, F. Dalaudier, C. Cot, O. Korablev, O. Fanton d'Andon, G. Barrot, A. Mangin, B. Theodore, M. Guirlet, F. Etanchaud, P. Snoeij, R. Koopman, L. Saavedra, R. Fraisse, D. Fussen, and F. Vanhellemont. GOMOS on Envisat - an overview. Advances in Space Research, 33:1020–1028, 2004.

J. Lenoble, Atmospheric Radiative Transfer – Studies in Geophysical Optics and Remote Sensing, A. Deepak Publishing, Hampton, Virginia, USA, 1993.

J. D. Lumpe, R. M. Bevilacqua, K. W. Hoppel, and C. E. Randall. POAM III retrieval algorithm and error analysis. Journal of Geophysical Research, 107(D21), 2002.

S. K. Meilinger, T. Koop, B. P. Luo, T. Huthwelker, K. S. Carslaw, U. Krieger, P. J. Crutzen, and T. Peter. Size-dependent stratospheric droplet composition in leewave temperature fluctuations and their potential role in PSC freezing. Geophysical Research Letters, 22:3031–3034, 1995.

NIMA. World geodetic system 1984, its definition and relationships with local geodetic systems. Technical Report TR8350.2, Department of Defense, 1984.

J. Orphal, C. E. Fellows, P.-M. Flaud, The visible absorption spectrum of NO3 measured by high-resolution Fourier transform spectroscopy, J. Geophys. Res, 108(D3), 4077, doi:10.1029/2002JD002489., 2003.

E. R. Peck and K. Reeder. Dispersion of air. Journal of the Optical Society of America, 62:958–962, 1972.

T. Popp, G. de Leeuw, C. Bingen, C. Brühl, V. Capelle, A. Chedin, A.; L. Clarisse, O. Dubovik, R. Grainger, J. Griesfeller, et al. Development, Production and Evaluation of Aerosol Climate Data Records from European Satellite Observations (Aerosol_cci), Remote Sens. 8, 421; doi:10.3390/rs8050421, 2016.

W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. Numerical Recipes in FORTRAN - The Art of scientific computing. Cambridge University Press, second edition, 1992.

H. R. Pruppacher and J. D. Klett, Microphysics of Clouds and Precipitation, Kluwers Academic Publishers, 2nd edition 1997.

R. J. Purser and H.-L. Huang, Estimating effective data density in a satellite retrieval or an objective analysis, J. Appl. Meteor., 32, 1092-1107, 1992.

C. E. Robert, C. Bingen, F. Vanhellemont, N. Mateshvili, E. Dekemper, C. Tétard, D. Fussen, C. Zehner, L. W. Thomason, C. T. McElroy, et al. AerGOM, an improved algorithm for stratospheric aerosol extinction retrieval from GOMOS observations Part 2: Intercomparisons, Atmos. Meas. Tech,. 2016, 9, 4701-4718, doi:10.5194/amt-9-4701-2016.

C.D. Rodgers. Inverse Methods for Atmospheric Sounding - Theory and Practice. World Scientific, first edition, 2000.

S. P. Sander, Temperature dependence of the NO3 absorption spectrum, J. Phys. Chem., 90, 4135–4142., 1986.

R. W. Saunders, S. Dhomse, W. S. Tian, M. P. Chipperfield, and J. M. C. Plane, Interactions of meteoric smoke particles with sulphuric acid in the Earth's stratosphere, Atmos. Chem. Phys., 12, 4387-4398, doi:10.5194/acp-12-4387-2012, 2012.

A. Serdyuchenko, V. Gorshelev, M. Weber, W. Chehade, J. P. Burrows, 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.

V. Sofieva, J. Tamminen, H. Haario, E. Kyrölä, and M. Lehtinen, Ozone profil smoothness as a priori information in the inversion of limb measurements, Annales Geophysicae, 22, 3411-3420, 2004.

J. Tamminen, E. Kyrölä, V. F. Sofieva, M. Laine, J.-L. Bertaux, A. Hauchecorne, F. Dalaudier, D. Fussen, F. Vanhellemont, O. Fanton-d'Andon, G. Barrot, A. Mangin, M. Guirlet, L. Blanot, T. Fehr, L. Saavedra de Miguel, and R. Fraisse, GOMOS data characterisation and error estimation. Atmos. Chem. Phys 10, 9505–9519, 2010.

L.W. Thomason, L.R. Poole, and C.E. Randall. SAGE III aerosol extinction validation in the Arctic winter: comparisons with SAGE II and POAM III. Atmospheric Chemistry and Physics, 7:1423–1433, 2007.

S. Twomey. Comparison of constrained linear inversion and an iterative nonlinear algorithm applied to the indirect estimation of particle size distributions. Journal of Computational Physics, 18:188–200, 1975.

A. C. Vandaele, C. Hermans, P.C. Simon, M. Carleer, R. Colin, S. Fally, M.-F. Mérienne, A. Jenouvrier, B. Coquart, Measurements of the NO2 absorption cross-section from 42 000 cm−1 to 10 000 cm−1 (238–1000 nm) at 220 K and 294 K, JQSRT, 59, 171–84, 1998.

A. C. Vandaele, C. Hermans, P.C. Simon, M. Carleer, R. Colin, S. Fally, M.-F. Mérienne, A. Jenouvrier, B. Coquart. High-resolution Fourier transform measurement of the NO2 visible and near-infrared absorption cross-section: temperature and pressure effects. J. Geophys. Res, 107(D18), 4348, doi:10.1029/2001JD000971, 2002.

A. C. Vandaele, C. Hermans, S. Fally, M. Carleer, M-F. Mérienne, A. Jenouvrier, B. Coquart, R. Colin, Absorption cross-section of NO2: simulation of temperature and pressure effects, JQSRT, 76, 373-391, 2003.

H.C. van de Hulst, Light Scattering by Small Particles, Dover Publications, New York, USA, 1981.

F. Vanhellemont and D. Fussen. Derivation of stratospheric aerosol properties from the ORA experiment. Advances in Space Research, 29 (11):1741–1746, 2002.

F. Vanhellemont, N. Mateshvili, L. Blanot, C. E. Robert, C. Bingen, C. Tétard, D. Fussen, E. Dekemper, V. Sofieva, E. Kyrölä, et al. AerGOM, an improved algorithm for stratospheric aerosol extinction retrieval from GOMOS observations Part 1: Algorithm development, Atmos. Meas. Tech,. 2016, 9, 4687-4700, doi:10.5194/amt-9-4687-2016.

S. G. Warren. Optical constants of ice from the ultraviolet to the microwave. Applied Optics, 23(8):1206–1225, 1984.

S. Woods, R. Lawson, E. Jensen, T. Bul, T. Thornberry, A. Rollins, L. Pfister, and M. Avery, Microphysical Properties of Tropical Tropopause Layer Cirrus, J. Geophys. Res, 123, https://doi.org/10.1029/2017JD028068, 2017.

R. J. Yokelson, J. B. Burkholder, R. W. Fox, R. K. Talukdar, A. R. Ravishankara, Temperature dependence of the NO3 absorption spectrum, J. Phys. Chem., 1994, 98, 13.144– 13.150.

Copernicus Climate Change Service

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: info@copernicus-climate.eu

climate.copernicus.eu copernicus.eu ecmwf.int